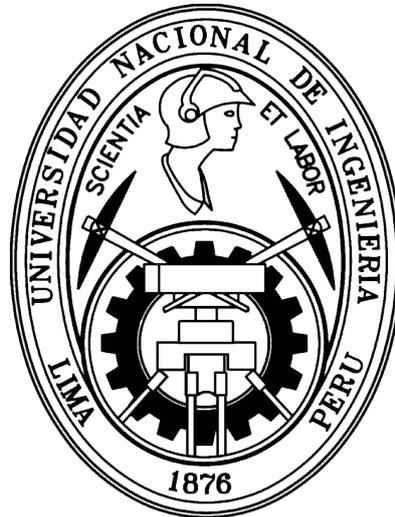


**UNIVERSIDAD NACIONAL DE INGENIERÍA**

**FACULTAD DE CIENCIAS**



TESIS

**EL PROBLEMA DE EQUILIBRIO DE NASH  
GENERALIZADO**

PARA OBTENER EL TÍTULO PROFESIONAL DE  
**LICENCIADO EN MATEMÁTICA**

ELABORADA POR  
**CESAR AUGUSTO HILARIO POMA**

ASESOR  
**WILLIAM CARLOS ECHEGARAY CASTILLO**

**LIMA - PERÚ**

**2019**

## Resumen

Los equilibrios de Nash constituyen un modelo matemático que desde su creación en los años 50 ha encontrado múltiples aplicaciones en varias áreas del conocimiento.

La idea intuitiva de equilibrio es la siguiente. Varios *agentes* o *jugadores* interactúan entre sí intentando maximizar sus beneficios mediante el uso de una estrategia. Los beneficios de cada agente dependen de su estrategia propia y de las de sus rivales. En todo momento los jugadores conocen las estrategias de sus contrincantes y, sobre la base de ello, cada agente escoge aquella estrategia que le reporte los mayores beneficios. Se habrá alcanzado un *equilibrio* cuando cada jugador emplee precisamente aquella estrategia que le proporcione más beneficios. El Problema de Equilibrio de Nash Generalizado (GNEP) formaliza matemáticamente esto que acabamos de describir. A manera de ejemplo, podemos pensar en varios bancos que compiten entre sí en un determinado país.

Alcanzar el equilibrio no es fácil. En este sentido, las reformulaciones del GNEP son útiles porque permiten analizar el problema desde nuevas perspectivas, utilizando herramientas de otras áreas de la matemática. Ahora bien, la teoría de las desigualdades variacionales está muy desarrollada, tanto teórica como algorítmicamente (véase [13]). Igualmente, los problemas de optimización (maximización o minimización de funciones) han sido estudiados exhaustivamente, ya desde la época de Newton, y por ello también existe una teoría muy desarrollada al respecto (véase [18]). De manera que las reformulaciones del GNEP como problemas de optimización y como problemas de desigualdad variacional son especialmente importantes.

En esta tesis estudiamos el GNEP y presentamos varias reformulaciones del GNEP como problemas de optimización y como problemas de desigualdad variacional. Las herramientas básicas para obtener las reformulaciones como problemas de optimización serán la función de Nikaidô-Isoda y su regularización, mientras que para obtener las reformulaciones como problemas de desigualdad variacional utilizaremos las nociones clásicas de gradientes y subdiferenciales, y la recientemente introducida noción de subniveles ajustados y sus operadores normales (véase [4]).

*A mi padre, a quien le debo todo.*

## Agradecimientos

Deseo agradecer ante todo a mi familia, por su apoyo constante e incondicional. Agradezco también al profesor William Echegaray por su ayuda en la realización de esta tesis. También deseo expresar un agradecimiento especial al profesor John Cotrina, quien me introdujo en la temática de los equilibrios de Nash.

# Índice general

<b>Antecedentes, objetivos, hipótesis y marco teórico</b>	<b>1</b>
<b>Introducción</b>	<b>3</b>
<b>1. Preliminares y resultados previos</b>	<b>7</b>
1. Conceptos básicos . . . . .	7
2. Análisis convexo . . . . .	18
2.1. Conjuntos convexos y conos . . . . .	19
2.2. Funciones convexas . . . . .	29
2.3. Subdiferenciales . . . . .	35
2.4. Funciones pseudoconvexas . . . . .	40
3. Problemas de optimización . . . . .	40
Condiciones KKT . . . . .	42
4. Análisis multivaluado . . . . .	46
El teorema de von Neumann . . . . .	55
5. Análisis cuasiconvexo . . . . .	58
5.1. Funciones cuasiconvexas . . . . .	58
5.2. Subniveles ajustados y operadores normales . . . . .	59
6. Funciones definidas a partir de supremos . . . . .	64
7. Desigualdades variacionales . . . . .	68
7.1. El problema de desigualdad variacional . . . . .	68
Condiciones KKT . . . . .	69
Existencia de solución . . . . .	71
7.2. El problema de desigualdad cuasivariacional . . . . .	73
8. Condiciones de optimalidad . . . . .	74
8.1. Caso convexo y diferenciable . . . . .	74
8.2. Caso convexo . . . . .	75
8.3. Caso cuasiconvexo . . . . .	76
<b>2. Equilibrios de Nash</b>	<b>81</b>
9. El NEP . . . . .	81
10. El GNEP . . . . .	86
<b>3. Reformulaciones como problemas de desigualdad variacional</b>	<b>95</b>
11. Caso convexo y diferenciable . . . . .	96
Condiciones KKT . . . . .	100
12. Caso convexo . . . . .	103

13. Caso cuasiconvexo . . . . .	107
<b>4. Reformulaciones como problemas de optimización</b>	<b>114</b>
14. Primeras reformulaciones . . . . .	115
15. Reformulación diferenciable y con restricciones . . . . .	122
Puntos estacionarios . . . . .	126
16. Reformulación diferenciable y sin restricciones . . . . .	130
Puntos estacionarios . . . . .	132
<b>5. Conclusiones y recomendaciones</b>	<b>135</b>

# Antecedentes, objetivos, hipótesis y marco teórico

## Antecedentes y Planteamiento del Problema

El Problema de Equilibrio de Nash Generalizado (GNEP) puede describirse informalmente de la siguiente manera. Supongamos que estamos viendo un partido de fútbol. Los dos equipos interactúan e intentan lograr la victoria marcando goles. Cada uno utiliza una estrategia, la forma en que juegan para derrotar a su rival, que está sujeta a la estrategia de dicho rival. Por ejemplo, si el equipo A emplea una estrategia agresiva, entonces la estrategia del equipo B podría consistir en esperar a que el equipo A ataque, para luego realizar rápidos contraataques cuando la mayoría de los jugadores del equipo A estén en su área. Puede llegarse a un *equilibrio* cuando la siguiente situación ocurra: ninguno de los equipos ataca por temor a sufrir el contraataque del equipo rival.

El Problema de Equilibrio de Nash (NEP, por sus siglas en inglés) es un modelo matemático que sirve para modelar problemas de la vida real como el que acabamos de describir. Introducido por John Nash [26, 27] en 1950, el NEP encontró numerosas aplicaciones, convirtiéndose en un tema de investigación importante dentro de la Economía Matemática. Su versión generalizada (el Problema de Equilibrio de Nash Generalizado o GNEP) fue introducida por Debreu [10] en un marco casi estrictamente económico (con bienes y servicios), y fue estudiado por diversos investigadores de áreas no necesariamente afines. De ahí que el GNEP haya tenido varias denominaciones a lo largo de su historia (*economía abstracta*, *problema de equilibrio social*, *pseudojuego*). Al igual que el NEP, pronto el GNEP mostró cuán útil era para lidiar con diversos tipos de problemas, convirtiéndose también en un tema de amplio interés.

El problema principal que se presenta al tratar con los equilibrios de Nash es decidir si existen soluciones para el GNEP o el NEP, y, en el caso afirmativo, cómo encontrarlas. En la segunda mitad del siglo XX se dieron grandes avances a este respecto. El NEP o el GNEP es difícil de analizar directamente, y lo que a menudo se hizo (y se hace) es buscar reformulaciones del problema. Por ejemplo, John Nash [26] utilizó una reformulación del NEP para mostrar que este admite soluciones en ciertos casos (lo que hizo, en realidad, fue probar que la reformulación admite soluciones). Así, las reformulaciones de los equilibrios de Nash son importantes porque nos ayudan a analizarlos desde nuevos puntos de vista y con nuevas herramientas. Desde luego, el tipo de reformulación obtenida depende de las características del pro-

blema en cuestión. Y, lógicamente, mientras más restricciones imponamos menos reformulaciones tendremos.

Estos son los problemas que nos planteamos en esta tesis. Primero, queremos comprender a fondo el GNEP. Y segundo, deseamos ver qué reformulaciones existen y qué información del GNEP podemos obtener a partir de ellas.

## Objetivos

Esta tesis tiene dos objetivos. El primero es estudiar en detalle el GNEP. El segundo es dar cuenta de las reformulaciones del GNEP. En consecuencia, además de dar una breve descripción de los equilibrios de Nash, vamos a presentar dos tipos de reformulaciones del GNEP: como problemas de desigualdad variacional y como problemas de optimización.

## Hipótesis

Existe una amplia bibliografía sobre el GNEP (véase [12] y las referencias allí citadas). De modo que no hay mucho que decir respecto de nuestro primer objetivo (estudiar el GNEP).

Ahora bien, como ya dijimos, en [26] John Nash usó una reformulación del NEP como problema de desigualdad variacional para mostrar que posee soluciones en cierto casos. Por consiguiente, aquí nuestra hipótesis natural es que existen reformulaciones de esta clase, ya no del NEP sino del GNEP.

Igualmente, como el GNEP consiste en resolver varios problemas de optimización al mismo tiempo (véase la Introducción o el Capítulo 2), se espera que existan reformulaciones del GNEP como problemas de optimización. Esta es precisamente nuestra hipótesis en este caso.

## Marco teórico

Nuestro estudio del GNEP (Capítulo 2) y la obtención de sus reformulaciones (Capítulos 3 y 4) van a estar basadas en gran medida en el análisis convexo y cuasiconvexo, así como en la teoría de las desigualdades variacionales. Consecuentemente, en el Capítulo 1 realizamos una revisión general de varios conceptos y teoremas conocidos, y probamos varios resultados sofisticados (véase por ejemplo la Sección 6). Este será nuestro marco teórico, es decir, nos servirá como soporte para el posterior estudio del GNEP, y para conseguir las reformulaciones deseadas.

# Introducción

Los equilibrios de Nash constituyen un modelo matemático en la Teoría de Juegos, que ha encontrado múltiples aplicaciones en las últimas décadas. Informalmente hablando, dicho modelo consiste en  $N$  jugadores que interactúan y que, haciendo uso de una estrategia, intentan maximizar sus beneficios (o minimizar sus pérdidas). Bajo el supuesto de que cada jugador conoce las estrategias de sus rivales, se habrá alcanzado un *equilibrio* cuando ninguno de ellos pueda aumentar sus beneficios (o reducir sus pérdidas) cambiando unilateralmente de estrategia.

Si el *juego* se tratara de un partido de fútbol, los jugadores serían los directores técnicos y las estrategias serían la manera como hacen que sus futbolistas se desempeñen durante el partido. Cada jugador “ve” (en el campo de juego) la estrategia de su contrincante, a partir de la cual decide qué hacer. Por ejemplo, si uno de los equipos está ganando y comienza a jugar defensivamente, el director técnico del otro podría intentar mejorar su situación haciendo ingresar a un delantero en reemplazo de un defensa, pues es menos probable que su rival lo ataque. Así, la estrategia de cada equipo depende de la estrategia de su contrincante. Y cada cambio de estrategia de cualquiera de los jugadores puede inducir un cambio de estrategia en el otro, hasta que, posiblemente, se alcance el equilibrio. Una situación de equilibrio podría ser aquella en que ninguno de los equipos se atreve a atacar por miedo a quedar descompensado en defensa y sufrir los contraataques del equipo rival.

Hagamos una descripción más precisa. Supongamos que  $n, n_1, \dots, n_N$  son enteros positivos tales que  $n = n_1 + \dots + n_N$ , de modo que podamos escribir

$$\mathbb{R}^n = \mathbb{R}^{n_1} \oplus \dots \oplus \mathbb{R}^{n_N}.$$

Cada jugador  $\nu$  controla la variable

$$x^\nu \in \mathbb{R}^{n_\nu},$$

conocida como su *estrategia* en el juego, y posee una *función de pérdidas* (o *función de coste*)

$$\theta_\nu : \mathbb{R}^n \rightarrow \mathbb{R}$$

cuyos valores dependen de las estrategias  $x^\mu$  de *todos* los jugadores. Dado un vector

$$x = (x^1, \dots, x^N) \in \mathbb{R}^n$$

de estrategias del juego, usamos el símbolo  $x^{-\nu}$  para denotar las estrategias de todos los jugadores excepto la del  $\nu$ -ésimo jugador, es decir

$$x^{-\nu} = (x^1, \dots, x^{\nu-1}, x^{\nu+1}, \dots, x^N) \in \mathbb{R}^{n-n_\nu},$$

y escribimos  $x = (x^\nu, x^{-\nu})$ <sup>1</sup>. El objetivo del  $\nu$ -ésimo jugador es el siguiente: dado un conjunto de estrategias empleadas por sus rivales, él debe minimizar su función de pérdidas haciendo variar su estrategia en un conjunto

$$X_\nu \subset \mathbb{R}^{n_\nu},$$

llamado *conjunto de estrategias factible* o *región factible* del  $\nu$ -ésimo jugador. Esto es, para cada  $x^{-\nu}$  fijo él intenta resolver el siguiente problema de minimización

$$\min_{x^\nu} \theta_\nu(x^\nu, x^{-\nu}), \quad x^\nu \in X_\nu. \quad (0.1)$$

El *Problema de Equilibrio de Nash* (NEP, por sus siglas en inglés) consiste justamente en encontrar aquellos puntos en que cada jugador ha logrado su objetivo, es decir, vectores  $\bar{x} \in \mathbb{R}^n$  con las siguientes propiedades:  $\bar{x}^\nu \in X_\nu$  para cada  $\nu$ , y

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu}) \text{ para cualesquiera } \nu \text{ y } x^\nu \in X_\nu.$$

Si estos puntos existen, son llamados *equilibrios* de Nash o *soluciones* del NEP. Notemos que el equilibrio se habrá alcanzado cuando a ninguno de los jugadores le convenga cambiar de estrategia unilateralmente<sup>2</sup>.

Tras ser introducido en 1950 por John Nash [26, 27] en su tesis de doctorado, el NEP se convirtió rápidamente en una herramienta muy útil en economía y en varias áreas relacionadas con las matemáticas, lo cual a su vez impulsó el desarrollo de una teoría amplia y fructífera. Sin embargo, pronto se hizo patente que en los “juegos de la vida real” los jugadores también interactúan a nivel de sus estrategias; o sea que las posibles estrategias que un jugador pueda elegir en un determinado momento *dependen* de las estrategias rivales. Nuestro ejemplo informal de fútbol ilustra este punto: cuando un equipo ataca, lo más probable es que el otro defienda o “haga tiempo”, pero no que ataque. Consecuentemente, la falta de una generalización adecuada del NEP que sirviera para lidiar con este tipo de situaciones quedó en evidencia, y el *Problema de Equilibrio de Nash Generalizado* (GNEP por sus siglas en inglés), introducido por Debreu [10], resultó ser tal generalización. El GNEP es un NEP que incorpora la condición ya mencionada: cada región factible  $X_\nu$  no permanece fija, sino que depende de las estrategias rivales  $x^{-\nu}$ , por lo que se la representa mediante  $X_\nu(x^{-\nu})$ . Así, el problema de minimización (0.1) del  $\nu$ -ésimo jugador se convierte en

$$\min_{x^\nu} \theta_\nu(x^\nu, x^{-\nu}), \quad x^\nu \in X_\nu(x^{-\nu}), \quad (0.2)$$

y los equilibrios correspondientes son llamados *equilibrios generalizados* de Nash o *soluciones* del GNEP.

En las décadas que siguieron a la publicación del artículo de Debreu, se encontraron numerosas aplicaciones del GNEP en múltiples áreas de las Matemáticas

<sup>1</sup> La igualdad  $x = (x^\nu, x^{-\nu})$  no significa que las componentes de  $x$  hayan sido reordenadas de modo que  $x^\nu$  sea la primera componente. Se trata tan solo de enfatizar la componente  $x^\nu$  de  $x$ .

<sup>2</sup> Recalamos que en el NEP (y también en el GNEP) cada jugador conoce las estrategias de sus rivales y, sobre la base de ese conocimiento, decide conservar o modificar la suya, siempre pensando en minimizar sus pérdidas. Así, los partidos de fútbol y de ajedrez son ejemplos de NEPs, mientras que las partidas de póquer no lo son.

Aplicadas. A pesar de lo cual el GNEP permaneció más vinculado a la Economía Matemática, debido a su origen económico. Puesto que en economía las funciones convexas (o cóncavas) y cuasiconvexas (o cuasicóncavas) surgen naturalmente, los GNEPs cuyas funciones de coste satisfacen estas hipótesis han sido históricamente importantes<sup>3</sup>. En general, mientras más condiciones verifique el GNEP analizado, más podremos afirmar sobre él.

La interacción entre los jugadores del GNEP es compleja. En efecto, como los valores de cada función de pérdidas dependen de *todas* las estrategias, si en un determinado momento un jugador cambiara la suya, entonces las estrategias óptimas que escogerían sus rivales para minimizar sus respectivas funciones de coste podrían ser diferentes de las que están empleando en ese momento. O sea que cada cambio de estrategia va a inducir, posiblemente, cambios de estrategias en los demás jugadores, los que a su vez van a provocar nuevos cambios de estrategias. Por ello, decidir cuándo un GNEP específico admite soluciones es en general difícil. Más difícil aun es determinar los equilibrios explícitamente (sabiendo de antemano que existen). En este sentido, las reformulaciones del GNEP (y de cualquier otro problema matemático) son útiles porque proporcionan nuevas perspectivas del problema original, haciendo posible utilizar métodos de otras áreas de las matemáticas para abordarlo. Uno puede, por ejemplo, mostrar que un GNEP admite soluciones probando que alguna de sus reformulaciones las admite<sup>4</sup>.

En esta tesis hacemos un estudio más o menos detallado del GNEP (Capítulo 2) y presentamos varias reformulaciones del GNEP como problemas de desigualdad variacional (Capítulo 3) y como problemas de optimización (Capítulo 4). Con ese fin, nos enfocamos en una clase especial de GNEPs, a saber, aquellos para los que las regiones factibles  $X_\nu(x^{-\nu})$  están dadas por

$$X_\nu(x^{-\nu}) = \{x^\nu \in \mathbb{R}^{n_\nu} \mid (x^\nu, x^{-\nu}) \in X\},$$

donde  $X$  es un subconjunto no vacío (posiblemente convexo o cerrado) de  $\mathbb{R}^n$ . Estos GNEPs, introducidos por Rosen [29], han sido intensamente estudiados en las últimas décadas, y han alcanzado tal trascendencia que con frecuencia se les confunde con los GNEPs originales.

La motivación del Capítulo 3 es simple. Como la teoría de las Desigualdades Variacionales es amplia (véase por ejemplo [13]), existen muchas más herramientas para analizar reformulaciones del GNEP como problemas de desigualdad variacional que para estudiar el GNEP directamente. Por ejemplo, la mayoría de los resultados de existencia (teoremas que garantizan que un GNEP con determinadas propiedades posee solución) han sido obtenidos aplicando teoremas de punto fijo (véase [14]) a tales reformulaciones<sup>5</sup>. En este capítulo presentamos varias reformulaciones de este tipo, bajo hipótesis cada vez más débiles. Suponiendo primero que las funciones

<sup>3</sup>En [26], además de definir el NEP, John Nash probó que el NEP admite solución bajo hipótesis razonables, incluyendo la convexidad de las funciones de coste (véase el Teorema 9.4).

<sup>4</sup>La prueba original del teorema de John Nash, que no aparece en [26], sigue este patrón: aplicando el teorema del punto fijo de Brouwer a una reformulación de un cierto NEP, él mostró que este admite soluciones.

<sup>5</sup>Los resultados de existencia de solución para el GNEP que vamos a probar (véanse por ejemplo los Teoremas 10.6, 10.8 o 13.6) estarán basados en el teorema de punto fijo de Brouwer (Teorema 4.18) y en el teorema de Kakutani(-Fan) (Teoremas 4.19 y 4.20).

objetivo  $\theta_\nu$  son convexas<sup>6</sup> y diferenciables, obtenemos una reformulación del GNEP como un problema de desigualdad variacional, el cual estará dado en función de los gradientes de los  $\theta_\nu$ 's, para luego analizar sus condiciones KKT y relacionarlas con las condiciones KKT de los problemas de minimización (0.2) asociados al GNEP. Después generalizamos esta reformulación al caso en que las funciones  $\theta_\nu$  son solo convexas<sup>7</sup>, haciendo uso de sus subdiferenciales (en vez de sus gradientes). Presentamos luego una reformulación todavía más general, que vale tan solo bajo la hipótesis de cuasiconvexidad<sup>8</sup> de las funciones  $\theta_\nu$ , para lo cual empleamos una noción recientemente introducida por D. Aussel y N. Hadjisavvas [4] en el análisis cuasiconvexo, a saber, la de los Subniveles Ajustados y sus Operadores Normales. En cada caso, la idea básica será utilizar las condiciones de optimalidad del análisis convexo y cuasiconvexo (véase la Sección 8) para traducir los problemas de minimización (0.2) del GNEP en un problema de desigualdad variacional.

Por otra parte, como el GNEP consiste básicamente en resolver varios problemas de minimización al mismo tiempo, es natural que existan reformulaciones del GNEP como problemas de optimización (o minimización). En el Capítulo 4 utilizamos la función de Nikaidô-Isoda y su versión regularizada para transformar el GNEP en varios problemas de optimización, con restricciones y sin restricciones. Si bien las primeras reformulaciones caracterizan completamente a los equilibrios generalizados, ellas tienen el defecto de que sus funciones objetivo no son diferenciables (cuando las funciones de coste  $\theta_\nu$  lo son). Superamos este problema trabajando con una subclase de los equilibrios generalizados, a saber, los *equilibrios normalizados*. Modificando el método usado en las primeras reformulaciones, obtendremos caracterizaciones de tales equilibrios normalizados como las soluciones de problemas de optimización con función objetivo *diferenciable*, con restricciones y sin restricciones. Aunque estas últimas reformulaciones no caracterizan a todas las soluciones del GNEP, el que sus funciones objetivo sean diferenciables es una gran ventaja, ya que, como se sabe, la mayoría de algoritmos usados para resolver problemas de minimización funcionan mejor cuando su función objetivo es diferenciable.

Terminamos indicando nuestras fuentes principales. La referencia principal utilizada al presentar los equilibrios de Nash en el Capítulo 2 ha sido el artículo [12]. El Capítulo 3 está basado en varias fuentes: la reformulación de la Sección 11 es una generalización natural de un resultado clásico sobre NEPs (véase por ejemplo [13, Proposition 1.4.2]), la de la Sección 12 proviene de la referencia [17], la de la Sección 13 del artículo [3], y la parte sobre condiciones KKT en la Sección 11 proviene del artículo [11]. Las reformulaciones del Capítulo 4 están basadas en [16].

---

<sup>6</sup>Más precisamente, las funciones  $\theta_\nu$  deben ser convexas *respecto a la  $\nu$ -ésima variable*, esto es, para cada  $\nu$  y  $x^{-\nu}$  la aplicación  $x^\nu \mapsto \theta_\nu(x^\nu, x^{-\nu})$  debe ser convexa.

<sup>7</sup>De nuevo, las funciones  $\theta_\nu$  deben ser convexas respecto a la  $\nu$ -ésima variable.

<sup>8</sup>Las funciones  $\theta_\nu$  deben ser cuasiconvexas respecto a la  $\nu$ -ésima variable.

# Capítulo 1

## Preliminares y resultados previos

En este capítulo presentamos y establecemos todos los conceptos y resultados que serán necesarios para el desarrollo del presente trabajo. Comenzando con las definiciones esenciales del álgebra lineal y del análisis real, y pasando por el análisis convexo y el análisis multivaluado, nuestro objetivo es proporcionar todos los requisitos para el estudio de los equilibrios de Nash en los siguientes capítulos. Debemos indicar, sin embargo, que el teorema de punto fijo de Brouwer, resultado fundamental para establecer resultados de existencia de solución de desigualdades variacionales, no será mostrado, debido a que su prueba requiere de herramientas que están más allá del alcance y del objetivo de esta tesis, no obstante que existan en la literatura varias demostraciones de dicho teorema<sup>1</sup>.

### 1. Conceptos básicos

En esta sección revisamos los principales conceptos y resultados del análisis real y del álgebra lineal que precisaremos luego. Como se trata de material estándar, la mayoría de las proposiciones y teoremas serán enunciados sin demostración, si bien algunos resultados relevantes serán probados. Las referencias que el lector puede consultar sobre el contenido de esta sección son numerosas, por ejemplo [23], [19], [30] o [25].

#### Espacios vectoriales

Un **espacio vectorial (real)** es un conjunto  $V$  en el que están definidas dos operaciones: una adición  $+$  :  $V \times V \rightarrow V$ , que asocia a cada par  $(x, y)$  de elementos de  $V$  otro elemento de  $V$  denotado por  $x + y$ ; y una multiplicación por escalares  $\cdot$  :  $\mathbb{R} \times V \rightarrow V$ , que asocia a cada  $(a, x) \in \mathbb{R} \times V$  un elemento de  $V$  denotado por  $a \cdot x$  o simplemente  $ax$ . Estas operaciones cumplen los siguientes axiomas

- (1)  $x + y = y + x$  para todo  $x, y \in V$ ;
- (2)  $(x + y) + z = x + (y + z)$  para todo  $x, y, z \in V$ ;
- (3) existe un elemento  $0$  en  $V$  tal que  $x + 0 = x$  para todo  $x \in V$ ;

---

<sup>1</sup>Véanse por ejemplo las referencias [8, Theorem 8.1.3], [24, Teorema 11, pág. 447] o [14, Theorem (7.2), pág. 95].

- (4) para todo  $x \in V$  existe un elemento  $-x$  en  $V$  tal que  $x + (-x) = 0$ ;
- (5)  $(a + b)x = ax + bx$  y  $a(x + y) = ax + ay$  para todo  $a, b \in \mathbb{R}$  y todo  $x, y \in V$ ;
- (6)  $1x = x$  para todo  $x \in V$ ;
- (7)  $a(bx) = (ab)x$  para todo  $a, b \in \mathbb{R}$  y todo  $x \in V$ .

**Ejemplo 1.1.** Sea  $n$  un número natural. Denotamos por  $\mathbb{R}^n$  al conjunto de las  $n$ -uplas  $(x_1, \dots, x_n)$  donde  $x_i \in \mathbb{R}$  para cada  $i \in \{1, \dots, n\}$ . Este conjunto es un espacio vectorial real, donde las operaciones  $+$  :  $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  y  $\cdot$  :  $\mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  están definidas así: si  $x = (x_1, \dots, x_n)$ ,  $y = (y_1, \dots, y_n)$  y  $a \in \mathbb{R}$ , entonces  $x + y = (x_1 + y_1, \dots, x_n + y_n)$  y  $ax = (ax_1, \dots, ax_n)$ . Este espacio es conocido como **espacio euclidiano**.

A los elementos del espacio vectorial  $V$  se les llama **vectores** o **puntos**, y a los elementos de  $\mathbb{R}$  se les llama **escalares**.

Un vector  $x \in V$  es **combinación lineal** de los vectores  $v_1, \dots, v_r \in V$  si existen escalares  $a_1, \dots, a_r$  tales que  $x = a_1v_1 + \dots + a_rv_r$ . Decimos que  $v_1, \dots, v_r \in V$  son **linealmente independientes (l.i.)** si ninguno de ellos es combinación lineal de los elementos restantes.

Una **base** de un espacio vectorial  $V$  es un conjunto linealmente independiente  $B$  tal que todo elemento de  $V$  se puede expresar (de manera única) como combinación lineal de elementos de  $B$ . Todo espacio vectorial  $V$  posee una base. Además, dos bases cualesquiera de  $V$  tienen la misma cardinalidad, esto es, existe una correspondencia biyectiva entre ellas. Si  $V$  posee una base de cardinalidad finita diremos que  $V$  es de **dimensión finita** y en el caso contrario decimos que  $V$  tiene **dimensión infinita**.

**Ejemplo 1.2.** El conjunto  $\{e_1, \dots, e_n\}$  es una base del espacio vectorial real  $\mathbb{R}^n$ , donde  $e_i$  es el vector de coordenada  $i$ -ésima 1 y demás coordenadas 0. Esta base es llamada la **base canónica** de  $\mathbb{R}^n$ .

Sean  $V$  y  $W$  espacios vectoriales. Una **transformación lineal** es una aplicación  $T : V \rightarrow W$  tal que para todo  $x, y \in V$  y todo  $a \in \mathbb{R}$  se cumple  $T(ax + y) = aT(x) + T(y)$ . Cuando  $W = \mathbb{R}$  decimos que  $T$  es una **funcional lineal**.

Si  $T : V \rightarrow W$  es una transformación lineal entre espacios vectoriales de dimensión finita, con bases  $\{v_1, \dots, v_n\} \subset V$  y  $\{w_1, \dots, w_m\} \subset W$ , entonces existe una matriz  $(t_{ij})$  de entradas reales con  $m$  filas y  $n$  columnas tal que

$$T(v_j) = \sum_{i=1}^m t_{ij}w_i \quad \text{para todo } j = 1, \dots, n.$$

Decimos que  $(t_{ij})$  es la **matriz de la transformación lineal**  $T$  con respecto a las bases  $\{v_1, \dots, v_n\}$  y  $\{w_1, \dots, w_m\}$ . Esta matriz posee la siguiente propiedad: Dado  $v = a_1v_1 + \dots + a_nv_n$  tenemos  $T(v) = b_1w_1 + \dots + b_mw_m$ , donde los  $a_i$  y los  $b_j$  satisfacen la relación

$$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = (t_{ij}) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

**Ejemplo 1.3.** Sea  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  una transformación lineal. Trabajando con respecto a las bases canónicas de  $\mathbb{R}^n$  y  $\mathbb{R}^m$  vemos que la matriz  $(t_{ij})$  de  $T$  determina  $T$ , esto es, si pensamos en  $\mathbb{R}^p$  como el espacio  $\mathbb{R}^{p \times 1}$  de matrices  $p \times 1$  (véase la Observación 1.19), entonces para todo  $v \in \mathbb{R}^n$  tenemos

$$T(v) = (t_{ij}) \cdot v.$$

Una aplicación  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$  sobre el espacio vectorial  $V$  es un **producto interno** si satisface los siguientes axiomas

- (1)  $\langle ax + y, z \rangle = a\langle x, z \rangle + \langle y, z \rangle$  para todo  $x, y \in V$  y  $a \in \mathbb{R}$ ;
- (2)  $\langle x, y \rangle = \langle y, x \rangle$  para todo  $x, y \in V$ ;
- (3)  $\langle x, x \rangle > 0$  para todo  $x \in V$  distinto de 0.

**Ejemplo 1.4.** En el espacio euclidiano  $\mathbb{R}^n$ , dados  $x = (x_1, \dots, x_n)$  e  $y = (y_1, \dots, y_n)$  el **producto interno canónico** está dado por

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i.$$

Una **norma** sobre el espacio vectorial  $V$  es una función  $\| \cdot \| : V \rightarrow \mathbb{R}$  que tiene las siguientes propiedades:

- (1)  $\|x\| > 0$  para todo  $x \in V$  distinto de 0;
- (2)  $\|\lambda x\| = |\lambda| \cdot \|x\|$  para todo  $x \in V$  y todo  $\lambda \in \mathbb{R}$ ;
- (3)  $\|x + y\| \leq \|x\| + \|y\|$  para todo  $x, y \in V$ .

Al espacio vectorial  $V$  que posee una norma se le denomina **espacio normado**.

Si el espacio vectorial  $V$  posee un producto interno  $\langle \cdot, \cdot \rangle$ , entonces la función  $\| \cdot \| : V \rightarrow \mathbb{R}$ ,  $\|x\| = \sqrt{\langle x, x \rangle}$  es una norma. Se suele decir que esta norma es **inducida** por el producto interno  $\langle \cdot, \cdot \rangle$ .

**Ejemplo 1.5.** El espacio euclidiano  $\mathbb{R}^n$  posee varias normas. Por ejemplo, las funciones  $\| \cdot \|, \| \cdot \|_s, \| \cdot \|_\infty : \mathbb{R}^n \rightarrow \mathbb{R}$  dadas por

$$\|x\| = \sqrt{\langle x, x \rangle}, \quad \|x\|_s = \sum_{i=1}^n |x_i|, \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|,$$

donde  $x = (x_1, \dots, x_n)$ , son normas en  $\mathbb{R}^n$ . Estas normas son denominadas la **norma euclidiana**, la **norma de la suma** y la **norma del máximo**, respectivamente. La norma euclidiana es inducida por el producto interno canónico.

## Espacios topológicos

Una **topología** sobre un conjunto  $X$  es una familia de subconjuntos  $\tau$  de  $X$ , llamados **conjuntos abiertos** o simplemente **abiertos**, que tiene las siguientes propiedades

- (1)  $X$  y  $\emptyset$  son abiertos.
- (2) La unión de cualquier familia de conjuntos abiertos es abierto.
- (3) La intersección de una familia finita de conjuntos abiertos es abierto.

Al par  $(X, \tau)$  que consta de un conjunto  $X$  y de una topología  $\tau$  sobre  $X$  se le llama **espacio topológico**. Si no hay ambigüedad respecto a la topología escribimos  $X$  en vez de  $(X, \tau)$ .

**Observación 1.6.** Las propiedades (2) y (3) de la definición de topología pueden resumirse señalando simplemente que  $\tau$  es cerrada por uniones arbitrarias e intersecciones finitas.

Un subconjunto  $C$  del espacio topológico  $X$  es **cerrado** si su complemento  $X \setminus C$  es abierto. La familia de subconjuntos cerrados de  $X$  contiene a  $\{\emptyset, X\}$  y es cerrada por intersecciones arbitrarias y uniones finitas.

Decimos que  $V \subset X$  es una **vecindad** de  $x \in X$  si existe un conjunto abierto  $U \subset V$  que contiene a  $x$ . Similarmente,  $V \subset X$  es una **vecindad** de  $A \subset X$  si existe un conjunto abierto  $U \subset V$  tal que  $A \subset U$ . Un punto  $x \in X$  es llamado **punto de acumulación** del conjunto  $A \subset X$  si toda vecindad de  $x$  contiene algún punto  $a$  de  $A$  distinto de  $x$ .

Sea  $A \subset X$ . El **interior** de  $A$ , denotado por  $\text{int}(A)$ , es la unión de todos los abiertos  $U$  que están contenidos en  $A$ . La **clausura** de  $A$ , denotada por  $\overline{A}$ , es la intersección de todos los conjuntos cerrados  $C$  que contienen a  $A$ . Notemos que

$$A \text{ es abierto (resp. cerrado) si y solo si } A = \text{int}(A) \text{ (resp. } A = \overline{A}\text{)}.$$

Siempre se cumple que  $\text{int}(A) \subset A \subset \overline{A}$ . Al conjunto de puntos de  $X$  que están en  $\overline{A}$  pero no en  $\text{int}(A)$  se le llama **frontera** de  $A$  y se le denota por  $\partial A$ , es decir,  $\partial A = \overline{A} \setminus \text{int}(A)$ . La frontera de un conjunto es siempre cerrada. Claramente,

$$\begin{aligned} \text{int}(A) &= \{x \in X \mid \text{existe una vecindad } V \text{ de } x \text{ tal que } V \subset A\}, \\ \overline{A} &= \{x \in X \mid \text{toda vecindad } V \text{ de } x \text{ intersecta } A\}, \\ \partial A &= \{x \in X \mid \text{toda vecindad } V \text{ de } x \text{ intersecta } A \text{ y } X \setminus A\}. \end{aligned}$$

Un **cubrimiento abierto** de  $A \subset X$  es una familia  $\{U_\lambda\}_{\lambda \in \Lambda}$  de conjuntos abiertos cuya unión contiene a  $A$ .

Una familia  $\{C_\lambda\}_{\lambda \in \Lambda}$  de subconjuntos de  $X$  tiene la **propiedad de la intersección finita** si  $\bigcap_{\lambda \in \Lambda'} C_\lambda$  es no vacío para todo  $\Lambda' \subset \Lambda$  finito.

Un espacio topológico  $X$  es **compacto** si todo cubrimiento abierto  $\{U_\lambda\}_{\lambda \in \Lambda}$  de  $X$  admite un subcubrimiento finito, esto es, existe  $\Lambda' \subset \Lambda$  finito tal que  $X = \bigcup_{\lambda \in \Lambda'} U_\lambda$ . Equivalentemente,  $X$  es compacto si toda familia  $\{C_\lambda\}_{\lambda \in \Lambda}$  de subconjuntos cerrados de  $X$  que posee la propiedad de la intersección finita tiene intersección  $\bigcap_{\lambda \in \Lambda} C_\lambda$  no vacía.

Si  $Y$  es un subconjunto del espacio topológico  $(X, \tau)$  entonces

$$\tau_Y := \{Y \cap U \mid U \in \tau\}$$

es una topología sobre  $Y$ , llamada **topología relativa** de  $Y$ . Los subconjuntos cerrados de  $(Y, \tau_Y)$  están dados por

$$\{Y \cap C \mid C \text{ es cerrado en } X\}.$$

Un subconjunto  $Y$  de  $(X, \tau)$  es **compacto** si el espacio topológico  $(Y, \tau_Y)$  es compacto. Es fácil verificar que  $Y$  es compacto si y solo si todo cubrimiento de  $Y$  (por abiertos de  $X$ ) admite un subcubrimiento finito. Todo subconjunto cerrado de un espacio topológico compacto es compacto.

Una función  $f : (X, \tau_X) \rightarrow (Y, \tau_Y)$  entre espacios topológicos es **continua** si la preimagen de todo abierto es abierto, es decir,  $f^{-1}(B) \in \tau_X$  para todo  $B \in \tau_Y$ . Si  $X$  es compacto y  $f$  es continua, entonces la imagen  $f(X)$  de  $X$  es un subconjunto compacto de  $Y$ .

## Espacios métricos

Dado un conjunto  $X$ , decimos que una función  $d : X \times X \rightarrow \mathbb{R}$  es una **métrica** si satisface las siguientes propiedades

- (1)  $d(x, y) \geq 0$  para todo  $x, y \in X$ ;
- (2)  $d(x, y) = 0$  si y solo si  $x = y$ ;
- (3)  $d(x, z) \leq d(x, y) + d(y, z)$  para todo  $x, y, z \in X$ ;
- (4)  $d(x, y) = d(y, x)$  para todo  $x, y \in X$ .

Al conjunto  $X$  que posee una métrica  $d$  se le llama **espacio métrico**, y se le denota por  $(X, d)$ . Si no hay ambigüedad respecto a la métrica escribimos  $X$  en vez de  $(X, d)$ .

**Ejemplo 1.7.** Si  $X$  es un espacio normado entonces la función  $d : X \times X \rightarrow \mathbb{R}$  dada por  $d(x, y) = \|x - y\|$  es una métrica. Decimos entonces que la norma  $\|\cdot\|$  **induce** una métrica en  $X$ . Así, todo espacio normado es un espacio métrico. En particular, el espacio euclidiano  $\mathbb{R}^n$  es un espacio métrico con la métrica  $d(x, y) = \|x - y\|$  inducida por la norma euclidiana.

En el espacio métrico  $(X, d)$ , dados  $a \in X$  y  $r > 0$ , definimos los conjuntos

$$B(a, r) = \{x \in X \mid d(a, x) < r\},$$

$$\overline{B}(a, r) = \{x \in X \mid d(a, x) \leq r\},$$

que son llamados **bola abierta** de centro  $a$  y radio  $r$ , y **bola cerrada** de centro  $a$  y radio  $r$ . La familia de conjuntos

$$\tau_d := \{A \subset X \mid \text{para todo } a \in A \text{ existe } r > 0 \text{ tal que } B(a, r) \subset A\}$$

define una topología sobre  $X$ . Luego  $(X, d)$  es un espacio topológico, de manera que la terminología introducida en la parte anterior se aplica para  $(X, d)$ .

**Ejemplo 1.8.** Toda bola del tipo  $B(a, r)$  es un conjunto abierto, mientras que las bolas del tipo  $\overline{B}(a, r)$  son conjuntos cerrados. Todo conjunto abierto es una vecindad de cada uno de sus puntos.

Dado un subconjunto  $Y$  de  $(X, d)$ , la restricción de  $d$  a  $Y \times Y$  es una métrica en  $Y$ , y por tanto  $Y$  es un espacio métrico. Decimos que  $d$  induce una métrica en  $Y$ , y, en un abuso de notación, escribimos  $(Y, d)$  en vez de  $(Y, d|_{Y \times Y})$ . La topología del espacio métrico  $(Y, d)$  coincide con la topología relativa de  $Y$ . Usamos el símbolo  $B_Y(a, r)$  para denotar las bolas en  $Y$  de centro  $a \in Y$  y radio  $r > 0$ ; explícitamente,

$$B_Y(a, r) = \{y \in Y \mid d(a, y) < r\} = B(a, r) \cap Y.$$

Decimos que  $A \subset Y$  es **abierto (cerrado) en  $Y$**  si  $A$  es abierto (cerrado) como subconjunto del espacio métrico  $(Y, d)$ . Notemos en particular que al hablar de conjuntos abiertos en  $Y$  se sobreentiende que ellos están contenidos en  $Y$ . Análogamente,  $U \subset X$  es una **vecindad de  $A \subset Y$  en  $Y$**  si  $U$  es una vecindad de  $A$  en el espacio métrico  $(Y, d)$ .

Una sucesión  $(x_n)$  en  $(X, d)$  **converge** a  $a \in X$  si para todo  $\varepsilon > 0$  existe  $n_0 \in \mathbb{N}$  tal que  $n > n_0$  implica  $d(x_n, a) < \varepsilon$ . En este caso escribimos  $x_n \rightarrow a$  o  $\lim_{n \rightarrow \infty} x_n = a$ . Una sucesión  $(x_n) \subset X$  es **de Cauchy** si para todo  $\varepsilon > 0$  existe  $n_0 \in \mathbb{N}$  tal que  $m, n > n_0$  implica  $d(x_m, x_n) < \varepsilon$ . Toda sucesión convergente es de Cauchy, pero no toda sucesión de Cauchy es convergente (la sucesión  $x_n = 1/n$  en  $X := (0, 1]$  es de Cauchy pero no es convergente). Cuando toda sucesión de Cauchy en  $X$  es convergente decimos que  $X$  es un **espacio métrico completo**. Un punto  $x \in X$  es **punto de acumulación** de la sucesión  $(x_n) \subset X$  si existe una subsucesión  $(x_j)$  de  $(x_n)$  tal que  $x_j \rightarrow x$ .

**Ejemplo 1.9.** El espacio euclidiano  $\mathbb{R}^n$ , con la métrica del Ejemplo 1.7, es un espacio métrico completo.

La convergencia de sucesiones determina la topología de  $(X, d)$ . En efecto, un conjunto  $C \subset X$  es cerrado si y solo si para toda sucesión  $(x_n) \subset C$  que converge a  $x \in X$  se cumple que  $x \in C$ . Más aun, para cada subconjunto  $A$  de  $X$  tenemos

$$\begin{aligned} \text{int}(A) &= \{x \in X \mid \text{existe } r > 0 \text{ tal que } B(x, r) \subset A\}, \\ \bar{A} &= \{x \in X \mid \text{existe una sucesión en } A \text{ que converge a } x\}, \\ \partial A &= \{x \in X \mid \text{existen sucesiones en } A \text{ y en } X \setminus A \text{ que convergen a } x\}; \end{aligned}$$

además,  $x \in X$  es punto de acumulación de  $A$  si y solo si existe una sucesión  $(x_n)$  contenida en  $A \setminus \{x\}$  tal que  $x_n \rightarrow x$ .

La continuidad de funciones entre espacios métricos admite la siguiente caracterización<sup>2</sup>.

**Proposición 1.10.** *Dada una función  $f : (X, d_X) \rightarrow (Y, d_Y)$  entre espacios métricos, las siguientes afirmaciones son equivalentes*

- (1)  $f$  es continua (la pre-imagen de todo abierto en  $Y$  es abierto en  $X$ );
- (2)  $f^{-1}(C)$  es cerrado en  $X$  para todo  $C \subset Y$  cerrado en  $Y$ ;
- (3) para cualesquiera  $\bar{x} \in X$  y  $\varepsilon > 0$  existe  $\delta > 0$  tal que  $d_X(x, \bar{x}) < \delta$  implica  $d_Y(f(x), f(\bar{x})) < \varepsilon$ ;

<sup>2</sup>La definición de continuidad que uno normalmente encuentra al estudiar los espacios métricos es la condición que aparece en el tercer ítem de la Proposición 1.10.

(4) para todo  $\bar{x} \in X$  y  $(x_n)$  en  $X$  tal que  $x_n \rightarrow x$  se cumple que  $f(x_n) \rightarrow f(\bar{x})$ .

**Ejemplo 1.11** (Función lipschitziana). Una función  $f : (X, d_X) \rightarrow (Y, d_Y)$  es **lipschitziana** si existe  $L \geq 0$  tal que  $d_Y(f(x), f(x')) \leq L \cdot d_X(x, x')$  para todo  $x, x' \in X$ . Toda función lipschitziana es continua.

**Ejemplo 1.12** (Función distancia  $d_C$ ). Dado un elemento  $a$  de  $(X, d)$ , la aplicación  $x \mapsto d(x, a)$  define una función continua de  $X$  en  $\mathbb{R}$ . En particular, si  $X$  es un espacio normado y  $a = 0$ , obtenemos que la función norma  $\|\cdot\|$  es continua en  $X$ . Si  $C$  es un subconjunto no vacío de  $(X, d)$ , la **función distancia**  $d_C : X \rightarrow \mathbb{R}$ , definida por  $d_C(x) = \inf_{y \in C} d(x, y)$ , también es una función continua.

Decimos que un subconjunto  $A$  del espacio métrico  $(X, d)$  es **acotado** si está contenido en alguna bola  $B(x, r)$ . Todo conjunto compacto es acotado.

A menudo utilizaremos la siguiente proposición para mostrar que un conjunto dado es compacto.

**Proposición 1.13.** *Sea  $Y$  un subconjunto del espacio métrico  $(X, d)$ . Las siguientes afirmaciones son equivalentes:*

- (1)  $Y$  es un conjunto compacto;
- (2) toda sucesión  $(y_n) \subset Y$  posee una subsucesión convergente en  $Y$ ;
- (3) todo cubrimiento abierto de  $Y$  (por abiertos de  $X$ ) admite un subcubrimiento finito;
- (4) toda familia de conjuntos cerrados en  $Y$  con la propiedad de la intersección finita tiene intersección no vacía.

Los dos teoremas que siguen son resultados clásicos que ilustran la relevancia de los espacios métricos compactos.

**Teorema 1.14** (Weierstrass). *Toda función continua  $f : X \rightarrow \mathbb{R}$  definida sobre un espacio métrico compacto  $(X, d)$  alcanza su mínimo y su máximo en  $X$ , es decir, existen  $x_0, x_1 \in X$  tales que  $f(x_0) \leq f(x) \leq f(x_1)$  para todo  $x \in X$ .*

**Teorema 1.15** (Cantor). *Sea  $K_1 \supset K_2 \supset \dots$  una sucesión decreciente de subconjuntos compactos y no vacíos del espacio métrico completo  $(X, d)$ . Entonces la intersección  $\bigcap_{n \geq 1} K_n$  es un conjunto compacto y no vacío.*

## El espacio euclidiano $\mathbb{R}^n$

En adelante, trabajaremos solamente con el espacio euclidiano  $\mathbb{R}^n$ . De lo visto previamente sabemos que  $\mathbb{R}^n$  es un espacio métrico completo con la métrica inducida por la norma euclidiana  $\|\cdot\|$ . Luego, todo lo afirmado hasta ahora se cumple para  $\mathbb{R}^n$ , y aquí solo estableceremos algunos hechos adicionales.

Cualquier otra norma de  $\mathbb{R}^n$  define la misma topología que la norma euclidiana, es decir, los abiertos en ambas topologías coinciden. Esto es consecuencia del siguiente teorema, que nos dice que la convergencia de sucesiones en  $\mathbb{R}^n$  no depende de la norma considerada.

**Teorema 1.16** (Equivalencia de normas). *Dada cualquier otra norma  $\|\cdot\|_1$  en  $\mathbb{R}^n$ , existen constantes positivas  $C$  y  $D$  tales que  $C\|x\| \leq \|x\|_1 \leq D\|x\|$  para todo  $x \in \mathbb{R}^n$ .*

El conjunto  $B(0, 1)$  es la **bola unitaria abierta** de  $\mathbb{R}^n$ , y su clausura  $\overline{B}(0, 1)$  es la **bola unitaria cerrada** de  $\mathbb{R}^n$ . La clausura de toda bola abierta  $B(a, r)$  es la bola cerrada  $\overline{B}(a, r)$ , si bien esto no siempre es cierto en espacios métricos. La **esfera unitaria** de  $\mathbb{R}^n$  es el conjunto

$$S(0, 1) = \{x \in \mathbb{R}^n \mid \|x\| = 1\}.$$

Dos puntos cualesquiera  $x$  e  $y$  de  $\mathbb{R}^n$  satisfacen la **desigualdad de Cauchy-Schwarz**

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|;$$

esta desigualdad es una igualdad si y solo si  $x$  e  $y$  “tienen la misma dirección”, es decir, existe un escalar  $t \geq 0$  tal que  $y = tx$  o  $x = ty$ .

Dados dos subconjuntos  $A$  y  $B$  de  $\mathbb{R}^n$  y un escalar  $\lambda \in \mathbb{R}$  definimos la suma de  $A$  y  $B$  mediante

$$A + B = \{a + b \mid a \in A, b \in B\}$$

y el producto de  $\lambda$  y  $A$  mediante

$$\lambda A = \{\lambda a \mid a \in A\}.$$

El siguiente teorema es un resultado clásico que usaremos con mucha frecuencia.

**Teorema 1.17** (Bolzano-Weierstrass). *Toda sucesión acotada en  $\mathbb{R}^n$  posee una sub-sucesión convergente.*

Combinando el teorema de Bolzano-Weierstrass con la Proposición 1.13 obtenemos una caracterización de los subconjuntos compactos de  $\mathbb{R}^n$ .

**Proposición 1.18.** *Un subconjunto  $K$  de  $\mathbb{R}^n$  es compacto si y solo si es cerrado y acotado.*

Por ejemplo, el espacio  $\mathbb{R}^n$ , que es cerrado pero no acotado, no es compacto. Podemos ver esto directamente: la familia  $\{B(0, m)\}_{m \geq 1}$  es un cubrimiento abierto de  $\mathbb{R}^n$  que no admite subcubrimiento finito.

**Observación 1.19.** Por cuestiones de espacio, en ocasiones será conveniente denotar a los vectores de  $\mathbb{R}^n$  como arreglos verticales. Es decir, dado  $v = (v_1, \dots, v_n) \in \mathbb{R}^n$ ,

la matriz  $\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$  también representará a  $v$ . Esto no debería dar lugar a confusión.

Por ejemplo, si  $x$  e  $y$  están en  $\mathbb{R}^n$  entonces  $x^T y$  es igual al producto interno canónico  $\langle x, y \rangle$  de  $x$  e  $y$  ( $A^T$  denota la traspuesta de la matriz  $A$ ).

Consideremos la base canónica  $\{e_1, \dots, e_n\}$  de  $\mathbb{R}^n$ . Sea la función  $f : U \rightarrow \mathbb{R}$  definida sobre el abierto  $U$  de  $\mathbb{R}^n$ . Definimos la **derivada direccional** de  $f$  en  $a \in U$  en la dirección  $v \in \mathbb{R}^n$  como el número

$$\frac{\partial f}{\partial v}(a) := \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t},$$

siempre que este límite exista; también definimos la **derivada direccional unilateral** de  $f$  en  $a$  en la dirección  $v$  por

$$f'(a; v) := \lim_{t \downarrow 0} \frac{f(a + tv) - f(a)}{t},$$

siempre que el límite exista. Esta derivada direccional será especialmente útil al trabajar con funciones convexas.

La  $i$ -ésima **derivada parcial** de  $f$  en  $a$ , denotada por  $\frac{\partial f}{\partial x_i}(a)$ , es la derivada direccional de  $f$  en  $a$  en la dirección canónica  $e_i$ , esto es,

$$\frac{\partial f}{\partial x_i}(a) := \lim_{t \rightarrow 0} \frac{f(a + te_i) - f(a)}{t}.$$

Cuando  $f$  tiene todas sus derivadas parciales en  $a$  el vector

$$\nabla f(a) := \left( \frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right) \in \mathbb{R}^n$$

es el **gradiente** de  $f$  en  $a$ , y decimos que  $f$  es **derivable** en  $a$ . Si  $f$  es derivable en cada punto de  $U$  decimos simplemente que  $f$  es **derivable**.

Una función  $f : U \rightarrow \mathbb{R}$  es **diferenciable** en  $a \in U$  si existe una transformación lineal  $f'(a) : \mathbb{R}^n \rightarrow \mathbb{R}$ , llamada **derivada** de  $f$  en  $a$ , tal que

$$f(a + h) = f(a) + f'(a)(h) + r(h), \quad a + h \in U, \quad (1.1)$$

donde  $\lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0$ . Si  $f$  es diferenciable en todos los puntos de  $U$  decimos simplemente que  $f$  es **diferenciable**. Por ejemplo, toda transformación lineal  $T : \mathbb{R}^n \rightarrow \mathbb{R}$  es diferenciable con  $T'(x) = T$ .

Toda función diferenciable es derivable. En efecto, dados  $a \in U$  y  $v \in \mathbb{R}^n \setminus \{0\}$ , tomando  $h = tv$  y haciendo  $t \rightarrow 0$  en (1.1) obtenemos

$$\frac{\partial f}{\partial v}(a) = \lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = f'(a)(v),$$

de donde

$$\frac{\partial f}{\partial x_i}(a) = f'(a)(e_i)$$

para todo  $i$ , es decir,  $f$  es derivable en  $a$ . Esto muestra en particular que  $\nabla f(a)$  es la matriz de la funcional  $f'(a)$  y que por tanto

$$f'(a; v) = \frac{\partial f}{\partial v}(a) = f'(a)(v) = \langle \nabla f(a), v \rangle.$$

Toda función diferenciable es continua, mas existen funciones derivables que no son continuas (ni diferenciables). De hecho, la función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  dada por  $f(x, y) = \frac{xy}{x^2 + y^2}$  si  $(x, y) \neq (0, 0)$  y  $f(0, 0) = 0$  es derivable pero no es continua en  $(0, 0)$ .

Una función derivable  $f : U \rightarrow \mathbb{R}$  es **de clase  $C^1$**  si cada función  $\frac{\partial f}{\partial x_i} : U \rightarrow \mathbb{R}$  es continua. Toda función de clase  $C^1$  es diferenciable, mas no toda función diferenciable es de clase  $C^1$ . En resumen,

$$f : \text{de clase } C^1 \Rightarrow f : \text{diferenciable} \Rightarrow f : \text{derivable}.$$

**Observación 1.20.** Si tenemos una función  $f : U \times V \rightarrow \mathbb{R}$ , donde  $U \subset \mathbb{R}^m$  y  $V \subset \mathbb{R}^n$ , los símbolos  $\nabla_x f(x, y)$  y  $\nabla_y f(x, y)$  representarán a los vectores

$$\left( \frac{\partial f}{\partial x_1}(x, y), \dots, \frac{\partial f}{\partial x_m}(x, y) \right) \in \mathbb{R}^m \quad \text{y} \quad \left( \frac{\partial f}{\partial y_1}(x, y), \dots, \frac{\partial f}{\partial y_n}(x, y) \right) \in \mathbb{R}^n$$

respectivamente, donde los  $x_i$  representan a las variables en  $\mathbb{R}^m$  y los  $y_j$  representan a las variables en  $\mathbb{R}^n$ .

Dada una función  $f : U \rightarrow \mathbb{R}^n$  donde  $U$  es un subconjunto abierto de  $\mathbb{R}^m$ , tenemos  $f = (f_1, \dots, f_n)$ , donde cada  $f_i$  es una función de  $\mathbb{R}^m$  en  $\mathbb{R}$ . Decimos que  $f$  es **diferenciable** en  $a \in U$  si existe una (única) transformación lineal  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$  tal que para todo  $h \in \mathbb{R}^m$  tal que  $a + h \in U$  se cumple

$$f(a + h) = f(a) + f'(a)(h) + r(h), \tag{1.2}$$

donde  $\lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0$ . Decimos también que  $f$  es **de clase  $C^1$**  si cada  $f_i$  es derivable y si las funciones  $\frac{\partial f_i}{\partial x_j} : U \rightarrow \mathbb{R}$  son continuas.

Claramente, si  $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$ , donde  $U \subset \mathbb{R}^m$ , entonces

$f$  es diferenciable si y solo si cada  $f_i$  es diferenciable,  
 $f$  es de clase  $C^1$  si y solo si cada  $f_i$  es de clase  $C^1$ .

Cuando  $f$  es diferenciable se puede mostrar que

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_n}{\partial x_1}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_1}{\partial x_m}(a) & \dots & \frac{\partial f_n}{\partial x_m}(a) \end{pmatrix}$$

es la matriz de la transformación lineal  $f'(a) : \mathbb{R}^m \rightarrow \mathbb{R}^n$ . Más aun, si consideramos la composición de dos funciones diferenciables  $f : U \rightarrow \mathbb{R}^n$  y  $g : V \rightarrow \mathbb{R}^p$  donde  $U \subset \mathbb{R}^m$  y  $f(U) \subset V \subset \mathbb{R}^n$ , entonces la matriz de  $(g \circ f)'(a)$  es igual al producto de las matrices de  $g'(f(a))$  y  $f'(a)$ . Esto es lo que nos dice la Regla de la Cadena.

**Teorema 1.21** (Regla de la Cadena). *Sean  $U \subset \mathbb{R}^m$  y  $V \subset \mathbb{R}^n$  abiertos. Sea  $f = (f_1, \dots, f_n) : U \rightarrow \mathbb{R}^n$  tal que  $f(U) \subset V$  y cada función coordenada  $f_k : U \rightarrow \mathbb{R}$  es diferenciable en el punto  $a \in U$ . Sea también  $g : V \rightarrow \mathbb{R}$  una función diferenciable en el punto  $b = f(a)$ . Entonces la composición  $g \circ f : U \rightarrow \mathbb{R}$  es diferenciable en el punto  $a$  y sus derivadas parciales son*

$$\frac{\partial (g \circ f)}{\partial x_i}(a) = \sum_{j=1}^n \frac{\partial g}{\partial y_j}(b) \cdot \frac{\partial f_j}{\partial x_i}(a).$$

Ahora mostraremos el lema de Urysohn, que nos dice que se pueden “separar” dos subconjuntos cerrados y disjuntos de  $\mathbb{R}^n$  con funciones continuas.

**Teorema 1.22** (Lema de Urysohn). *Para todo par de conjuntos cerrados, disjuntos y no vacíos  $U, V \subset \mathbb{R}^n$ , existe una función continua  $f : \mathbb{R}^n \rightarrow [0, 1]$  tal que  $f(x) = 0$  si  $x \in U$  y  $f(x) = 1$  si  $x \in V$ .*

PRUEBA. Sabemos que cuando  $C \subset \mathbb{R}^n$  es cerrado, la función distancia  $x \mapsto d_C(x)$  es continua; además,  $d_C(x) = 0$  si y solo si  $x \in C$ . Entonces  $0 < d_U(x) + d_V(x)$  para todo  $x \in \mathbb{R}^n$  pues siendo  $U$  y  $V$  disjuntos y cerrados, las funciones  $d_U$  y  $d_V$  no pueden anularse al mismo tiempo. Por tanto, la función  $f : \mathbb{R}^n \rightarrow [0, 1]$  dada por

$$f(x) = \frac{d_V(x)}{d_U(x) + d_V(x)}$$

está bien definida y es continua pues es el cociente de dos funciones continuas. Finalmente, observemos que  $f(x) = 0$  si  $x \in U$  y  $f(x) = 1$  si  $x \in V$ .  $\square$

**Corolario 1.23.** *Si  $U$  y  $C$  son subconjuntos de  $\mathbb{R}^n$  tales que  $U$  es abierto,  $C$  es cerrado y  $C \subset U$ , entonces existe una función  $f : \mathbb{R}^n \rightarrow [0, 1]$  tal que  $f(x) = 1$  si  $x \in C$  y  $f(x) = 0$  si  $x \notin U$ .*

PRUEBA. Si  $C = \emptyset$  definimos  $f = 0$ , y si  $U = \mathbb{R}^n$  tomamos  $f = 1$ . En ambos casos, la función  $f$  tiene la propiedad deseada. Cuando  $C \neq \emptyset$  y  $U \neq \mathbb{R}^n$ , la función  $f$  se puede conseguir aplicando el Teorema 1.22 a los conjuntos cerrados, disjuntos y no vacíos  $C$  y  $\mathbb{R}^n \setminus U$ .  $\square$

## Particiones de la unidad

Terminamos esta sección mostrando un caso particular del teorema de existencia de particiones de la unidad. Nuestra demostración está basada en [30] y [31].

Dada una función  $f : U \rightarrow \mathbb{R}$  definida en el conjunto abierto  $U \subset \mathbb{R}^n$ , el **soporte** de  $f$ , denotado por  $\text{sop}(f)$ , es la clausura (relativa a  $U$ ) de los puntos de  $U$  donde  $f$  no se anula, esto es,

$$\text{sop}(f) = \overline{\{x \in U \mid f(x) \neq 0\}} \cap U.$$

Notemos que  $x \in U$  no está en el soporte de  $f$  si y solo si existe  $\delta > 0$  tal que

$$B(x, \delta) \cap \{x \in U \mid f(x) \neq 0\} = \emptyset.$$

En particular,  $f(x) = 0$  para todo  $x \in U \setminus \text{sop}(f)$ .

Sean el conjunto  $K \subset \mathbb{R}^n$  y  $U_1, \dots, U_r$  un cubrimiento abierto de  $K$ . Decimos que la familia de funciones continuas  $\varphi_1, \dots, \varphi_r : \mathbb{R}^n \rightarrow [0, 1]$  es una **partición de la unidad en  $K$** , subordinada al cubrimiento  $U_1, \dots, U_r$ , si se cumplen las siguientes condiciones:

$$(1) \sum_{i=1}^r \varphi_i(x) = 1 \text{ para todo } x \in K;$$

$$(2) \text{sop}(\varphi_i) \subset U_i \text{ para cada } i = 1, \dots, r.$$

**Teorema 1.24.** *Para todo conjunto compacto  $K \subset \mathbb{R}^n$  y todo cubrimiento abierto y finito de  $K$ , existe una partición de la unidad subordinada a dicho cubrimiento.*

PRUEBA. Sea  $U_1, \dots, U_r$  un cubrimiento abierto y finito de  $K$ . Sabemos que cada  $x \in K$  está contenido en algún conjunto abierto  $U_i$ . Entonces, para cada  $x \in K$  existen  $\delta_x, \delta'_x > 0$  tales que  $\delta_x < \delta'_x$  y

$$B(x, \delta_x) \subset \overline{B(x, \delta_x)} \subset B(x, \delta'_x) \subset \overline{B(x, \delta'_x)} \subset U_{i(x)},$$

donde  $i(x) \in \{1, \dots, r\}$ . Escribiendo  $B_x$  en vez de  $B(x, \delta_x)$  y  $B'_x$  en vez de  $B(x, \delta'_x)$ , vemos que  $\{B_x\}_{x \in X}$  constituye un cubrimiento abierto del conjunto compacto  $K$ . Luego existen  $x_1, \dots, x_m$  en  $K$  tales que  $K \subset B_{x_1} \cup \dots \cup B_{x_m}$ . Para cada  $i = 1, \dots, r$  sea  $H_i$  la unión de todos los conjuntos  $\overline{B_{x_j}}$  tales que  $\overline{B'_{x_j}} \subset U_i$ , y sea  $W_i$  la unión de todos los conjuntos  $B'_{x_j}$  tales que  $\overline{B'_{x_j}} \subset U_i$  (nótese que  $H_i$  y  $W_i$  pueden ser vacíos). Entonces  $K \subset H_1 \cup \dots \cup H_r$ , y cada  $H_i$  es un conjunto cerrado contenido en  $W_i$ , que es un abierto cuya clausura está contenida en  $U_i$ :

$$\overline{W_i} = \overline{\bigcup_{\overline{B'_{x_j}} \subset U_i} B'_{x_j}} = \bigcup_{\overline{B'_{x_j}} \subset U_i} \overline{B'_{x_j}} \subset U_i.$$

Por el Corolario 1.23, existen funciones continuas  $\psi_1, \dots, \psi_r : \mathbb{R}^n \rightarrow [0, 1]$  tales que  $\psi_i(x) = 1$  si  $x \in H_i$  y  $\psi_i(x) = 0$  si  $x \notin W_i$ . De  $\{x \in \mathbb{R}^n \mid \psi_i(x) \neq 0\} \subset W_i$  se sigue que  $\text{sop}(\psi_i) \subset U_i$  para cada  $i$ . Luego, si definimos las funciones  $\varphi_1, \dots, \varphi_r$  mediante  $\varphi_1 = \psi_1$  y

$$\varphi_{i+1} = (1 - \psi_1) \cdots (1 - \psi_i) \psi_{i+1} \tag{1.3}$$

cuando  $i = 1, \dots, r - 1$ , es claro que  $\text{sop}(\varphi_i) \subset \text{sop}(\psi_i) \subset U_i$  y que  $0 \leq \varphi_i \leq 1$  para cada  $i$ . Además, la relación

$$\varphi_1 + \cdots + \varphi_i = 1 - (1 - \psi_1) \cdots (1 - \psi_i) \tag{1.4}$$

se satisface trivialmente para  $i = 1$ ; y si (1.4) se verifica para  $i < r$ , sumando (1.3) y (1.4) vemos que (1.4) también se verifica para  $i + 1$ . Así, concluimos por inducción que

$$\varphi_1 + \cdots + \varphi_r = 1 - (1 - \psi_1) \cdots (1 - \psi_r),$$

y como  $K \subset H_1 \cup \dots \cup H_r$ , al menos uno de los  $\psi_i(x)$  será 1 cuando  $x \in K$ , esto es,

$$\varphi_1(x) + \cdots + \varphi_r(x) = 1 \quad \text{para todo } x \in K.$$

Esto muestra que  $\varphi_1, \dots, \varphi_r$  es la partición de unidad que buscábamos.  $\square$

## 2. Análisis convexo

El objeto de estudio del análisis convexo son los conjuntos convexos y las funciones convexas. Una propiedad geométrica tan simple como la convexidad puede tener implicancias profundas. Por ejemplo, los conjuntos convexos pueden *separarse con hiperplanos*; y todo mínimo local de una función convexa es un mínimo global, lo cual significa que para minimizar funciones convexas basta con aplicar métodos de estudio locales.

Todos los resultados de esta sección serán demostrados. Para esto, las principales referencias que utilizaremos serán los libros [5], [8], [9] y [32]. Por supuesto, la referencia canónica de análisis convexo es el libro clásico de Rockafellar [28]. La definición de pseudoconvexidad proviene de [22].

## 2.1. Conjuntos convexos y conos

Dados dos puntos  $x, y \in \mathbb{R}^n$ , denotaremos mediante  $[x, y]$  al segmento que los une, es decir,

$$[x, y] := \{\lambda x + (1 - \lambda)y \mid 0 \leq \lambda \leq 1\}.$$

Un conjunto  $C \subset \mathbb{R}^n$  es **convexo** si para cualesquiera  $x, y \in C$  se verifica que  $[x, y] \subset C$ .

**Ejemplo 2.1.** El espacio  $\mathbb{R}^n$  y las bolas  $B(a, r) = \{x \in \mathbb{R}^n \mid \|x - a\| < r\}$  son ejemplos de conjuntos convexos. El conjunto  $A$  de la Figura 2.1 es convexo mientras que el conjunto  $B$  no lo es.

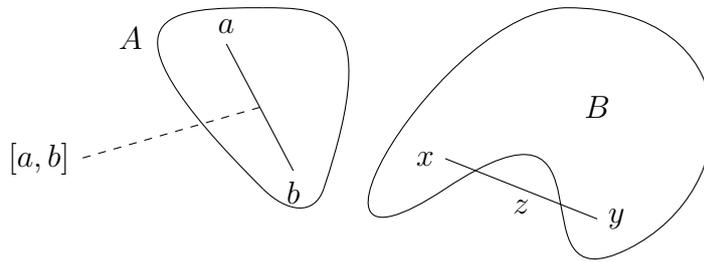


Figura 2.1: Un conjunto convexo y uno que no es convexo.

Un vector  $z \in \mathbb{R}^n$  es una **combinación convexa** de los vectores  $x_1, \dots, x_r \in \mathbb{R}^n$  si existen  $t_1, \dots, t_r \in [0, 1]$  tales que  $\sum_{i=1}^r t_i = 1$  y  $z = \sum_{i=1}^r t_i x_i$ . Por ejemplo, los puntos de  $[x, y]$  son combinaciones convexas de  $x$  e  $y$ .

Se puede probar por inducción que todo conjunto convexo contiene las combinaciones convexas de sus elementos. Recíprocamente, si  $C \subset \mathbb{R}^n$  contiene todas las combinaciones convexas de sus elementos, entonces es convexo. En particular, dado cualquier  $X \subset \mathbb{R}^n$ , el conjunto

$$\text{co}(X) := \left\{ \sum_{i=1}^r t_i x_i \mid r \in \mathbb{N}, (t_i)_{i=1}^r \subset [0, 1], \sum_{i=1}^r t_i = 1, (x_i)_{i=1}^r \subset X \right\},$$

es convexo y está contenido en todos los conjuntos convexos  $C \subset \mathbb{R}^n$  tales que  $X \subset C$  (notemos que  $\text{co}(\emptyset) = \emptyset$ ). Luego,  $\text{co}(X)$  es el *menor* conjunto convexo que contiene a  $X$ . Decimos entonces que  $\text{co}(X)$  es la **cápsula convexa** de  $X$ .

### Proposición 2.2.

- (1) Si  $\{C_\lambda\}_{\lambda \in \Lambda}$  es una familia de subconjuntos convexos de  $\mathbb{R}^n$ , entonces  $\bigcap_{\lambda \in \Lambda} C_\lambda$  es convexo.
- (2) Si  $C \subset \mathbb{R}^n$  es convexo, entonces  $\overline{C}$  es convexo.
- (3) Si  $C_1, C_2 \subset \mathbb{R}^n$  son convexos, entonces  $C_1 + C_2$  y  $C_1 - C_2$  son convexos.
- (4) Si  $X \subset \mathbb{R}^n$ , entonces  $\text{co}(X) = \bigcap_{C \in \mathcal{F}} C$ , donde  $\mathcal{F}$  es la colección de todos los conjuntos convexos  $C \subset \mathbb{R}^n$  que contienen a  $X$ . (Obsérvese que  $\mathbb{R}^n \in \mathcal{F} \neq \emptyset$ .)  
En particular,

$$X \text{ es convexo si y solo si } X = \text{co}(X).$$

PRUEBA. El primer ítem es claro. Para verificar el segundo ítem, sean dos sucesiones  $(x_k)$  e  $(y_k)$  en  $C$  tales que  $x_k \rightarrow x$  e  $y_k \rightarrow y$ . Debemos probar que  $[x, y] \subset \overline{C}$ . De hecho, si  $t \in [0, 1]$ , entonces la sucesión definida por  $tx_k + (1-t)y_k$  está contenida en  $C$  y converge a  $z_t := tx + (1-t)y$ , es decir,  $z_t \in \overline{C}$ . Esto muestra que  $[x, y] \subset \overline{C}$ . Veamos el tercer ítem. Sean  $z := x_1 \pm x_2$  y  $z' := x'_1 \pm x'_2$  elementos de  $C_1 \pm C_2$  con  $x_i, x'_i \in C_i$ . Entonces para cada  $t \in [0, 1]$  tenemos  $tx_i + (1-t)x'_i \in C_i$ , de donde

$$tz + (1-t)z' = (tx_1 + (1-t)x'_1) \pm (tx_2 + (1-t)x'_2) \in C_1 \pm C_2.$$

Luego,  $C_1 \pm C_2$  es convexo. El cuarto ítem es consecuencia del primer ítem y de lo expresado en el párrafo anterior.  $\square$

**Observación 2.3.** La unión de conjuntos convexos no es necesariamente un conjunto convexo. Por ejemplo, los puntos son convexos, pero la unión de dos puntos distintos no es un conjunto convexo.

Dado un subconjunto no vacío  $X$  de  $\mathbb{R}^n$ , cada elemento  $x$  de  $\text{co}(X)$  puede expresarse como combinación convexa de elementos de  $X$ . Podemos de hecho escoger  $n+1$  (o menos) elementos de  $X$  de modo que  $x$  sea combinación convexa de dichos elementos.

**Teorema 2.4** (Carathéodory). *Sea  $X$  un subconjunto no vacío de  $\mathbb{R}^n$  y consideremos  $x \in \text{co}(X)$ . Entonces  $x$  se puede expresar como una combinación convexa de  $n+1$  (o menos) elementos de  $X$ .*

PRUEBA. Basta con mostrar que si  $x$  puede escribirse como combinación convexa de  $r$  elementos de  $X$ , donde  $r > n+1$ , entonces también puede escribirse como combinación convexa de  $r-1$  elementos de  $X$ .

Supongamos que existen  $x_1, \dots, x_r \in X$  y  $\lambda_1, \dots, \lambda_r \in [0, 1]$  tales que  $r > n+1$ ,  $\sum_{i=1}^r \lambda_i = 1$  y  $x = \sum_{i=1}^r \lambda_i x_i$ . Entonces los  $r-1$  vectores  $x_2 - x_1, \dots, x_r - x_1$  son linealmente dependientes (pues  $r-1 > n$ ) y por tanto existen escalares  $t_2, \dots, t_r$  no todos nulos tales que  $\sum_{i=2}^r t_i(x_i - x_1) = 0$ . Definiendo  $t_1 := -\sum_{i=2}^r t_i$  tenemos

$$\sum_{i=1}^r t_i x_i = 0 \quad \text{y} \quad \sum_{i=1}^r t_i = 0.$$

Como no todos los  $t_i$  son nulos existe algún  $t_i$  positivo, de modo que

$$\alpha := \min_{1 \leq i \leq r} \left\{ \frac{\lambda_i}{t_i} \mid t_i > 0 \right\}$$

es un número real no negativo. Sin pérdida de generalidad supongamos que  $\alpha = \frac{\lambda_r}{t_r}$ . Ahora bien, si  $t_i > 0$  entonces  $\alpha t_i \leq \lambda_i$ ; también, si  $t_i \leq 0$  entonces  $\alpha t_i \leq 0 \leq \lambda_i$ . En definitiva, los números  $\beta_i := \lambda_i - \alpha t_i$  son no negativos, tienen la propiedad de que  $\sum_{i=1}^r \beta_i = 1$  y además

$$x = x - \alpha \cdot 0 = \sum_{i=1}^r \lambda_i x_i - \alpha \sum_{i=1}^r t_i x_i = \sum_{i=1}^r \beta_i x_i = \sum_{i=1}^{r-1} \beta_i x_i,$$

donde la última igualdad se debe a que  $\beta_r = 0$ . Esto muestra que  $x$  es combinación convexa de los  $r-1$  elementos  $x_1, \dots, x_{r-1}$  de  $X$ .  $\square$

Una consecuencia importante de este teorema es que la cápsula convexa de todo conjunto compacto es un conjunto compacto.

**Corolario 2.5.** *Si  $K$  es un subconjunto compacto y no vacío de  $\mathbb{R}^n$  entonces  $\text{co}(K)$  es un conjunto compacto. En particular,  $\text{co}(\{v_1, \dots, v_r\})$  es compacto para cualesquiera  $v_1, \dots, v_r \in \mathbb{R}^n$ .*

PRUEBA. Mostraremos que toda sucesión en  $\text{co}(K)$  posee una subsucesión convergente en  $\text{co}(K)$ . Sea  $(z_k) \subset \text{co}(K)$ . Por el teorema de Carathéodory, cada  $z_k$  se expresa como  $z_k = \sum_{i=1}^{n+1} \lambda_i^k v_i^k$ , donde  $v_i^k \in K$  y  $\sum_{i=1}^{n+1} \lambda_i^k = 1$ . La sucesión acotada

$$(\lambda_1^k, \dots, \lambda_{n+1}^k, v_1^k, \dots, v_{n+1}^k)_k$$

posee una subsucesión que converge a algún  $(\lambda_1, \dots, \lambda_{n+1}, v_1, \dots, v_{n+1})$ , donde cada  $\lambda_i$  está en  $[0, 1]$  y cada  $v_i$  está en  $K$ , ya que  $[0, 1]$  y  $K$  son cerrados. Como  $\sum_{i=1}^{n+1} \lambda_i = 1$  (pues  $\sum_{i=1}^{n+1} \lambda_i^k = 1$  para todo  $k$ ), esto muestra que  $\sum_{i=1}^{n+1} \lambda_i v_i \in \text{co}(K)$  es el límite de una subsucesión de  $(z_k)$ .  $\square$

**Observación 2.6.** Un conjunto compacto es un conjunto que es cerrado y acotado. La cápsula convexa de un conjunto acotado es acotado, mas la cápsula convexa de un conjunto cerrado no siempre es cerrado. En efecto, el subconjunto cerrado  $X = \{(0, 0)\} \cup \{(x, y) \in \mathbb{R}^2 \mid x > 0, y \geq 1/x\}$  de  $\mathbb{R}^2$  tiene como cápsula convexa a

$$\text{co}(X) = \{(x, y) \in \mathbb{R}^2 \mid x, y > 0\} \cup \{(0, 0)\},$$

que no es cerrado.

Ahora veremos una propiedad especial de los conjuntos convexos, que intuitivamente nos dice que el segmento abierto que une un punto del interior de un conjunto convexo y otro de su clausura está contenido en el interior de dicho conjunto. Para ello definamos los segmentos abiertos y semiabiertos

$$\begin{aligned} (x, y) &:= \{\lambda x + (1 - \lambda)y \mid 0 < \lambda < 1\}, \\ [x, y) &:= \{\lambda x + (1 - \lambda)y \mid 0 < \lambda \leq 1\}. \end{aligned}$$

que unen  $x$  e  $y$ .

**Proposición 2.7.** *Sea  $C$  un subconjunto convexo y no vacío de  $\mathbb{R}^n$ . Si  $x \in \text{int}(C)$  e  $y \in \overline{C}$ , entonces  $[x, y) \subset \text{int}(C)$ .*

PRUEBA. Sea  $(y_k)$  una sucesión en  $C$  que converge a  $y$ , y tomemos una bola  $B(x, \varepsilon)$  contenida en  $C$ . Fijemos  $t \in (0, 1]$  y escribamos  $z_t := tx + (1 - t)y$ . Para cada  $k \in \mathbb{N}$  escribimos  $z_t^k := tx + (1 - t)y_k$ . Luego la convexidad de  $C$  implica que  $B(z_t^k, t\varepsilon) \subset C$  para todo  $k$  pues si  $z \in B(z_t^k, t\varepsilon)$  entonces  $\|z - (tx + (1 - t)y_k)\| < t\varepsilon$ , esto es,  $\frac{z}{t} + (1 - \frac{1}{t})y_k \in B(x, \varepsilon) \subset C$  y por tanto

$$z = t\left(\frac{z}{t} + (1 - \frac{1}{t})y_k\right) + (1 - t)y_k \in C.$$

Como la sucesión  $(z_t^k)_k$  converge a  $z_t$ , para  $k$  suficientemente grande  $z_t$  estará en la bola  $B(z_t^k, t\varepsilon)$ , que está contenida en  $C$ , lo que implica que  $z_t \in \text{int}(C)$ . Esto muestra que  $[x, y) \subset C$ .  $\square$

**Corolario 2.8.** *Sea  $C$  un subconjunto convexo de  $\mathbb{R}^n$  tal que  $\text{int}(C) \neq \emptyset$ . Entonces  $\text{int}(\overline{C}) = \text{int}(C)$  y  $\overline{C} = \overline{\text{int}(C)}$ .*

PRUEBA. Claramente,  $\text{int}(C) \subset \text{int}(\overline{C})$  y  $\overline{\text{int}(C)} \subset \overline{C}$ . Consideremos  $x \in \text{int}(C)$ . Dado  $y \in \text{int}(\overline{C})$ , existe  $t > 0$  suficientemente pequeño tal que  $z := y + t(y - x)$  está en  $\overline{C}$ , de modo que  $y = \frac{1}{1+t}z + \frac{t}{1+t}x$  está en  $(x, z)$ , que por la Proposición 2.7 está contenido en  $\text{int}(C)$ . Luego,  $y \in \text{int}(C)$  y obtenemos  $\text{int}(C) = \text{int}(\overline{C})$ .

Sea ahora  $y \in \overline{C}$ . Por la Proposición 2.7 tenemos  $[x, y) \subset \text{int}(C)$ , y como hay puntos de  $[x, y)$  arbitrariamente cercanos a  $y$ , podemos inferir que  $y \in \text{int}(C)$ . Así,  $\overline{C} = \overline{\text{int}(C)}$ .  $\square$

La igualdad  $\overline{C} = \overline{\text{int}(C)}$  en el corolario anterior no se cumple si el conjunto en cuestión tiene interior vacío. Por ejemplo, si  $C$  es una recta en  $\mathbb{R}^2$  entonces es un conjunto convexo tal que

$$\overline{C} = C \neq \emptyset = \overline{\emptyset} = \overline{\text{int}(C)}.$$

Sin embargo,  $\text{int}(\overline{C}) = \text{int}(C)$  se cumple incluso cuando  $\text{int}(C) = \emptyset$ . Queriendo ver esto, introducimos la siguiente definición.

Decimos que  $x_0, x_1, \dots, x_r \in \mathbb{R}^n$  son **afínmente independientes (a.i.)** si los vectores  $x_1 - x_0, \dots, x_r - x_0$  son linealmente independientes (l.i.). Esta definición es buena, es decir, si  $x_1 - x_0, \dots, x_r - x_0$  son l.i. entonces también  $\{x_i - x_j\}_{0 \leq i \leq r, i \neq j}$  son l.i. para cualquier  $j \in \{1, \dots, r\}$ . En efecto, si los escalares  $\{c_i\}_{0 \leq i \leq r, i \neq j}$  satisfacen la relación  $\sum_{0 \leq i \leq r, i \neq j} c_i(x_i - x_j) = 0$  entonces

$$\sum_{\substack{i=1 \\ i \neq j}}^r c_i(x_i - x_0) + \sum_{\substack{i=0 \\ i \neq j}}^r (-c_i)(x_j - x_0) = \sum_{\substack{i=0 \\ i \neq j}}^r c_i(x_i - x_j) = 0,$$

de donde, como  $x_1 - x_0, \dots, x_r - x_0$  son l.i., obtenemos  $c_i = 0$  para cada  $i$ .

Un  **$n$ -simplex** en  $\mathbb{R}^n$  es un conjunto  $\{x_0, \dots, x_n\}$ , donde  $x_0, \dots, x_n$  son a.i. Al ser  $x_1 - x_0, \dots, x_n - x_0$  l.i., todo  $x \in \mathbb{R}^n$  puede expresarse de manera única como combinación lineal de los vectores  $x_i - x_0$ , lo que implica que existen escalares únicos  $t_0, \dots, t_n$  tales que  $\sum_{i=0}^n t_i x_i = x$  y  $\sum_{i=0}^n t_i = 1$ . Definamos  $\|\cdot\|_* : \mathbb{R}^n \rightarrow \mathbb{R}$  así: si  $x = \sum_{i=1}^n \lambda_i(x_i - x_0)$  entonces  $\|x\|_* = \max_{1 \leq i \leq n} |\lambda_i|$ . La función  $\|\cdot\|_*$  está bien definida y es una norma. Por el teorema de equivalencia de normas en  $\mathbb{R}^n$  tenemos que una sucesión  $(x_k)$  dada por  $x_k = \sum_{i=1}^n \lambda_i^k(x_i - x_0)$  converge a  $x = \sum_{i=1}^n \lambda_i(x_i - x_0)$  en el sentido usual (con la norma euclidiana) si y solo si converge con la norma  $\|\cdot\|_*$ , esto es, si y solo si  $\lim_{k \rightarrow \infty} \lambda_i^k = \lambda_i$  para cada  $i$ . Una consecuencia de esto es que toda sucesión  $(x_k)$  dada por  $x_k = \sum_{i=0}^n \lambda_i^k x_i$ , con  $\sum_{i=0}^n \lambda_i^k = 1$ , converge a  $x = \sum_{i=0}^n \lambda_i x_i$ , con  $\sum_{i=0}^n \lambda_i = 1$ , si y solo si cada sucesión  $(\lambda_i^k)_k$  converge a  $\lambda_i$ .

Dado el  $n$ -simplex  $S = \{x_0, \dots, x_n\}$  consideremos el conjunto

$$U := \left\{ \sum_{i=0}^n t_i x_i \mid \sum_{i=0}^n t_i = 1, 0 < t_i < 1, \forall i = 0, \dots, n \right\}, \quad (2.1)$$

que está contenido en  $\text{co}(S)$ . Afirmamos que  $U$  es abierto. Para probar ello mostremos que  $\mathbb{R}^n \setminus U$  es cerrado. Tomemos una sucesión  $(z_k)$  en  $\mathbb{R}^n \setminus U$  que converge a  $z \in \mathbb{R}^n$ . Entonces, cada  $z_k$  se expresa como  $z_k = \sum_{i=0}^n t_i^k x_i$  con  $\sum_{i=0}^n t_i^k = 1$ . Supongamos que  $z \in U$ , es decir, que  $z = \sum_{i=0}^n t_i x_i$  con  $0 < t_i < 1$  para cada  $i$ . Por

el resultado del párrafo anterior tenemos  $t_i^k \rightarrow t_i$  para todo  $i$ , de modo que si  $k$  es suficientemente grande se tiene  $0 < t_i^k < 1$  para cada  $i$ , lo que significa que  $z_k \in U$ , una contradicción. Luego,  $U$  es abierto.

Concluimos así que todo conjunto convexo que contiene algún  $n$ -simplex tiene interior no vacío.

Ahora establecemos el resultado análogo al Corolario 2.8 que buscábamos.

**Proposición 2.9.** *Sea  $C$  un subconjunto convexo de  $\mathbb{R}^n$ . Entonces  $\text{int}(C) = \text{int}(\overline{C})$ .*

PRUEBA. Por el Corolario 2.8, solo falta probar la proposición cuando  $\text{int}(C) = \emptyset$ . Supongamos esto y supongamos también que  $\text{int}(\overline{C}) \neq \emptyset$ . Entonces  $\overline{C}$  contiene alguna bola, que a su vez contendrá algún  $n$ -simplex  $S'$ . Tomando puntos de  $C$  suficientemente cercanos a los puntos de  $S'$  obtenemos un  $n$ -simplex  $S$  compuesto de puntos de  $C$ . Esto significa que  $\text{int}(C) \neq \emptyset$ , una contradicción.  $\square$

Dados un subconjunto no vacío  $C$  de  $\mathbb{R}^n$  y un vector  $a \in \mathbb{R}^n$ , decimos que un punto  $\bar{x} \in C$  es una **proyección** de  $a$  en  $C$  si  $d_C(a) = \|a - \bar{x}\|$ .

**Proposición 2.10.** *Si  $C \subset \mathbb{R}^n$  es cerrado, convexo y no vacío, entonces para todo vector  $a \in \mathbb{R}^n$  existe una única proyección de  $a$  en  $C$ .*

PRUEBA. Escribamos  $d := d_C(a)$ . Veamos primero la unicidad de la proyección. Supongamos que  $\bar{x}, x' \in C$  son tales que  $\|\bar{x} - a\| = \|x' - a\| = d$  y sea  $z := \frac{\bar{x} + x'}{2}$ , que estará en  $C$  porque  $C$  es convexo. Entonces  $2z - 2a = (\bar{x} - a) + (x' - a)$ ,  $\bar{x} - x' = (\bar{x} - a) - (x' - a)$  y se tiene

$$4d^2 + \|\bar{x} - x'\|^2 \leq \|2z - 2a\|^2 + \|\bar{x} - x'\|^2 = 2\|\bar{x} - a\|^2 + 2\|x' - a\|^2 = 4d^2,$$

lo cual implica  $\bar{x} = x'$ . Mostremos ahora la existencia de la proyección. Sea  $(x_k)$  una sucesión en  $C$  tal que  $\|a - x_k\| \rightarrow d$ . Como  $(x_k)$  es una sucesión acotada, podemos suponer que converge a algún vector  $\bar{x} \in \mathbb{R}^n$ , que estará en  $C$  porque  $C$  es cerrado. Luego  $\|a - \bar{x}\| = \lim_{k \rightarrow \infty} \|a - x_k\| = d$ .  $\square$

Así, cuando  $C \subset \mathbb{R}^n$  es convexo, cerrado y no vacío podemos definir una función  $P_C : \mathbb{R}^n \rightarrow C$  que asocia a cada  $a \in \mathbb{R}^n$  con su proyección  $P_C(a)$  en  $C$ .

**Proposición 2.11.** *Sea  $C$  un subconjunto convexo, cerrado y no vacío de  $\mathbb{R}^n$ . Sean  $a \in \mathbb{R}^n$  y  $\bar{x} \in C$ . Entonces*

- (1)  $\bar{x} = P_C(a)$  si y solo si  $\langle x - \bar{x}, a - \bar{x} \rangle \leq 0$  para todo  $x \in C$ ;
- (2) la función  $P_C$  es continua;

PRUEBA. (1) Supongamos que  $\bar{x} = P_C(a)$ . Sea  $x \in C$ . Para cada  $\lambda \in (0, 1)$  el punto  $x_\lambda := \lambda x + (1 - \lambda)\bar{x}$  está en  $C$  (pues  $C$  es convexo) y en consecuencia

$$\begin{aligned} \|\bar{x} - a\|^2 &\leq \|x_\lambda - a\|^2 = \|\lambda(x - a) + (1 - \lambda)(\bar{x} - a)\|^2 \\ &= \lambda^2\|x - a\|^2 + (1 - \lambda)^2\|\bar{x} - a\|^2 + 2\lambda(1 - \lambda)\langle x - a, \bar{x} - a \rangle. \end{aligned}$$

Como  $\|\bar{x} - a\|^2 - \langle x - a, \bar{x} - a \rangle = \langle x - \bar{x}, a - \bar{x} \rangle$  lo anterior significa que

$$2\lambda(1 - \lambda)\langle x - \bar{x}, a - \bar{x} \rangle \leq \lambda^2\|x - a\|^2,$$

para todo  $\lambda \in (0, 1)$ . Simplificando  $\lambda$  en ambos lados de la desigualdad y haciendo  $\lambda \rightarrow 0$ , obtenemos  $\langle x - \bar{x}, a - \bar{x} \rangle \leq 0$ .

Recíprocamente, si  $\bar{x} \in C$  es tal que  $\langle x - \bar{x}, a - \bar{x} \rangle \leq 0$  para todo  $x \in C$  entonces

$$\|\bar{x} - a\|^2 \leq \|x - \bar{x}\|^2 + \|\bar{x} - a\|^2 + 2\langle x - \bar{x}, \bar{x} - a \rangle = \|x - a\|^2$$

para cada  $x \in C$ , esto es,  $\bar{x} = P_C(a)$ .

(2) Dados  $x, y \in \mathbb{R}^n$ , de la parte (1) tenemos

$$\langle x - P_C(x), P_C(y) - P_C(x) \rangle \leq 0 \quad \text{y} \quad \langle y - P_C(y), P_C(x) - P_C(y) \rangle \leq 0.$$

Puesto que  $\langle x - P_C(x), P_C(x) - P_C(y) \rangle - \langle y - P_C(y), P_C(x) - P_C(y) \rangle$  es igual a  $\langle x - y, P_C(x) - P_C(y) \rangle - \|P_C(x) - P_C(y)\|^2$ , lo anterior nos da

$$\|P_C(x) - P_C(y)\|^2 \leq \langle x - y, P_C(x) - P_C(y) \rangle \leq \|x - y\| \cdot \|P_C(x) - P_C(y)\|,$$

y por tanto  $\|P_C(x) - P_C(y)\| \leq \|x - y\|$  (incluso cuando  $x = y$ ). Así,  $P_C$  es lipshitziana y en particular continua.  $\square$

Un **hiperplano** en  $\mathbb{R}^n$  es un conjunto de la forma  $H = \{x \in \mathbb{R}^n \mid \langle v, x \rangle = b\}$ , donde  $v \in \mathbb{R}^n$  es no nulo y  $b \in \mathbb{R}$ .

Decimos que dos subconjuntos  $A$  y  $B$  de  $\mathbb{R}^n$  están **separados** por el hiperplano  $H = \{x \in \mathbb{R}^n \mid \langle v, x \rangle = b\}$  si

$$\langle v, x \rangle \leq b \leq \langle v, y \rangle \quad \text{para cualesquiera } x \in A \text{ y } y \in B.$$

Decimos que  $A$  y  $B$  están **separados fuertemente** por el hiperplano  $H$  si existe  $\delta > 0$  tal que

$$\langle v, x \rangle \leq b - \delta < b + \delta \leq \langle v, y \rangle \quad \text{para cualesquiera } x \in A \text{ y } y \in B. \quad (2.2)$$

Por la continuidad del producto interno tenemos que si dos conjuntos  $A$  y  $B$  están separados (fuertemente) por un hiperplano  $H$  entonces sus clausuras  $\bar{A}$  y  $\bar{B}$  también lo están.

No siempre es posible separar dos conjuntos con un hiperplano (por ejemplo, si  $x \in \text{int}(M)$  entonces  $M$  y  $\{x\}$  no pueden separarse). No obstante, cuando los conjuntos  $A$  y  $B$  son convexos, bajo ciertas condiciones se puede asegurar la existencia de un hiperplano que los separa. Es por ello que la noción de separabilidad está muy presente en el análisis convexo; además, los hiperplanos tienen la propiedad especial de que su complemento es la unión de dos conjuntos convexos disjuntos. El siguiente teorema es un resultado clásico de análisis convexo.

**Teorema 2.12** (Teorema de Separación). *Sean  $A$  y  $B$  dos subconjuntos convexos y no vacíos de  $\mathbb{R}^n$ . Entonces bajo cualquiera de las siguientes hipótesis existe un hiperplano que separa  $A$  y  $B$ .*

- (1) El conjunto  $B$  es unitario, digamos  $B = \{\bar{x}\}$ , y además  $\bar{x} \notin \text{int}(A)$ .
- (2) Los conjuntos  $A$  y  $B$  no tienen puntos en común.
- (3) El interior de  $A$  es no vacío y  $B \cap \text{int}(A) = \emptyset$ .

PRUEBA. (1) Como  $\bar{x} \notin \text{int}(A)$ , es claro que  $\bar{x} \notin \text{int}(\bar{A})$  (Proposición 2.9). Luego existe una sucesión  $(x_k) \subset \mathbb{R}^n \setminus \bar{A}$  que converge a  $\bar{x}$ . Sea  $(z_k) \subset \bar{A}$  la sucesión definida por  $z_k = P_{\bar{A}}(x_k)$ . Entonces

$$\langle x_k - z_k, x - z_k \rangle \leq 0, \quad \forall x \in \bar{A}, \forall k \geq 1$$

y como  $\langle x_k - z_k, x - x_k \rangle \leq \langle x_k - z_k, x - x_k \rangle + \langle x_k - z_k, x_k - z_k \rangle = \langle x_k - z_k, x - z_k \rangle$ , se sigue que

$$\langle x_k - z_k, x - x_k \rangle \leq 0, \quad \forall x \in \bar{A}, \forall k \geq 1.$$

Por otro lado, como  $\|x_k - z_k\| > 0$  para todo  $k$  (pues  $x_k \notin \bar{A}$  y  $z_k \in \bar{A}$ ), podemos dividir ambos lados de la ecuación anterior entre  $\|x_k - z_k\|$  para obtener

$$\langle v_k, x - x_k \rangle \leq 0, \quad \forall x \in \bar{A}, \forall k \geq 1, \quad (2.3)$$

donde  $v_k := \frac{x_k - z_k}{\|x_k - z_k\|}$  define una sucesión acotada que, sin pérdida de generalidad, podemos suponer que converge a algún vector  $v$  de norma 1. Tomando límite cuando  $k \rightarrow \infty$  obtenemos  $\langle v, x - \bar{x} \rangle \leq 0$  para todo  $x \in \bar{A}$ , lo cual prueba (1).

(2) El conjunto convexo  $C = A - B = \{a - b \mid a \in A, b \in B\}$  no contiene a 0 pues  $A$  y  $B$  son disjuntos. Luego, por la parte (1) existe un vector  $v \in \mathbb{R}^n$  no nulo tal que  $0 \leq \langle v, c \rangle$  para todo  $c \in C$ . Esto prueba (2).

(3) Por la parte (2), existe un hiperplano  $H$  que separa  $\text{int}(A)$  y  $B$ . Luego  $H$  también separa  $\text{int}(\bar{A})$  y  $B$ . Finalmente,  $\bar{A} = \overline{\text{int}(A)}$  por el Corolario 2.8.  $\square$

**Teorema 2.13** (Separación Fuerte). *Sean  $A$  y  $B$  dos subconjuntos convexos y no vacíos de  $\mathbb{R}^n$ . Entonces bajo cualquiera de las siguientes hipótesis existe un hiperplano que separa fuertemente  $A$  y  $B$ .*

- (1) La distancia  $d(A, B)$  entre  $A$  y  $B$  es positiva.
- (2)  $A$  es cerrado,  $B$  es compacto y  $A \cap B = \emptyset$ .

PRUEBA. (1) Sea  $d := d(A, B) > 0$ . Como el conjunto convexo  $C := A - B$  no tiene puntos en común con la bola  $B(0, d)$ , la clausura  $\bar{C}$  de  $C$  tampoco los tiene. Luego, si  $v$  es la proyección de 0 en  $\bar{C}$  entonces  $\|v\| \geq d$  y, por la Proposición 2.11,

$$\langle c - v, 0 - v \rangle \leq 0 \quad \text{para todo } c \in \bar{C},$$

lo cual implica que  $\|v\|^2 + \langle v, y \rangle \leq \langle v, x \rangle$  para todo  $x \in A$  e  $y \in B$ . Definiendo  $b := \sup_{y \in B} \langle y, v \rangle + \frac{\|v\|^2}{2}$  y  $\delta := \frac{\|v\|^2}{2}$  vemos que (2.2) se cumple.

(2) La función distancia  $d_A$  es continua y por tanto alcanza un mínimo, digamos  $y \in B$ , sobre el conjunto compacto  $B$ . Luego  $d(A, B) = d_A(y)$  es positiva pues  $A \cap B = \emptyset$ , y el resultado se sigue de la parte (1).  $\square$

Un conjunto  $C \subset \mathbb{R}^n$  es un **cono** si  $\lambda x \in C$  para todo  $x \in C$  y todo  $\lambda \geq 0$ . Un cono no es necesariamente convexo. Tanto la intersección arbitraria como la unión arbitraria de conos es un cono.

**Ejemplo 2.14.** El espacio  $\mathbb{R}^n$  es un cono, así como cualquier semirecta del tipo  $\{\lambda z \mid \lambda \geq 0\}$  para cada  $z \in \mathbb{R}^n$  fijo. Los conjuntos  $E$  y  $F$  de la Figura 2.3 son conos convexos,  $E$  es cerrado y  $F$  es abierto.

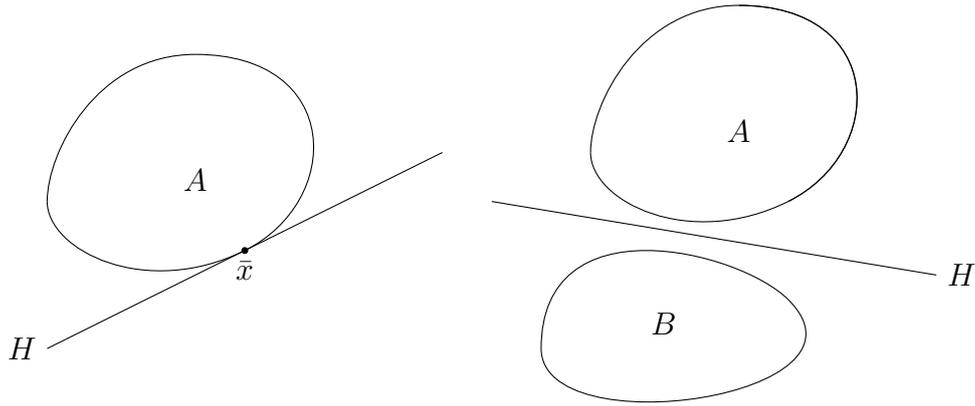


Figura 2.2: Ilustración gráfica de los teoremas de separación. A la izquierda, el hiperplano  $H$  separa al punto  $\bar{x}$  del conjunto convexo  $A$ . A la derecha, el hiperplano  $H$  separa fuertemente a los conjuntos convexos y disjuntos  $A$  y  $B$ .

Sea  $X \subset \mathbb{R}^n$ . El conjunto

$$\text{con}(X) := \left\{ \sum_{i=1}^r t_i x_i \mid r \in \mathbb{N}, (t_i)_{i=1}^r \subset [0, \infty), (x_i)_{i=1}^r \subset X \right\}$$

es un cono convexo que satisface la siguiente propiedad: está contenido en cada cono convexo  $C \subset \mathbb{R}^n$  tal que  $X \subset C$ . En otras palabras,  $\text{con}(X)$  es el menor cono convexo que contiene a  $X$ , o, equivalentemente,

$$\text{con}(X) = \bigcap \{ C \mid C \text{ es un cono convexo tal que } X \subset C \}.$$

Como la intersección de conos es un cono, y como la intersección de conjuntos convexos (cerrados) es un conjunto convexo (cerrado), el conjunto

$$\overline{\text{con}}(X) := \bigcap \{ C \mid C \text{ es un cono convexo y cerrado tal que } X \subset C \}$$

es el menor conjunto convexo y cerrado que contiene a  $X$ .

**Lema 2.15.** Dado  $X \subset \mathbb{R}^n$ , tenemos

$$\overline{\text{con}}(X) = \overline{\text{con}(X)}.$$

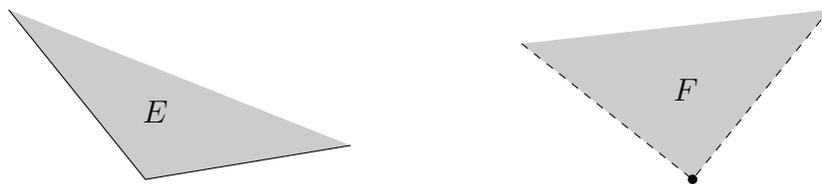


Figura 2.3: Ejemplos de conos convexos, uno cerrado y otro no cerrado

Dado un conjunto no vacío  $C$  en  $\mathbb{R}^n$ , el conjunto

$$C^* := \{ z \in \mathbb{R}^n \mid \langle z, x \rangle \leq 0, \forall x \in C \}$$

es llamado el **cono polar** de  $C$ . Evidentemente,  $C^*$  es un cono convexo y cerrado que contiene a 0. Además, si  $C_1 \subset C_2$  entonces  $C_2^* \subset C_1^*$ .

**Ejemplo 2.16.** El cono polar de  $\mathbb{R}^n$  es  $\{0\}$ ; más generalmente, si  $0 \in \text{int}(C)$  entonces  $C^* = \{0\}$ . El cono polar del primer cuadrante en  $\mathbb{R}^2$  es el tercer cuadrante.

Una propiedad básica e importante de los conos polares es la siguiente proposición.

**Proposición 2.17.** *Para todo subconjunto no vacío  $C$  de  $\mathbb{R}^n$ , se cumple que*

$$C^* = \text{co}(C)^* = (\overline{C})^* = \text{con}(C)^* = \overline{\text{con}(C)}^*.$$

PRUEBA. Ya sabemos que  $\text{con}(C)^* \subset \text{co}(C)^* \subset C^*$  y que  $(\overline{C})^* \subset C^*$ . Sea  $z \in C^*$ . Como  $\langle z, x \rangle \leq 0$  para todo  $x \in C$ , tenemos  $\langle z, x \rangle \leq 0$  para todo  $x \in \overline{C}$  (por la continuidad del producto interno); también, si  $x \in \text{con}(C)$  entonces  $x = \sum_{i=1}^r t_i x_i$ , donde  $t_i \geq 0$  y  $x_i \in C$ , y por tanto  $\langle z, x \rangle = \sum_i t_i \langle z, x_i \rangle \leq 0$ . Así,  $z \in (\overline{C})^* \cap \text{con}(C)^*$ , y concluimos que

$$C^* = \text{co}(C)^* = (\overline{C})^* = \text{con}(C)^*.$$

Por último,  $\overline{\text{con}(C)} = \overline{\text{con}(C)}$  implica  $\overline{\text{con}(C)}^* = \text{con}(C)^* = C^*$ .  $\square$

**Proposición 2.18.** *Si  $C \subset \mathbb{R}^n$  es un cono no vacío, convexo y cerrado, entonces  $C^{**} = C$ .*

PRUEBA. La inclusión  $C \subset C^{**}$  es clara. Sea  $z \in C^{**}$  y consideremos su proyección  $x = P_C(z)$  en  $C$ . Queremos mostrar que  $z = x$ . Utilizando la Proposición 2.11 vemos que  $\langle z - x, 0 - x \rangle$  y  $\langle z - x, 2x - x \rangle$  son menores o iguales que 0 ya que  $0, 2x \in C$ , es decir,  $\langle z - x, x \rangle = 0$ . Luego,  $\langle z - x, y \rangle = \langle z - x, y - x \rangle \leq 0$  para todo  $y \in C$ , lo cual significa  $z - x \in C^*$ . Como  $z \in C^{**}$ , tenemos  $\langle z - x, z \rangle \leq 0$ , de donde  $x = z$ , ya que

$$\|z - x\|^2 = \langle z - x, z - x \rangle \leq 0. \quad \square$$

**Proposición 2.19.** *Si  $C$  es un subconjunto no vacío de  $\mathbb{R}^n$ , entonces  $C^{**} = \overline{\text{con}(C)}$ .*

PRUEBA. Consecuencia directa de la Proposiciones 2.17 y 2.18.  $\square$

**Proposición 2.20.** *Supongamos que  $v_1, \dots, v_r$  son vectores no nulos en  $\mathbb{R}^n$ . Entonces el cono convexo*

$$\text{con}(\{v_1, \dots, v_r\}) = \left\{ \sum_{i=1}^r t_i v_i \mid t_i \geq 0, \forall i = 1, \dots, r \right\}$$

*es cerrado. En particular,*

$$\{x \in \mathbb{R}^n \mid \langle x, v_i \rangle \leq 0, \forall i = 1, \dots, r\}^* = \text{con}(\{v_1, \dots, v_r\}). \quad (2.4)$$

PRUEBA. Probaremos que  $\text{co}(\{v_1, \dots, v_r\})$  es cerrado haciendo inducción sobre  $r$ . Si  $r = 1$ , es claro que  $\text{con}(\{v_1\}) = \{tv_1 \mid t \geq 0\}$  es cerrado. Supongamos que la afirmación es cierta para  $r - 1$ , donde  $r > 1$ . Sean  $v_1, \dots, v_r$  vectores no nulos en  $\mathbb{R}^n$  y escribamos  $C := \text{con}(\{v_1, \dots, v_r\})$ . Podemos asumir que  $\|v_i\| = 1$  para cada  $i$ . Debido a que  $-v_1, \dots, -v_r \in C$  implica que  $C$  es el subespacio vectorial (cerrado) de  $\mathbb{R}^n$  generado por  $v_1, \dots, v_r$ , podemos suponer además, reindexando los vectores si fuera necesario, que  $-v_r \notin C$ . El cono convexo

$$C' := \text{con}(\{v_1, \dots, v_{r-1}\})$$

es cerrado por hipótesis, y en particular  $\{x \in C' \mid \|x\| = 1\}$  es compacto. Luego, existe  $\hat{x} \in C'$  de norma 1 tal que

$$\gamma := \min_{x \in C', \|x\|=1} \langle x, v_r \rangle = \langle \hat{x}, v_r \rangle$$

y tal que  $\hat{x} \neq -v_r$ , pues  $-v_r \notin C$ . Por la desigualdad de Cauchy-Schwarz tenemos

$$\gamma = \langle \hat{x}, v_r \rangle > -\|\hat{x}\| \|v_r\| = -1,$$

donde la desigualdad estricta se debe a que  $\hat{x} \neq -v_r$ . Además,  $\gamma \|x\| \leq \langle x, v_r \rangle$  para todo  $x \in C'$ .

Mostraremos ahora que si una sucesión  $(z_k)$  en  $C$  converge a  $z \in \mathbb{R}^n$ , entonces  $z \in C$ . Para cada  $k$  existen  $z'_k \in C'$  y  $t_k \geq 0$  tales que  $z_k = z'_k + t_k v_r$ . La sucesión  $(z_k)$  es acotada, de modo que  $\gamma + 1 > 0$  y

$$\begin{aligned} \|z_k\|^2 &= \|z'_k + t_k v_r\|^2 \\ &= \|z'_k\|^2 + t_k^2 + 2t_k \langle z'_k, v_r \rangle \\ &\geq \|z'_k\|^2 + t_k^2 + 2t_k \gamma \|z'_k\| \\ &= (\|z'_k\| - t_k)^2 + 2(\gamma + 1)t_k \|z'_k\| \end{aligned}$$

implican que  $(z'_k)$  y  $(t_k)$  son acotadas. Podemos por tanto suponer que  $z'_k$  converge a algún  $z' \in C'$  y que  $t_k$  converge a algún  $t \geq 0$ . En consecuencia,  $z = \lim z_k = \lim(z'_k + t_k v_r) = z' + t v_r \in C$ , y concluimos que  $C$  es cerrado. Esto completa el proceso inductivo.

Finalmente, una aplicación directa de la Proposición 2.18 nos da (2.4).  $\square$

De modo análogo al cono polar, dados un subconjunto no vacío  $C$  de  $\mathbb{R}^n$  y un vector  $x \in \mathbb{R}^n$ , el conjunto

$$N(x, C) := \{z \in \mathbb{R}^n \mid \langle z - x, y \rangle \leq 0, \forall y \in C\}$$

es llamado **cono normal** de  $C$  en el punto  $x$ <sup>3</sup>. Notemos que  $N(x, C) = (C - x)^*$ . Luego obtenemos para conos normales un resultado similar a la Proposición 2.17.

**Proposición 2.21.** *Para todo subconjunto no vacío  $C$  de  $\mathbb{R}^n$  y todo  $x \in \mathbb{R}^n$  se cumple*

$$N(x, C) = N(x, \text{co}(C)) = N(x, \overline{C}).$$

Otro resultado concerniente a conos normales es la siguiente proposición.

**Proposición 2.22.** *Sea  $C$  un subconjunto de  $\mathbb{R}^n$  con interior no vacío y sea  $x \in \mathbb{R}^n$ . Escribamos  $K := N(x, C)$ . Entonces  $K \cap (-K) = \{0\}$ .*

<sup>3</sup>Aquí adoptamos estas definiciones, no sin dejar de señalar que, dependiendo de la referencia que se use, el término *cono normal* puede significar cosas distintas. Por ejemplo, en [5], dados  $C \subset \mathbb{R}^n$  y  $x \in C$  se llama cono normal a un conjunto obtenido a partir del cono tangente  $T_C(x)$  (conjunto previamente definido en [5]).

PRUEBA. Supongamos que existe  $z \in K \cap (-K)$  no nulo. Entonces  $z$  y  $-z$  están en  $K$  de modo que para cualquier  $y \in C$  tenemos  $\langle z, y - x \rangle \leq 0$  y  $\langle -z, y - x \rangle \leq 0$ , es decir,  $\langle z, y - x \rangle = 0$ . Luego  $C$  está contenido en el conjunto  $\{y \in \mathbb{R}^n \mid \langle z, y - x \rangle = 0\}$ , que tiene interior vacío pues es la traslación de un subespacio de dimensión  $n-1$ . Esto implica que  $\text{int}(C) = \emptyset$ , una contradicción. Concluimos que  $K \cap (-K) = \{0\}$ .  $\square$

Definamos el cono tangente de  $C$  en  $x \in C$ . Un vector  $v \in \mathbb{R}^n$  es un elemento del **cono tangente** de  $C$  en  $x \in C$ , denotado por  $T(x, C)$ , si existen sucesiones  $(x_k)$  en  $C$  y  $(t_k)$  en  $(0, \infty)$  tales que

$$x_k \rightarrow x, \quad t_k \rightarrow 0, \quad \frac{x_k - x}{t_k} \rightarrow v. \quad (2.5)$$

Es claro que  $0 \in T(x, C)$  siempre ocurre. Afirmamos que un vector no nulo  $v \in \mathbb{R}^n$  pertenece a  $T(x, C)$  si y solo si existe una sucesión  $(x_k)$  en  $C$  tal que

$$x_k \neq x \text{ para todo } k, \quad x_k \rightarrow x, \quad \frac{x_k - x}{\|x_k - x\|} \rightarrow \frac{v}{\|v\|}. \quad (2.6)$$

En efecto, si  $(x_k) \subset C$  es tal que (2.6) se cumple, escogiendo  $t_k := \frac{\|x_k - x\|}{\|v\|} > 0$  vemos que  $(x_k)$  y  $(t_k)$  verifican (2.5).

Recíprocamente, si suponemos que  $(x_k) \subset C$  y  $(t_k) \subset (0, \infty)$  verifican (2.5), entonces  $\frac{x_k - x}{t_k} \rightarrow v \neq 0$  implicará que  $x_k \neq x$  para  $k$  suficientemente grande, de modo que podemos suponer que  $x_k \neq x$  para todo  $k$ ; luego  $\frac{\|x_k - x\|}{t_k} \rightarrow \|v\|$ , y tomando límite en  $\frac{x_k - x}{\|x_k - x\|} = \frac{(x_k - x)/t_k}{\|x_k - x\|/t_k}$  obtenemos  $\frac{x_k - x}{\|x_k - x\|} \rightarrow \frac{v}{\|v\|}$ .

Así, los elementos de  $T(x, C)$  son aquellos vectores  $v \in \mathbb{R}^n$  que son *tangentes* a  $C$  en  $x$ , es decir, existe una sucesión  $(x_k)$  en  $C$  que converge a  $x$  y tal que la dirección de  $x_k - x$  converge a la dirección de  $v$ . La condición (2.6) ilustra mejor esta propiedad.

## 2.2. Funciones convexas

Dados una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  y un número  $\lambda \in \mathbb{R}$  definimos los conjuntos

$$\begin{aligned} \text{epi } f &:= \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq t\}, \\ S_\lambda(f) &:= \{x \in \mathbb{R}^n \mid f(x) \leq \lambda\}, \\ S_\lambda^<(f) &:= \{x \in \mathbb{R}^n \mid f(x) < \lambda\}, \end{aligned}$$

que son denominados el **epígrafo** de  $f$ , el **subnivel**  $\lambda$  de  $f$  y el **subnivel estricto**  $\lambda$  de  $f$ , respectivamente. Si no hubiera confusión respecto a la función  $f$  con la que estamos trabajando escribiremos  $S_\lambda$  y  $S_\lambda^<$  en lugar de  $S_\lambda(f)$  y  $S_\lambda^<(f)$  respectivamente. Definimos también

$$\arg \min f := \{x \in \mathbb{R}^n \mid f(x) \leq f(y), \forall y \in \mathbb{R}^n\},$$

que es el conjunto donde  $f$  alcanza su mínimo sobre  $\mathbb{R}^n$  (este conjunto puede ser vacío).

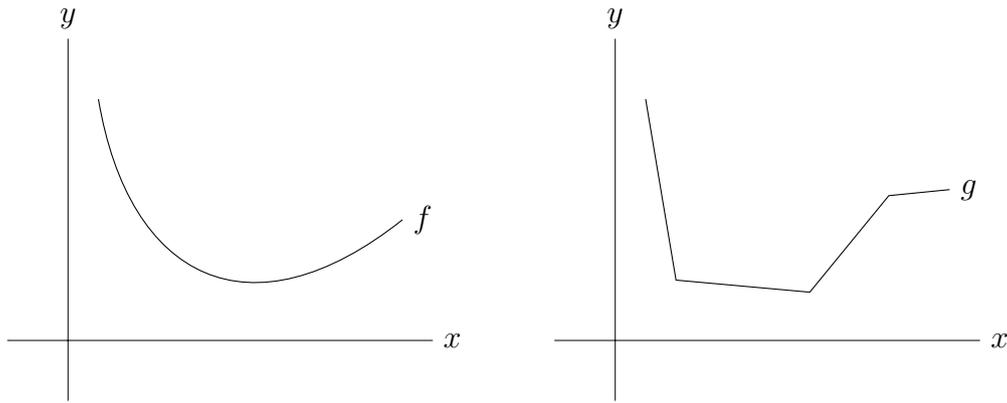


Figura 2.4: Ejemplo de funciones convexas y cuasiconvexas: la función  $f$  es convexa; y la función  $g$  es cuasiconvexa, pero no convexa.

Una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es **convexa** si para todo  $x, y \in \mathbb{R}^n$  y  $t \in [0, 1]$  se tiene

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

Si esta desigualdad fuera estricta para  $x \neq y$  y  $0 < t < 1$  entonces decimos que  $f$  es **estrictamente convexa**. Similarmente, decimos que  $f$  es **(estrictamente) cóncava** si  $-f$  es (estrictamente) convexa.

Gráficamente, una función  $f$  es convexa cuando la línea que une dos puntos cualesquiera del tipo  $(x, f(x))$  e  $(y, f(y))$  está por encima del gráfico de la función  $f$ . Esto equivale a que la convexidad del epígrafo de  $f$ . Así,

$f$  es convexa si y solo si  $\text{epi } f \subset \mathbb{R}^n \times \mathbb{R}$  es un conjunto convexo.

**Ejemplo 2.23** (La función distancia es convexa). Si  $C$  es un subconjunto convexo y no vacío de  $\mathbb{R}^n$  entonces la función distancia  $d_C : \mathbb{R}^n \rightarrow \mathbb{R}$ , que asocia a cada  $x \in \mathbb{R}^n$  su distancia  $d_C(x)$  a  $C$ , es convexa. En efecto, dados  $x, y \in \mathbb{R}^n$  y  $t \in [0, 1]$  tenemos que para cualesquiera  $z, w \in C$  el vector  $tz + (1 - t)w$  está en  $C$ , de modo que

$$d_C(tx + (1 - t)y) \leq \|tx + (1 - t)y - (tz + (1 - t)w)\| \leq t\|x - z\| + (1 - t)\|y - w\|$$

para todo  $z, w \in C$ , es decir,  $d_C(tx + (1 - t)y) \leq td_C(x) + (1 - t)d_C(y)$ .

En general no es cierto que la convexidad de una función  $f : C \rightarrow \mathbb{R}$  implique su continuidad. (Tómese por ejemplo la función convexa  $f : [0, 1] \rightarrow \mathbb{R}$  definida por  $f(x) = 0$  si  $x \in (0, 1]$  y  $f(0) = 1$ .) No obstante, sucede un hecho interesante; y es que toda función convexa definida en  $\mathbb{R}^n$  (o en un abierto de  $\mathbb{R}^n$ ) es continua.

**Proposición 2.24.** *Toda función convexa  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es continua.*

PRUEBA. Basta con mostrar que  $f$  es continua en 0. Para ello usaremos la norma del máximo, introducida en el Ejemplo 1.5. Sea el cubo unitario  $X = \{x \in \mathbb{R}^n \mid \|x\|_\infty \leq 1\}$  y sean  $e_1, \dots, e_{2^n}$  sus  $2^n$  esquinas. Cada  $e_i$  es un vector cuyas entradas son solo 1 y  $-1$ . Además, geoméricamente se observa que cada  $x \in X$  es combinación convexa

de los vectores  $e_i$ , es decir, existen  $\lambda_1, \dots, \lambda_{2^n}$  no negativos tales que  $x = \sum_{i=1}^{2^n} \lambda_i e_i$  y  $\sum_{i=1}^{2^n} \lambda_i = 1$ , de manera que

$$f(x) = f\left(\sum_{i=1}^{2^n} \lambda_i e_i\right) \leq \sum_{i=1}^{2^n} \lambda_i f(e_i) \leq \sum_{i=1}^{2^n} \lambda_i A = A,$$

donde  $A := \max\{f(e_1), \dots, f(e_{2^n})\}$ . Para ver que  $f$  es continua en 0 tomemos una sucesión  $(x_k) \subset \mathbb{R}^n$  que converge a 0. Sin pérdida de generalidad podemos asumir que  $(x_k) \subset X$  y que  $x_k \neq 0$  para todo  $k$ . Si definimos las sucesiones

$$y_k := \frac{x_k}{\|x_k\|_\infty} \quad \text{y} \quad z_k := -\frac{x_k}{\|x_k\|_\infty},$$

que están contenidas en  $X$ , observamos de inmediato que

$$\begin{aligned} x_k &= (1 - \|x_k\|_\infty) \cdot 0 + \|x_k\|_\infty \cdot y_k, \\ 0 &= \left(\frac{\|x_k\|_\infty}{\|x_k\|_\infty + 1}\right) z_k + \left(\frac{1}{\|x_k\|_\infty + 1}\right) x_k, \end{aligned}$$

lo cual, gracias a la convexidad de  $f$ , implica que

$$\begin{aligned} f(x_k) &\leq (1 - \|x_k\|_\infty)f(0) + \|x_k\|_\infty \cdot f(y_k) \leq (1 - \|x_k\|_\infty)f(0) + A\|x_k\|_\infty, \\ f(0) &\leq \left(\frac{\|x_k\|_\infty}{\|x_k\|_\infty + 1}\right) f(z_k) + \left(\frac{1}{\|x_k\|_\infty + 1}\right) f(x_k) \leq \frac{A\|x_k\|_\infty}{\|x_k\|_\infty + 1} + \left(\frac{1}{\|x_k\|_\infty + 1}\right) f(x_k). \end{aligned}$$

Como  $\|x_k\|_\infty$  converge a 0, tomando límite cuando  $k \rightarrow \infty$  vemos que

$$\limsup f(x_k) \leq f(0) \quad \text{y} \quad f(0) \leq \liminf f(x_k),$$

lo cual muestra que  $f(x_k) \rightarrow f(0)$ . Concluimos por tanto que  $f$  es continua en 0.  $\square$

Analicemos ahora las derivadas direccionales de las funciones convexas. Para ello probemos primero un lema.

**Lema 2.25.** *Sea  $f : \mathbb{R} \rightarrow \mathbb{R}$  una función convexa. Si  $a, b, c \in \mathbb{R}$  son tales que  $a < b < c$ , entonces*

$$\frac{f(b) - f(a)}{b - a} \leq \frac{f(c) - f(a)}{c - a} \leq \frac{f(c) - f(b)}{c - b}.$$

PRUEBA. Debido a que  $a < b < c$ , tenemos que  $b = (1 - t)a + tc$ , donde  $t = \frac{b-a}{c-a}$ . Luego, de la convexidad de  $f$  se sigue que  $f(b) \leq (1 - t)f(a) + tf(c)$ ; equivalentemente,  $f(b) \leq \lambda f(a) + (1 - \lambda)f(c)$ , donde  $\lambda = \frac{c-b}{c-a}$ . Operando y reemplazando los valores de  $t$  y  $\lambda$  obtenemos el resultado deseado.  $\square$

**Corolario 2.26.** *Supongamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función convexa,  $\bar{x} \in \mathbb{R}^n$  y  $v \in \mathbb{R}^n$ .*

(1) *La aplicación  $h : (0, +\infty) \rightarrow \mathbb{R}$  definida mediante*

$$t \mapsto h(t) := \frac{f(\bar{x} + tv) - f(\bar{x})}{t}$$

*es no decreciente.*

(2) La derivada direccional unilateral  $f'(\bar{x}; v) = \lim_{t \downarrow 0} h(t)$  existe, es finita y

$$f'(\bar{x}; v) \leq \frac{f(\bar{x} + tv) - f(\bar{x})}{t} \quad \text{para todo } t > 0. \quad (2.7)$$

PRUEBA. Para ver el primer ítem, basta con aplicar el Lema 2.25 a la función convexa  $g : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $g(t) = f(\bar{x} + tv)$ , tomando  $a := 0$ . Vemos en particular que cuando  $t \downarrow 0$ ,  $h(t)$  tiende a un número real o a  $-\infty$ . Aplicando nuevamente el Lema 2.25 a la función  $g$  con  $a = -1$  y  $b = 0$  obtenemos

$$\Delta := f(\bar{x}) - f(\bar{x} - v) \leq \frac{f(\bar{x} + tv) - f(\bar{x})}{t} \quad \text{para todo } t > 0,$$

de donde  $-\infty < \Delta \leq f'(\bar{x}; v)$ , lo que significa que  $f'(\bar{x}; v)$  es finito. La desigualdad (2.7) es consecuencia directa del primer ítem.  $\square$

Si una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es, además de convexa, diferenciable entonces el gradiente  $\nabla f(\bar{x})$  de  $f$  en  $\bar{x}$  nos da una aplicación afín<sup>4</sup>  $L : \mathbb{R}^n \rightarrow \mathbb{R}$  con la propiedad de que  $f \geq L$  y  $f(\bar{x}) = L(\bar{x})$ . Gráficamente, esto significa que en  $\mathbb{R}^n \times \mathbb{R}$  el hiperplano definido por  $L$  separa el punto  $(\bar{x}, f(\bar{x}))$  y el epígrafo  $\text{epi}(f)$  de  $f$ .

**Proposición 2.27.** *Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función convexa y diferenciable. Entonces, para cada  $\bar{x} \in \mathbb{R}^n$  se cumple*

$$f(x) \geq f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x} \rangle, \quad \forall x \in \mathbb{R}^n.$$

Además, esta propiedad caracteriza a  $\nabla f(\bar{x})$ , esto es, si  $v \in \mathbb{R}^n$  es tal que

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle, \quad \forall x \in \mathbb{R}^n, \quad (2.8)$$

entonces  $v = \nabla f(\bar{x})$ .

PRUEBA. Como  $f$  es convexa, para cada  $u \in \mathbb{R}^n$  la aplicación  $t \mapsto \frac{f(\bar{x} + tu) - f(\bar{x})}{t}$  es creciente en  $t > 0$ . De esta manera, si  $t > 0$  y  $u \in \mathbb{R}^n$  tenemos

$$\langle \nabla f(\bar{x}), u \rangle = \lim_{t' \rightarrow 0^+} \frac{f(\bar{x} + t'u) - f(\bar{x})}{t'} \leq \frac{f(\bar{x} + tu) - f(\bar{x})}{t},$$

de donde  $f(\bar{x} + tu) \geq f(\bar{x}) + \langle \nabla f(\bar{x}), tu \rangle$ . Esto prueba la primera parte de la proposición. Para ver la segunda parte, basta con notar que si  $v \in \mathbb{R}^n$  satisface (2.8) entonces

$$\langle \nabla f(\bar{x}), u \rangle = \lim_{t \rightarrow 0} \frac{f(\bar{x} + tu) - f(\bar{x})}{t} \geq \langle v, u \rangle$$

para todo  $u \in \mathbb{R}^n$ , lo cual implica  $v = \nabla f(\bar{x})$ .  $\square$

**Corolario 2.28.** *Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función convexa y diferenciable. Dados dos vectores  $x, \bar{x} \in \mathbb{R}^n$  tenemos que*

<sup>4</sup>Una función  $L : \mathbb{R}^n \rightarrow \mathbb{R}$  es afín si  $L$  y  $-L$  son convexas; equivalentemente,  $L$  es afín si es la traslación de una aplicación lineal  $T : \mathbb{R}^n \rightarrow \mathbb{R}$ , esto es,  $L = T + c$  para algún  $c \in \mathbb{R}$ . Los gráficos de aplicaciones afines de  $\mathbb{R}^n$  en  $\mathbb{R}$  son hiperplanos en  $\mathbb{R}^{n+1}$ .

- (a)  $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0$  implica  $f(x) \geq f(\bar{x})$ ;
- (b)  $f(x) \leq f(\bar{x})$  implica  $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \leq 0$ .

**Corolario 2.29.** Sea  $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$  una función continua tal que la aplicación  $f^y : \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $f^y(x) = f(x, y)$  es convexa y diferenciable para cada  $y \in \mathbb{R}^n$ . Entonces la función  $F : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^m$  dada por  $F(x, y) = \nabla f^y(x)$  es continua.

PRUEBA. Supongamos que  $F$  no es continua en  $(\bar{x}, \bar{y})$ . Entonces existen un abierto  $V \subset \mathbb{R}^m$  que contiene a  $F(\bar{x}, \bar{y}) = \nabla f^{\bar{y}}(\bar{x})$ , y una sucesión  $(x_k, y_k)$  en  $\mathbb{R}^m \times \mathbb{R}^n$  tal que  $(x_k, y_k) \rightarrow (\bar{x}, \bar{y})$  y  $v_k := \nabla f^{y_k}(x_k) \notin V$  para todo  $k$ . Por la Proposición 2.27, para cada  $k$  tenemos

$$f(x, y_k) - f(x_k, y_k) = f^{y_k}(x) - f^{y_k}(x_k) \geq \langle v_k, x - x_k \rangle, \quad \forall x \in \mathbb{R}^m. \quad (2.9)$$

Si  $(v_k)$  no es acotada podemos suponer, tomando subsucesiones y aplicando el teorema de Bolzano-Weierstrass, que  $\|v_k\| \rightarrow \infty$  y que  $\frac{v_k}{\|v_k\|}$  converge a algún vector  $v \in \mathbb{R}^m$  de norma 1. Dividiendo ambos miembros de (2.9) entre  $\|v_k\|$  para cada  $x$  fijo, y haciendo  $k \rightarrow \infty$  obtenemos que el lado izquierdo tiende a 0 (pues  $f$  es continua), mientras que el lado derecho tiende a  $\langle v, x - \bar{x} \rangle$ . Esto significa que  $0 \geq \langle v, x - \bar{x} \rangle$  para todo  $x \in \mathbb{R}^m$ , y por tanto  $0 \geq \langle v, v \rangle = 1$ , una contradicción.

Si  $(v_k)$  es acotada podemos suponer que converge a algún vector  $v \in \mathbb{R}^m$ , que no estará en  $V$  porque  $(v_k)$  está contenida en el conjunto cerrado  $\mathbb{R}^m \setminus V$ . No obstante,  $v = \nabla f^{\bar{y}}(\bar{x}) \in V$  en virtud de la Proposición 2.27, ya que tomando límite cuando  $k \rightarrow \infty$  en (2.9) obtenemos

$$f^{\bar{y}}(x) - f^{\bar{y}}(\bar{x}) = f(x, \bar{y}) - f(\bar{x}, \bar{y}) \geq \langle v, x - \bar{x} \rangle, \quad \forall x \in \mathbb{R}^m.$$

Esta contradicción completa la prueba del corolario. □

Una noción que parametriza la convexidad estricta es la idea de convexidad fuerte. Una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es **fuertemente convexa** con parámetro  $m > 0$  si para todo  $x, y \in \mathbb{R}^n$  y  $t \in [0, 1]$  se cumple

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) - \frac{m}{2}t(1-t)\|x - y\|^2.$$

Toda función fuertemente convexa es estrictamente convexa. Una característica importante de estas funciones es que si poseen mínimo, lo alcanzan solo en un punto. Para funciones estrictamente convexas no se puede garantizar que el mínimo se alcance en algún punto (puede que el mínimo ni siquiera exista como es el caso de  $f(x) = x$ ). No obstante, esto no acontece con las funciones fuertemente convexas: probaremos en el Corolario 2.31 que toda función fuertemente convexa  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  alcanza un único mínimo sobre cada subconjunto convexo y cerrado de  $\mathbb{R}^n$ .

**Proposición 2.30.** Sean las funciones  $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$  tales que  $f(x) = g(x) + \frac{m}{2}\|x - a\|^2$  para todo  $x \in \mathbb{R}^n$ , donde  $m > 0$  y  $a \in \mathbb{R}^n$  es un vector fijo. Entonces  $f$  es fuertemente convexa con parámetro  $m$  si y solo si  $g$  es convexa.

PRUEBA. Observemos primero que dados  $x, y \in \mathbb{R}^n$ , se cumple

$$\begin{aligned} \|x - a\|^2 - \|x - y\|^2 - \|y - a\|^2 &= \|x\|^2 + \|a\|^2 - 2\langle x, a \rangle - (\|x\|^2 + \|y\|^2 + \|y\|^2 \\ &\quad + \|a\|^2 - 2\langle x, y \rangle - 2\langle y, a \rangle) \\ &= -2\langle x, a \rangle + 2\langle x, y \rangle + 2\langle y, a \rangle - 2\|y\|^2 \\ &= 2\langle y - a, x - y \rangle, \end{aligned}$$

de modo que

$$\begin{aligned} \|tx + (1 - t)y - a\|^2 &= \|y - a + t(x - y)\|^2 \\ &= \|y - a\|^2 + t^2\|x - y\|^2 + 2t\langle y - a, x - y \rangle \\ &= t^2\|x - y\|^2 + \|y - a\|^2 + t(\|x - a\|^2 - \|x - y\|^2 - \|y - a\|^2) \\ &= (1 - t)\|y - a\|^2 - t(1 - t)\|x - y\|^2 + t\|x - a\|^2, \end{aligned}$$

para todo  $t \in (0, 1)$ . Luego,  $f$  es fuertemente convexa con parámetro  $m$  si y solo si para todo  $x, y \in \mathbb{R}^n$  y todo  $t \in (0, 1)$  se tiene

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y) - \frac{m}{2}t(1 - t)\|x - y\|^2,$$

lo que equivale a afirmar que

$$\begin{aligned} g(tx + (1 - t)y) + \frac{m}{2}\|tx + (1 - t)y - a\|^2 &\leq tg(x) + (1 - t)g(y) + \frac{m}{2}t\|x - a\|^2 \\ &\quad + \frac{m}{2}(1 - t)\|y - a\|^2 - \frac{m}{2}t(1 - t)\|x - y\|^2, \end{aligned}$$

para todo  $x, y \in \mathbb{R}^n$  y todo  $t \in (0, 1)$ . Pero por la expresión deducida líneas arriba, lo anterior es equivalente a que

$$g(tx + (1 - t)y) \leq tg(x) + (1 - t)g(y), \quad \forall x, y \in \mathbb{R}^n, \forall t \in [0, 1].$$

Así,  $f$  es fuertemente convexa con parámetro  $m$  si y solo si  $g$  es convexa.  $\square$

Antes de mostrar el siguiente corolario recordemos que toda función convexa  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es continua.

**Corolario 2.31.** *Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  fuertemente convexa. Entonces  $f$  alcanza mínimo sobre cualquier subconjunto convexo, cerrado y no vacío de  $\mathbb{R}^n$ . Además, el punto donde se alcanza tal mínimo es único.*

PRUEBA. Sea  $m > 0$  el parámetro de convexidad fuerte de  $f$ . Primero vamos a probar que  $f$  está acotada inferiormente. Sin pérdida de generalidad supongamos que  $f(0) = 0$ . La función  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  dada por  $g(x) = f(x) - \frac{m}{2}\|x\|^2$  satisface  $g(0) = 0$  y, por la Proposición 2.30, es convexa y continua. Luego,  $g$  alcanza un mínimo  $M \in \mathbb{R}$  en el conjunto compacto  $K := \{x \in \mathbb{R}^n \mid \|x\| \leq 1\}$ . Dado  $z \in \mathbb{R}^n$  con  $\|z\| \geq 1$  tenemos  $\frac{z}{\|z\|} \in K$ , y por la convexidad de  $g$ ,

$$M \leq g\left(\frac{z}{\|z\|}\right) = g\left(\frac{1}{\|z\|}z + \frac{\|z\|-1}{\|z\|} \cdot 0\right) \leq \frac{1}{\|z\|}g(z) + \frac{\|z\|-1}{\|z\|}g(0) = \frac{1}{\|z\|}g(z),$$

esto es,  $g(z) \geq M\|z\|$ . Luego,

$$f(z) = g(z) + \frac{m}{2}\|z\|^2 \geq M\|z\| + \frac{m}{2}\|z\|^2 \geq \frac{-M^2}{2m}, \quad \text{si } \|z\| \geq 1,$$

ya que la aplicación cuadrática  $t \mapsto Mt + \frac{m}{2}t^2$  está acotada inferiormente por  $\frac{-M^2}{2m}$ . Como  $f$  está acotada en el conjunto compacto  $K$  (por continuidad), lo anterior muestra que  $f$  está acotada inferiormente en  $\mathbb{R}^n$ .

Mostraremos ahora que los subniveles de  $f$  son compactos, para lo cual bastará con probar que son acotados ( $f$  es continua). Sea  $\alpha := \inf_{x \in \mathbb{R}^n} f(x) \in \mathbb{R}$ . Dado  $\lambda \in \mathbb{R}$  consideramos dos casos: si  $S_\lambda$  es vacío no hay nada que probar; y si  $S_\lambda$  es no vacío podemos escoger  $\bar{x} \in S_\lambda$  de modo que para todo  $x \in S_\lambda$  tenemos por la convexidad fuerte de  $f$  que

$$\alpha \leq f\left(\frac{x+\bar{x}}{2}\right) \leq \frac{1}{2}f(x) + \frac{1}{2}f(\bar{x}) - \frac{m}{2} \cdot \frac{1}{4}\|x - \bar{x}\|^2 \leq \lambda - \frac{m}{8}\|x - \bar{x}\|^2,$$

es decir,  $\|x - \bar{x}\|^2 \leq \frac{8}{m}(\lambda - \alpha)$ . Esto muestra que los subniveles de  $f$  son acotados (y por tanto compactos).

Sea ahora  $C$  un subconjunto convexo, cerrado y no vacío de  $\mathbb{R}^n$ . El ínfimo  $\theta := \inf_{x \in C} f(x)$  es un número real pues  $f$  está acotada inferiormente. Queremos ver que existe un único  $\bar{x} \in C$  tal que  $f(\bar{x}) = \theta$ . Sea la sucesión  $\lambda_k := \theta + \frac{1}{k}$ . Entonces cada conjunto  $C_k := C \cap S_{\lambda_k}$  es no vacío y compacto. Además,  $C_1 \supset C_2 \supset \dots$  de manera que aplicando el teorema de Cantor deducimos que el conjunto  $\bigcap_{k \geq 1} C_k$  es compacto y no vacío. Ahora bien, dado  $x \in C$  se cumple que  $x \in \bigcap_{k \geq 1} C_k$  si y solo si  $\theta \leq f(x) \leq \theta + \frac{1}{k}$  para todo  $k$ , esto es, si y solo si  $f(x) = \theta$ . En particular, existe  $\bar{x} \in C$  donde  $f$  alcanza el ínfimo  $\theta$ ; este  $\bar{x}$  es único porque  $f$  es estrictamente convexa.  $\square$

### 2.3. Subdiferenciales

En la Sección 3 veremos que si  $X \subset \mathbb{R}^n$  es convexo y la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es convexa y diferenciable entonces los mínimos globales de  $f$  sobre  $X$  son precisamente los vectores  $\bar{x} \in X$  tales que  $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0$  para todo  $x \in X$ . Esto ilustra la relevancia que tiene el gradiente para minimizar funciones convexas y diferenciables. En esta sección estudiamos el análogo del gradiente para funciones convexas no diferenciables: el subdiferencial.

Dada una función convexa  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , decimos que el vector  $v \in \mathbb{R}^n$  es un **subgradiente** de  $f$  en el punto  $\bar{x} \in \mathbb{R}^n$  si

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle \quad \text{para todo } x \in \mathbb{R}^n. \quad (2.10)$$

El **subdiferencial** de  $f$  en  $\bar{x}$ , denotado por  $\partial f(\bar{x})$ , es el conjunto de todos los subgradientes de  $f$  en  $\bar{x}$ , esto es,

$$\partial f(\bar{x}) = \{v \in \mathbb{R}^n \mid f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle \text{ para todo } x \in \mathbb{R}^n\}.$$

Intuitivamente, el subdiferencial es el conjunto de los vectores  $v \in \mathbb{R}^n$  que generan hiperplanos  $H_v = \{(x, f(\bar{x}) + \langle v, x - \bar{x} \rangle) \mid x \in \mathbb{R}^n\}$  en  $\mathbb{R}^n \times \mathbb{R}$  que se encuentran

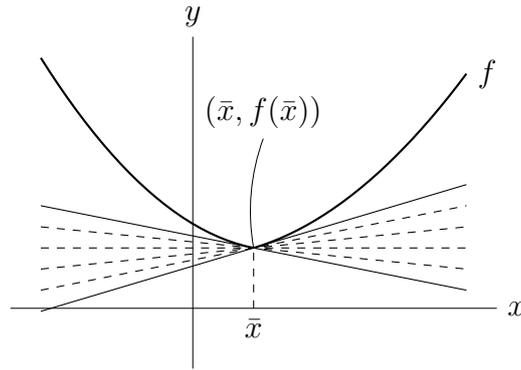


Figura 2.5: Ilustración gráfica del subdiferencial. Para la función  $f$ ,  $\partial f(\bar{x})$  es el conjunto de pendientes de las rectas dibujadas que pasan por  $(\bar{x}, f(\bar{x}))$  y que se encuentran por debajo del gráfico de  $f$ .

debajo del gráfico de  $f$ . Cuando  $n = 1$ ,  $\partial f(\bar{x})$  es el conjunto de todos los números  $m \in \mathbb{R}$  tal que existe una recta tangente de pendiente  $m$  que pasa por  $(\bar{x}, f(\bar{x}))$  y que está por debajo del gráfico de  $f$ .

**Ejemplo 2.32.** El subdiferencial de la función norma  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = \|x\|$  está dado por

$$\partial f(x) = \begin{cases} \{y \in \mathbb{R}^n \mid \|y\| \leq 1\}, & \text{si } x = 0, \\ \left\{ \frac{x}{\|x\|} \right\}, & \text{si } x \neq 0. \end{cases}$$

En efecto, cuando  $\bar{x} = 0$  tenemos que  $y \in \partial f(\bar{x})$  si y solo si  $\|x\| \geq \langle y, x \rangle$  para todo  $x \in \mathbb{R}^n$ , esto es, si y solo si  $\|y\| \leq 1$ . Cuando  $\bar{x} \neq 0$ , es claro que  $\frac{\bar{x}}{\|\bar{x}\|} \in \partial f(\bar{x})$  pues

$$\left\langle \frac{\bar{x}}{\|\bar{x}\|}, x - \bar{x} \right\rangle = \left\langle \frac{\bar{x}}{\|\bar{x}\|}, x \right\rangle - \left\langle \frac{\bar{x}}{\|\bar{x}\|}, \bar{x} \right\rangle \leq \|x\| - \|\bar{x}\| \text{ para todo } x \in \mathbb{R}^n.$$

Y si  $y \in \mathbb{R}^n$  es tal que  $\langle y, x - \bar{x} \rangle \leq \|x\| - \|\bar{x}\|$  para todo  $x \in \mathbb{R}^n$ , entonces  $y = \frac{x}{\|x\|}$ . De hecho, tomando  $x = 0$  y  $x = 2\bar{x}$  obtenemos  $\|\bar{x}\| = \langle y, \bar{x} \rangle$ , de donde  $\langle y, x \rangle \leq \|x\|$  para todo  $x \in \mathbb{R}^n$ , es decir  $\|y\| \leq 1$ , y

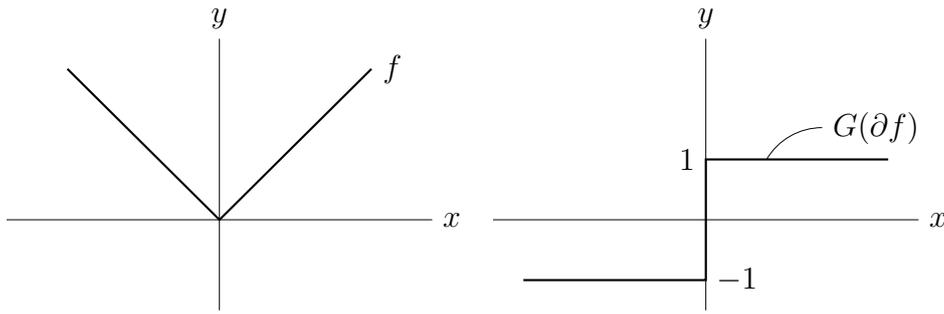
$$\left\| y - \frac{\bar{x}}{\|\bar{x}\|} \right\|^2 = \|y\|^2 + 1 - 2\langle y, \frac{\bar{x}}{\|\bar{x}\|} \rangle \leq 2 - 2 \cdot \frac{\langle y, \bar{x} \rangle}{\|\bar{x}\|} = 0.$$

**Ejemplo 2.33** (El subdiferencial es la generalización del gradiente). Sea  $f : \mathbb{R} \rightarrow \mathbb{R}$  una función convexa y diferenciable. Dado  $\bar{x} \in \mathbb{R}^n$ , la Proposición 2.27 nos dice que  $\nabla f(\bar{x})$  es el único subgradiente de  $f$  en  $\bar{x}$ , esto es,  $\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$ .

**Proposición 2.34.** Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función convexa, y sea  $\bar{x} \in \mathbb{R}^n$ . Entonces el subdiferencial  $\partial f(\bar{x})$  es convexo, compacto y no vacío.

PRUEBA. El subdiferencial  $\partial f(\bar{x})$  es siempre convexo: si  $v_1, v_2 \in \partial f(\bar{x})$  y  $t \in [0, 1]$  entonces  $y_t := tv_1 + (1-t)v_2$  está en  $\partial f(\bar{x})$  pues para cada  $x \in \mathbb{R}^n$  se cumple que

$$\begin{aligned} f(x) - f(\bar{x}) &= t(f(x) - f(\bar{x})) + (1-t)(f(x) - f(\bar{x})) \\ &\geq t\langle v_1, x - \bar{x} \rangle + (1-t)\langle v_2, x - \bar{x} \rangle = \langle y_t, x - \bar{x} \rangle. \end{aligned}$$


 Figura 2.6: Ilustración gráfica del Ejemplo 2.32 cuando  $n = 1$ .

Veamos que  $\partial f(\bar{x})$  es compacto, para lo cual mostraremos que  $\partial f(\bar{x})$  es cerrado y acotado. Es claro que  $\partial f(\bar{x})$  es cerrado, pues  $\partial f(\bar{x}) = \bigcap_{x \in \mathbb{R}^n} C_x$  y cada conjunto  $C_x := \{v \in \mathbb{R}^n \mid f(x) - f(\bar{x}) \geq \langle v, x - \bar{x} \rangle\}$  es cerrado. Supongamos que  $\partial f(\bar{x})$  no es acotado, esto es, existe una sucesión  $(v_k)$  en  $\partial f(\bar{x})$  tal que  $\|v_k\| \rightarrow \infty$ . Podemos suponer que  $\frac{v_k}{\|v_k\|}$  converge a algún vector  $v \in \mathbb{R}^n$  de norma 1. Luego

$$\frac{f(x) - f(\bar{x})}{\|v_k\|} \geq \left\langle \frac{v_k}{\|v_k\|}, x - \bar{x} \right\rangle \quad \text{para todo } x \in \mathbb{R}^n \text{ y todo } k,$$

y haciendo  $k \rightarrow \infty$  obtenemos  $0 \geq \langle v, x - \bar{x} \rangle$  para todo  $x \in \mathbb{R}^n$ , lo cual implica la contradicción  $0 \geq \langle v, v \rangle = 1$ . Así,  $\partial f(\bar{x})$  es compacto.

Veamos por último que  $\partial f(\bar{x})$  es no vacío. Como  $f$  es convexa y continua (Proposición 2.24),  $\text{epi}(f)$  es un conjunto convexo y cerrado en cuya frontera se encuentra  $z_0 := (\bar{x}, f(\bar{x}))$ , siendo esto así porque  $\lambda > f(\bar{x})$  implica  $(\bar{x}, \lambda) \in \text{epi}(f)$  y porque  $\lambda < f(\bar{x})$  implica  $(\bar{x}, \lambda) \notin \text{epi}(f)$ . Por el teorema de separación (Teorema 2.12), existe  $v' := (v, t) \in \mathbb{R}^n \times \mathbb{R}$  no nulo tal que  $\langle v', z \rangle \geq \langle v', z_0 \rangle$  para todo  $z \in \text{epi}(f)$ , lo cual significa que

$$\langle v, x \rangle + t\lambda \geq \langle v, \bar{x} \rangle + tf(\bar{x}) = \langle v', z_0 \rangle \quad \text{para todo } x \in \mathbb{R}^n \text{ y todo } \lambda \geq f(x). \quad (2.11)$$

En esta ecuación no es posible que  $t \leq 0$ , pues si  $t = 0$  entonces  $\langle v, x - \bar{x} \rangle \geq 0$  para todo  $x \in \mathbb{R}^n$ , lo que implica  $v = 0$ , absurdo ya que  $v'$  es no nulo; y si  $t < 0$ , haciendo  $\lambda \rightarrow \infty$  en (2.11) obtenemos  $-\infty \geq \langle v', z_0 \rangle$ , absurdo también. Tenemos por tanto que  $t > 0$ , y de la ecuación (2.11) deducimos

$$f(x) - f(\bar{x}) \geq \left\langle -\frac{v}{t}, x - \bar{x} \right\rangle \quad \text{para todo } x \in \mathbb{R}^n,$$

de donde  $-\frac{v}{t} \in \partial f(\bar{x})$ . Esto prueba que  $\partial f(\bar{x})$  es no vacío.  $\square$

Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función convexa, y sea  $\bar{x} \in \mathbb{R}^n$ . Dado  $v \in \mathbb{R}^n$  tenemos que  $v \in \partial f(\bar{x})$  si y solo si  $f(\bar{x} + tu) - f(\bar{x}) \geq \langle v, tu \rangle$  para todo  $t > 0$  y todo  $u \in \mathbb{R}^n$ , esto es, si y solo si,

$$\frac{f(\bar{x} + tu) - f(\bar{x})}{t} \geq \langle v, u \rangle, \quad \forall t > 0, \forall u \in \mathbb{R}^n.$$

Pero  $f'(\bar{x}; u) = \lim_{t \downarrow 0} \left( \frac{f(\bar{x}+tu) - f(\bar{x})}{t} \right)$ , y como la aplicación  $t \mapsto \frac{f(\bar{x}+tu) - f(\bar{x})}{t}$  es no decreciente en  $(0, +\infty)$  para cada  $u$  fijo, lo anterior significa

$$v \in \partial f(\bar{x}) \Leftrightarrow f'(\bar{x}; u) \geq \langle v, u \rangle, \quad \forall u \in \mathbb{R}^n.$$

Así, para cada  $u \in \mathbb{R}^n$  obtenemos

$$f'(\bar{x}; u) \geq \max_{v \in \partial f(\bar{x})} \langle v, u \rangle,$$

donde el símbolo  $\max$  está bien empleado porque  $\partial f(\bar{x})$  es compacto. Ahora bien, resulta que esta desigualdad es en realidad una igualdad.

**Proposición 2.35.** *Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función convexa, y sea  $\bar{x} \in \mathbb{R}^n$ . Entonces, para cada  $u \in \mathbb{R}^n$  se verifica*

$$f'(\bar{x}; u) = \max_{v \in \partial f(\bar{x})} \langle v, u \rangle.$$

PRUEBA. Basta con encontrar un elemento  $v$  de  $\partial f(\bar{x})$  tal que  $\langle v, u \rangle \geq f'(\bar{x}; u)$ . Consideremos los subconjuntos de  $\mathbb{R}^{n+1}$

$$\begin{aligned} W &= \{(x, \lambda) \mid f(x) < \lambda\}, \\ Z &= \{(\bar{x}, f(\bar{x})) + \alpha(u, f'(\bar{x}; u)) \mid \alpha \geq 0\}. \end{aligned}$$

Como  $W$  y  $Z$  son disjuntos (Corolario 2.26), convexos y no vacíos, por el teorema de separación (Teorema 2.12) existe un vector  $v' := (v, t) \in \mathbb{R}^n \times \mathbb{R}$  no nulo tal que  $\langle v', w \rangle \leq \langle v', z \rangle$  para cualesquiera  $w \in W$  y  $z \in Z$ , esto es,

$$\langle v, x \rangle + t\lambda \leq \langle v, \bar{x} + \alpha u \rangle + t(f(\bar{x}) + \alpha f'(\bar{x}; u)), \quad \forall \alpha \geq 0, x \in \mathbb{R}^n, \lambda > f(x). \quad (2.12)$$

En esta ecuación no es posible que  $t \geq 0$  pues si  $t = 0$  entonces  $\langle v, x - \bar{x} - \alpha u \rangle \leq 0$  para todo  $x \in \mathbb{R}^n$  y  $\alpha \geq 0$ , lo que implica  $v = 0$ , absurdo ya que  $v'$  es no nulo; y si  $t > 0$  podemos hacer  $\lambda \rightarrow \infty$  de modo que  $+\infty \leq \langle v', z \rangle$ , igualmente absurdo. Luego  $t < 0$ , y, dividiendo entre  $-t$  en (2.12) si fuera necesario, podemos suponer que  $t = -1$ :

$$\lambda \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle + \alpha f'(\bar{x}; u) - \alpha \langle v, u \rangle, \quad \forall \alpha \geq 0, x \in \mathbb{R}^n, \lambda > f(x). \quad (2.13)$$

Tomando  $\alpha = 0$  en (2.13) y haciendo tender  $\lambda$  a  $f(x)$  para cada  $x$  fijo obtenemos

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle, \quad \forall x \in \mathbb{R}^n,$$

es decir,  $v \in \partial f(\bar{x})$ . Finalmente, tomando  $x = \bar{x}$ ,  $\alpha = 1$  y haciendo tender  $\lambda$  a  $f(\bar{x})$  en (2.13) conseguimos  $\langle v, u \rangle \geq f'(\bar{x}; u)$ .  $\square$

Si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función convexa entonces cada  $\partial f(\bar{x})$  es un subconjunto de  $\mathbb{R}^n$ . Esto significa que el subdiferencial de  $f$  define de modo natural una multifunción<sup>5</sup>

$$\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$$

que aplica cada  $x \in \mathbb{R}^n$  en  $\partial f(x) \subset \mathbb{R}^n$ .

La Proposición 2.34 puede reescribirse del siguiente modo.

<sup>5</sup>Una multifunción u operador multivaluado  $\Omega : X \rightrightarrows Y$  asocia a cada elemento  $x$  de  $X$  un subconjunto  $\Omega(x)$  de  $Y$ . En otras palabras, se trata de una función  $\Omega : X \rightarrow 2^Y$ , donde  $2^Y$  es el conjunto potencia de  $Y$ . Para mayor información véase la Sección 4.

**Proposición 2.36.** *Si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función convexa entonces la multifunción  $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  tiene imágenes convexas, compactas y no vacías.*

El operador  $\partial f$  tiene buenas propiedades. Por ejemplo,  $\partial f$  es scs<sup>6</sup> (Corolario 2.38). Más generalmente, tenemos el siguiente resultado, que es una generalización del Corolario 2.29.

**Proposición 2.37.** *Sea  $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$  una función continua tal que la aplicación  $f^y : \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $f^y(x) = f(x, y)$  es convexa para cada  $y \in \mathbb{R}^n$ . Entonces, el operador multivaluado  $\Omega : \mathbb{R}^m \times \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  dado por  $\Omega(x, y) = \partial f^y(x)$  es scs.*

PRUEBA. Supongamos que  $\Omega$  no es scs en  $(x, y)$ . Entonces, por el Lema 4.10 existen  $V \subset \mathbb{R}^m$  abierto,  $(x_k, y_k)$  en  $\mathbb{R}^m \times \mathbb{R}^n$  y  $(v_k)$  en  $\mathbb{R}^m$  tales que  $\Omega(x, y) = \partial f^y(x) \subset V$ ,  $(x_k, y_k) \rightarrow (x, y)$  y

$$v_k \in \partial f^{y_k}(x_k), \quad v_k \notin V, \quad \forall k \geq 1.$$

Por la definición de subdiferencial, para cada  $k$  tenemos

$$f(x', y_k) - f(x_k, y_k) = f^{y_k}(x') - f^{y_k}(x_k) \geq \langle v_k, x' - x_k \rangle, \quad \forall x' \in \mathbb{R}^m. \quad (2.14)$$

Si  $(v_k)$  no es acotada podemos suponer, tomando subsucesiones y aplicando el teorema de Bolzano-Weierstrass, que  $\|v_k\| \rightarrow \infty$  y que  $\frac{v_k}{\|v_k\|}$  converge a algún vector  $v \in \mathbb{R}^m$  de norma 1. Dividiendo ambos miembros de (2.14) entre  $\|v_k\|$  para cada  $x'$  fijo, y haciendo  $k \rightarrow \infty$  obtenemos que el lado izquierdo tiende a 0 (pues  $f$  es continua), mientras que el lado derecho tiende a  $\langle v, x' - x \rangle$ . Esto significa que  $0 \geq \langle v, x' - x \rangle$  para todo  $x' \in \mathbb{R}^m$ , y por tanto  $v = 0$ , una contradicción.

Si  $(v_k)$  es acotada podemos suponer que converge a algún  $v \in \mathbb{R}^m$ , que no estará en  $V$  porque  $(v_k)$  está contenida en el conjunto cerrado  $\mathbb{R}^m \setminus V$ . No obstante,  $v$  está en  $\partial f^y(x) \subset V$  ya que tomando límite cuando  $k \rightarrow \infty$  en (2.14) obtenemos

$$f^y(x') - f^y(x) = f(x', y) - f(x, y) \geq \langle v, x' - x \rangle, \quad \forall x' \in \mathbb{R}^m.$$

Esta contradicción completa la prueba de la proposición. □

**Corolario 2.38.** *Si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función convexa entonces la multifunción  $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  es scs.*

PRUEBA. Como la función  $g : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$  dada por  $g(x, t) = f(x)$  es continua ( $f$  es continua por ser convexa) y tiene la propiedad de que cada  $g^t = f$  es convexa, la Proposición 2.37 nos dice que la multifunción  $(x, t) \mapsto \partial f(x)$  es scs. Luego, dados  $\bar{x} \in \mathbb{R}^n$  y  $V \subset \mathbb{R}^n$  abierto con  $\partial f(\bar{x}) \subset V$ , puesto que dicha multifunción es scs en  $(\bar{x}, 0)$ , existe  $\delta > 0$  tal que  $\|x - \bar{x}\| = \|(x, 0) - (\bar{x}, 0)\| < \delta$  implica  $\partial f(x) \subset V$ . Esto prueba que  $\partial f$  es scs. □

---

<sup>6</sup>Un operador  $\Omega : X \rightrightarrows Y$  es **semicontinuo superiormente (scs)** en  $\bar{x} \in X$  si para cada abierto  $V$  de  $Y$  tal que  $\Omega(\bar{x}) \subset V$  existe una vecindad abierta  $U$  de  $\bar{x}$  tal que  $\Omega(x) \subset V$  para todo  $x \in U$ ;  $\Omega$  es **scs** si es scs en cada  $x \in X$ . Véanse también la Definición 4.2 y el Lema 4.10.

## 2.4. Funciones pseudoconvexas

En el Corolario 2.28 mostramos que si  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función convexa y diferenciable y  $\bar{x}, x \in \mathbb{R}^n$ , entonces  $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0$  implica que  $f(x) \geq f(\bar{x})$ . Esta propiedad de las funciones convexas y diferenciables motiva la siguiente definición.

Una función diferenciable  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es **pseudoconvexa** si para todo par de puntos  $x, \bar{x} \in \mathbb{R}^n$  tenemos que

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0 \text{ implica } f(x) \geq f(\bar{x}).$$

Toda función diferenciable que es convexa también es pseudoconvexa. La función  $\mathbb{R} \rightarrow \mathbb{R}$  dada por  $x \mapsto x^3 + x$  es pseudoconvexa pero no convexa; más generalmente, si  $f : \mathbb{R} \rightarrow \mathbb{R}$  es una función creciente, diferenciable y tal que  $f'(x) \neq 0$  para todo  $x \in \mathbb{R}$ , entonces  $f$  es pseudoconvexa.

## 3. Problemas de optimización

La Optimización Matemática estudia las propiedades de máximos y mínimos de funciones. Preguntas típicas que la Optimización intenta responder son las siguientes: dada una función  $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ , ¿cuál es el ínfimo  $\inf_{x \in X} f(x)$  de  $f$ ?; si dicho ínfimo es un número real, ¿existe algún punto de  $X$  donde es efectivamente alcanzado?, ¿qué condiciones necesarias y/o suficientes deben satisfacer  $f$  y  $X$  para poder estar seguros de que existe?; y si existe tal punto, ¿es único?, ¿hay algún algoritmo para encontrarlo? Serán justamente estas cuestiones las que trataremos de resolver en el próximo capítulo, cuando estudiemos los equilibrios de Nash. Por ello hacemos aquí una revisión breve de las notaciones y terminología básicas en Optimización, basándonos principalmente en las referencias [8] y [6].

Sean la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  y el subconjunto  $X$  de  $\mathbb{R}^n$ .

Un vector  $\bar{x} \in \mathbb{R}^n$  es un **mínimo global** de  $f$  **sobre**  $X$  si  $\bar{x} \in X$  y si para todo  $x \in X$  tenemos  $f(x) \geq f(\bar{x})$ ; equivalentemente,  $f(\bar{x}) = \inf_{x \in X} f(x)$  y  $\bar{x} \in X$ . Decimos entonces que  $\bar{x}$  es una **solución óptima** (o simplemente **solución**) del **problema de optimización (minimización)**

$$\min f(x), \quad x \in X. \tag{3.1}$$

También decimos que  $f$  **alcanza su mínimo sobre  $X$  en  $\bar{x}$** , o que  $\bar{x}$  **minimiza  $f$  sobre  $X$** . Si  $X = \mathbb{R}^n$  entonces el problema (3.1) es llamado problema de optimización (minimización) **sin restricciones** o problema de optimización (minimización) **global**, y escribimos

$$\bar{x} \in \arg \min f$$

para indicar que  $\bar{x}$  es una solución óptima. En caso contrario, (3.1) es un problema de optimización **con restricciones**. Un vector  $x \in \mathbb{R}^n$  es llamado **factible** si satisface la **restricción** del problema, es decir, si  $x \in X$ ; la **región factible** del problema es el conjunto de todos los puntos factibles. La función  $f$  que se quiere minimizar es llamada **función objetivo**, **función de beneficio** o **función de coste**.

Un punto  $\bar{x} \in \mathbb{R}^n$  es un **mínimo local** de  $f$  sobre  $X$  si  $\bar{x}$  minimiza  $f$  sobre una vecindad  $U$  de  $\bar{x}$  en  $X$ . En otras palabras,  $\bar{x} \in \mathbb{R}^n$  es un mínimo local de  $f$  si es solución del problema

$$\min f(x), \quad x \in B_X(\bar{x}, \delta)$$

para algún  $\delta > 0$ . Todo mínimo global es un mínimo local.

Los **máximos globales** y **locales** son definidos de manera similar. Así, un vector dado  $\bar{x} \in \mathbb{R}^n$  es un máximo global de  $f$  sobre  $X$  si es solución del problema de optimización (o **maximización**)

$$\max f(x), \quad x \in X;$$

y  $\bar{x}$  es un máximo local de  $f$  sobre  $X$  si  $\bar{x}$  maximiza  $f$  sobre una vecindad de  $\bar{x}$  en  $X$ .

Así, un problema de optimización consiste en minimizar (maximizar) una función dada sobre una cierta región de su dominio. Por ejemplo, el problema

$$\min e^{-\|x\|}, \quad x \in \mathbb{R}^n$$

no posee solución, mientras que el problema

$$\min \|x\|, \quad x \in B(0, 1)$$

tiene una única solución en  $x = 0$ .

La existencia de solución de un problema de optimización a veces depende de las propiedades topológicas de  $f$  y  $X$ . Por ejemplo, si  $f$  es continua y  $X$  es compacto el problema (3.1) siempre posee solución por el teorema de Weierstrass (Teorema 1.14).

Una **condición de optimalidad** para el problema (3.1) es una condición necesaria y/o suficiente para que un vector factible  $x \in X$  sea un mínimo global o local.

La mayoría de algoritmos y condiciones de optimalidad con que se trabaja no distingue entre los mínimos locales y globales. No obstante, si  $f$  es convexa y  $X$  es convexo, entonces todo mínimo local (sobre  $X$ ) es un mínimo global (sobre  $X$ )<sup>7</sup>, y en consecuencia los mínimos locales y globales de  $f$  (sobre  $X$ ) coinciden.

Una condición de optimalidad básica que se enseña en los cursos de cálculo nos dice que todo mínimo local  $\bar{x}$  de una función diferenciable  $f$  sobre  $\mathbb{R}^n$  tiene la propiedad de que  $\nabla f(\bar{x}) = 0$ . La siguiente proposición es una generalización de este resultado.

**Proposición 3.1** (Condición de optimalidad necesaria de primer orden). *Si  $f$  es una función diferenciable y  $\bar{x}$  es un mínimo local de  $f$  sobre el conjunto convexo  $X \subset \mathbb{R}^n$ , entonces  $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0$  para todo  $x \in X$ .*

PRUEBA. Dado  $x \in X$ , tenemos que  $\bar{x} + t(x - \bar{x}) \in X$  si  $t \in [0, 1]$  pues  $X$  es convexo. Luego  $f(\bar{x} + t(x - \bar{x})) \geq f(\bar{x})$  para  $t > 0$  suficientemente pequeño y por tanto

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle = f'(\bar{x}; x - \bar{x}) = \lim_{t \downarrow 0} \frac{f(\bar{x} + t(x - \bar{x})) - f(\bar{x})}{t} \geq 0. \quad \square$$

<sup>7</sup>Más generalmente, si  $f$  es cuasiconvexa semiestricta entonces todo mínimo local es un mínimo global.

**Corolario 3.2.** *Supongamos que  $\bar{x}$  es un mínimo local de la función (no necesariamente diferenciable)  $f$  sobre el conjunto convexo  $X \subset \mathbb{R}^n$ . Si  $x \in X$  es tal que la derivada direccional unilateral  $f'(\bar{x}; x - \bar{x})$  existe, entonces  $f'(\bar{x}; x - \bar{x}) \geq 0$ .*

Supongamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es diferenciable y que  $X \subset \mathbb{R}^n$  es convexo. Los vectores factibles  $\bar{x} \in X$  tales que  $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0$  para todo  $x \in X$  son llamados **puntos estacionarios** del problema de optimización

$$\min f(x), \quad x \in X. \quad (3.2)$$

Evidentemente, si  $X = \mathbb{R}^n$ , los puntos estacionarios son los puntos  $\bar{x}$  tales que  $\nabla f(\bar{x}) = 0$ . En términos de desigualdades variacionales (véase la Sección 7), los puntos estacionarios son las soluciones del problema  $VI(\nabla f, X)$ .

Así, la Proposición 3.1 nos asegura que toda solución del problema de optimización (3.2) es un punto estacionario. Como la mayoría de algoritmos usados en Optimización sirven para hallar puntos estacionarios, es interesante saber bajo qué condiciones un punto estacionario es solución del problema de optimización.

**Proposición 3.3** (Condición de optimalidad suficiente de primer orden). *Supongamos que  $f$  es convexa y diferenciable (o, más generalmente, que  $f$  es pseudoconvexa) y que  $X \subset \mathbb{R}^n$  es convexo. Si  $\bar{x} \in X$  es tal que  $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0$  para todo  $x \in X$ , entonces  $\bar{x}$  minimiza  $f$  sobre  $X$ .*

PRUEBA. Como  $f$  es convexa, aplicando el Corolario 2.28 vemos que para cada  $x \in X$  se cumple que  $f(x) \geq f(\bar{x})$ . Si  $f$  es pseudoconvexa, en vez del Corolario 2.28 usamos la definición de función pseudoconvexa.  $\square$

**Corolario 3.4.** *Si  $f$  es convexa y diferenciable (o pseudoconvexa), entonces  $\bar{x} \in \mathbb{R}^n$  es un mínimo global de  $f$  sobre  $\mathbb{R}^n$  si y solo si  $\nabla f(\bar{x}) = 0$ .*

Los dos resultados anteriores ya ilustran las fuertes implicancias que tiene la convexidad. En particular, el Corolario 3.4 traduce el problema (3.1), donde  $X = \mathbb{R}^n$ , en el problema de encontrar los ceros de la función  $\nabla f : \mathbb{R}^n \rightarrow \mathbb{R}$ .

## Condiciones KKT

En lo que resta de esta sección intentamos establecer qué condiciones necesarias y/o suficientes (de primer orden) deben satisfacer los minimizadores locales del problema

$$\min f(x), \quad x \in X \quad (3.1)$$

cuando la región factible  $X$  tiene la forma

$$X = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \forall i = 1, \dots, r\}, \quad (3.3)$$

donde  $g_1, \dots, g_r : \mathbb{R}^n \rightarrow \mathbb{R}$  son funciones diferenciables. Veremos que estas condiciones están relacionadas con los conocidos “multiplicadores de Lagrange” que se estudian en los cursos de cálculo.

Sea  $g := (g_1, \dots, g_r) : \mathbb{R}^n \rightarrow \mathbb{R}^r$  y supongamos que la función  $f$  es diferenciable. Entonces  $g$  es diferenciable y el problema (3.1) puede escribirse de la siguiente forma

$$\begin{aligned} \min f(x) \\ \text{s.a. } g(x) \leq 0, \end{aligned}$$

donde “s.a.” significa “sujeto a”.

Dado un punto factible  $\bar{x} \in X$  definimos el **conjunto activo** de  $\bar{x}$  mediante

$$I(\bar{x}) = \{i \mid g_i(\bar{x}) = 0\}.$$

Un vector  $\lambda = (\lambda_1, \dots, \lambda_r)$  es llamado **vector de multiplicadores de Lagrange** en  $\bar{x}$  si las siguientes condiciones se satisfacen

- (1)  $\nabla f(\bar{x}) + \sum_i \lambda_i \nabla g_i(\bar{x}) = 0$ ;
- (2)  $\lambda_i \geq 0$  para todo  $i$ ;
- (3)  $\lambda_i = 0$  para todo  $i \notin I(\bar{x})$ .

Abreviadamente, las condiciones (2) y (3) pueden escribirse así

$$0 \leq \lambda \perp g(\bar{x}) \leq 0.$$

Una **condición de calificación (CQ)** para el problema de optimización (3.1) es una condición suficiente para que en un mínimo local dado exista un vector de multiplicadores de Lagrange.

**Lema 3.5.** *Supongamos que  $h_1, \dots, h_p : \mathbb{R}^n \rightarrow \mathbb{R}$  son funciones diferenciables y definamos  $h(x) := \max\{h_1(x), \dots, h_p(x)\}$ . Entonces las derivadas direccionales unilaterales de  $h$  existen y están dadas por*

$$h'(x; v) = \max\{\langle \nabla h_i(x), v \rangle \mid i \in A(x)\},$$

donde  $A(x) := \{i \mid h(x) = h_i(x)\}$ .

PRUEBA. Sea  $\bar{x} \in \mathbb{R}^n$  y consideremos el vector  $v \in \mathbb{R}^n$ . Por continuidad, podemos suponer que  $A(\bar{x}) = \{1, \dots, p\}$ . Luego para cada  $i$  tenemos

$$\liminf_{t \downarrow 0} \frac{h(\bar{x} + tv) - h(\bar{x})}{t} \geq \liminf_{t \downarrow 0} \frac{h_i(\bar{x} + tv) - h_i(\bar{x})}{t} = \langle \nabla h_i(\bar{x}), v \rangle,$$

de donde

$$\liminf_{t \downarrow 0} \frac{h(\bar{x} + tv) - h(\bar{x})}{t} \geq \max\{\langle \nabla h_i(\bar{x}), v \rangle \mid i \in A(\bar{x})\}.$$

Por otro lado, existe una sucesión  $(t_k)$  en  $(0, \infty)$  tal que  $t_k \rightarrow 0$  y

$$\lim_{k \rightarrow \infty} \frac{h(\bar{x} + t_k v) - h(\bar{x})}{t_k} = \limsup_{t \downarrow 0} \frac{h(\bar{x} + tv) - h(\bar{x})}{t}.$$

Como  $A(\bar{x})$  es finito, existe algún  $j \in A(\bar{x})$  con la propiedad de que  $h_j(x + t_k v) = h(x + t_k v)$  para infinitos  $k$ , de modo que podemos tomar límite en

$$\frac{h(\bar{x} + t_k v) - h(\bar{x})}{t_k} \leq \frac{h_j(\bar{x} + t_k v) - h_j(\bar{x})}{t_k}$$

para obtener

$$\limsup_{t \downarrow 0} \frac{h(\bar{x} + tv) - h(\bar{x})}{t} \leq h'_j(\bar{x}; v) = \langle \nabla h_j(\bar{x}), v \rangle \leq \max\{\langle \nabla h_i(\bar{x}), v \rangle \mid i \in A(\bar{x})\}.$$

Esto completa la prueba del lema.  $\square$

**Teorema 3.6** (Condiciones necesarias de Fritz John). *Supongamos que  $\bar{x} \in X$  es un mínimo local de  $f$  sobre  $X$ . Entonces existen escalares no negativos  $\mu_0, \mu_i$  (con  $i \in I(\bar{x})$ ) no todos nulos tales que*

$$\mu_0 \nabla f(\bar{x}) + \sum_{i \in I(\bar{x})} \mu_i \nabla g_i(\bar{x}) = 0. \quad (3.4)$$

PRUEBA. Como  $\bar{x}$  es un mínimo local de  $f$  en  $X$ , y como  $g_i(x) \leq 0$  para todo  $x \in X$  y todo  $i$ , tenemos que  $\bar{x}$  es un mínimo local sobre  $\mathbb{R}^n$  de la función  $h$  definida por

$$h(x) = \max\{f(x) - f(\bar{x}), g_1(x), \dots, g_r(x)\}.$$

Del Lema 3.5 y del Corolario 3.2 deducimos que

$$0 \leq h'(\bar{x}; v) = \max\{\langle \nabla f(\bar{x}), v \rangle, \langle \nabla g_i(\bar{x}), v \rangle \mid i \in I(\bar{x})\}$$

para todo vector  $v \in \mathbb{R}^n$ . Como  $K := \text{co}(\{\nabla f(\bar{x}), \nabla g_i(\bar{x})\}_{i \in I(\bar{x})})$  es compacto por el Corolario 2.5, lo anterior nos indica que no es posible separar fuertemente  $K$  y  $\{0\}$ . Luego, el Teorema 2.13 nos asegura que  $0$  pertenece a  $K$ .  $\square$

Si  $\mu_0 \neq 0$  en (3.4) podemos dividir los  $\mu_i$  entre  $\mu_0$  para obtener  $\lambda = (\lambda_1, \dots, \lambda_r)$  que satisface las condiciones de la definición de multiplicadores de Lagrange. Usamos esta observación en la prueba del siguiente teorema.

**Teorema 3.7** (Condiciones de calificación). *Supongamos que  $\bar{x} \in X$  es un mínimo local de  $f$  sobre  $X$ . Entonces bajo cualquiera de las siguientes hipótesis existe un vector de multiplicadores de Lagrange en  $\bar{x}$ .*

- (1) Existe un vector  $v \in \mathbb{R}^n$  tal que  $\langle \nabla g_i(\bar{x}), v \rangle < 0$  para todo  $i \in I(\bar{x})$ .
- (2) Los vectores  $\{\nabla g_i(\bar{x})\}_{i \in I(\bar{x})}$  son linealmente independientes.
- (3) (a) las funciones  $g_i$  son convexas;  
(b) existe un vector factible  $x_0 \in X$  tal que  $g_i(x_0) < 0$  para todo  $i \in I(\bar{x})$ .

PRUEBA. Del Teorema 3.6 sabemos que existen  $\mu_0, \mu_i$  (con  $i \in I(\bar{x})$ ) tales que (3.4) se cumple. Luego, por la observación que hicimos antes de enunciar el teorema, bastará con probar que  $\mu_0 \neq 0$  en cada caso.

- (1) Tenemos  $\mu_0 \langle \nabla f(\bar{x}), v \rangle = - \sum_{i \in I(\bar{x})} \mu_i \langle \nabla g_i(\bar{x}), v \rangle > 0$ , y por tanto  $\mu_0 \neq 0$ .

(2) Si  $\mu_0 = 0$ , entonces  $\sum_{i \in I(\bar{x})} \mu_i \nabla g_i(\bar{x}) = 0$  donde no todos los  $\mu_i$  son nulos, lo que contradice nuestra hipótesis. Luego,  $\mu_0 \neq 0$ .

(3) Notemos que  $\nabla g_i(\bar{x}) \neq 0$  para cada  $i \in I(\bar{x})$ , ya que  $g_i(\bar{x}) = 0 > g(x_0)$  implica  $\langle \nabla g_i(\bar{x}), x_0 - \bar{x} \rangle < 0$  por el Corolario 2.28. Por continuidad, la condición (b) nos dice que existe  $x_0 \in \text{int}(X)$ . Al ser  $X$  convexo (las funciones  $g_i$  son convexas), el cono normal  $K := N(x, X)$  tiene la propiedad de que  $K \cap (-K) = \{0\}$  en virtud de la Proposición 2.22. Por el Corolario 2.28, para cada  $i \in I(\bar{x})$  tenemos

$$\langle \nabla g_i(\bar{x}), x - \bar{x} \rangle \leq 0, \quad \forall x \in X,$$

esto es,  $\nabla g_i(\bar{x}) \in K$ . Luego, si  $\mu_0 = 0$  en (3.4), entonces  $\mu_j > 0$  para algún  $j \in I(\bar{x})$  y tendríamos  $\mu_j \nabla g_j(\bar{x}) = -\sum_{i \in I(\bar{x}), i \neq j} \mu_i \nabla g_i(\bar{x}) \in -K$ , lo cual implicaría que  $\nabla g_j(\bar{x}) \in K \cap (-K) = \{0\}$ , una contradicción. Así,  $\mu_0 \neq 0$ .  $\square$

La hipótesis (1) es conocida como la **condición de calificación de Mangasarian-Fromovitz**. Intuitivamente, nos indica que el interior del cono convexo y cerrado  $\text{con}(\{\nabla g_i(\bar{x})\}_{i \in I(\bar{x})})^*$  es no vacío.

La hipótesis (2) es llamada **condición de calificación de independencia lineal**. Cuando ella se cumple, el vector  $\lambda$  de multiplicadores de Lagrange está únicamente determinado.

La hipótesis (3), o más específicamente, la condición (3) (b), es frecuentemente llamada **condición de calificación de Slater**. Ella nos dice que  $\text{int}(X) \neq \emptyset$ .

Veamos un ejemplo en el que no existe ningún vector de multiplicadores de Lagrange. Desde luego, en este ejemplo ninguna de las condiciones de calificación que acabamos de introducir vale.

**Ejemplo 3.8.** En  $\mathbb{R}^2$  consideremos  $f(x_1, x_2) = x_1$ ,  $g_1(x_1, x_2) = x_2$  y

$$g_2(x_1, x_2) = \begin{cases} x_1^2 - x_2, & \text{si } x_1 \leq 0, \\ -x_2, & \text{si } x_1 \geq 0. \end{cases}$$

Como  $X = \{x \in \mathbb{R}^2 \mid g_i(x) \leq 0, \forall i = 1, 2\} = [0, \infty) \times \{0\} \subset \mathbb{R}^2$ , el punto  $\bar{x} := (0, 0)$  es un mínimo local (en realidad global) de  $f$  sobre  $X$ . El conjunto activo  $I(\bar{x})$  es  $\{1, 2\}$ , y los gradientes son

$$\nabla f(\bar{x}) = (1, 0), \quad \nabla g_1(\bar{x}) = (0, 1) \quad \text{y} \quad \nabla g_2(\bar{x}) = (0, -1).$$

Ahora bien, observando la Figura 3.1 deducimos que ninguna CQ se satisface en  $\bar{x}$ , y que tampoco existen multiplicadores de Lagrange en  $\bar{x}$ .

Si la condición (3) (a) es satisfecha (y por tanto  $X$  es convexo), entonces la existencia de los multiplicadores de Lagrange garantiza la optimalidad cuando  $f$  es convexa y diferenciable (o pseudoconvexa).

**Teorema 3.9** (Condición suficiente de optimalidad). *Supongamos que las funciones  $g_i$  son convexas. Supongamos también que  $f$  es convexa<sup>8</sup> (o, más generalmente que  $f$  es pseudoconvexa). Si en un punto factible dado  $\bar{x} \in X$  existen los multiplicadores de Lagrange, entonces  $\bar{x}$  minimiza  $f$  sobre  $X$ .*

<sup>8</sup>Recordemos que  $f$  ya es diferenciable.

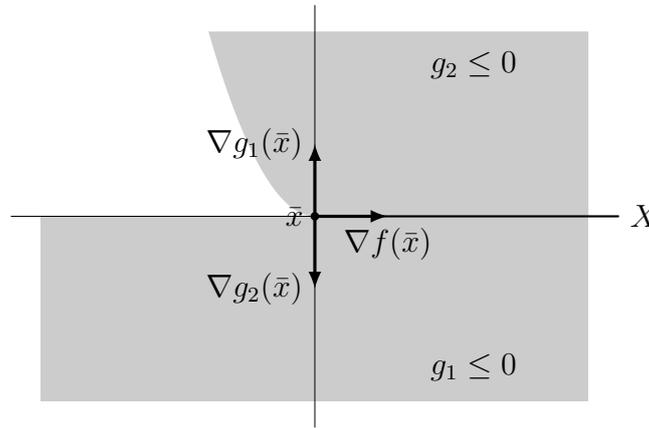


Figura 3.1: Gráfico del Ejemplo 3.8.

PRUEBA. Sea  $\lambda = (\lambda_1, \dots, \lambda_r)$  un vector de multiplicadores de Lagrange en  $\bar{x}$ . Sabemos que  $\nabla f(\bar{x}) = -\sum_{i \in I(\bar{x})} \lambda_i \nabla g_i(\bar{x})$ . Consideremos un vector arbitrario  $x$  de  $X$ . Para cada  $i \in I(\bar{x})$  la función  $g_i$  es convexa, de modo que  $g_i(x) \leq 0 = g_i(\bar{x})$  implica  $\langle \nabla g_i(\bar{x}), x - \bar{x} \rangle \leq 0$  por el Corolario 2.28. Luego

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle = \sum_{i \in I(\bar{x})} \lambda_i \cdot \langle -\nabla g_i(\bar{x}), x - \bar{x} \rangle \geq 0,$$

y el Corolario 2.28 nos permite concluir que  $f(x) \geq f(\bar{x})$ . Si  $f$  es pseudoconvexa, en vez del Corolario 2.28, usamos la definición de función pseudoconvexa.  $\square$

**Observación 3.10.** Hemos mostrado que cuando las funciones  $g_i$  son convexas y existe  $\lambda = (\lambda_1, \dots, \lambda_r)$  tal que  $0 \leq \lambda \perp g(\bar{x}) \leq 0$ , entonces todo vector de la forma  $v = -\sum_i \lambda_i \nabla g_i(\bar{x})$  satisface la propiedad

$$\langle v, x - \bar{x} \rangle \geq 0, \quad \forall x \in X.$$

## 4. Análisis multivaluado

La noción de operador multivaluado (o multifunción) generaliza la noción clásica de función. Una función  $X \rightarrow Y$  asigna a cada elemento de su dominio un *único* elemento de su codominio; o sea que no pueden existir  $y \neq y'$  en  $Y$  asociados con  $x \in X$ . Con las multifunciones esto *sí puede ocurrir*. De hecho, una multifunción  $X \rightrightarrows Y$  asocia a cada elemento de su dominio un *subconjunto* de su codominio.

El análisis multivaluado estudia los operadores multivaluados, con el mismo espíritu con que en análisis se estudian las funciones. Una gran parte del análisis multivaluado surgió del análisis convexo y la optimización (por ejemplo, los subdiferenciales). Siendo el análisis multivaluado una generalización del análisis matemático, existen generalizaciones de las nociones de continuidad, diferenciabilidad, puntos fijos, etc.

En esta sección hacemos un resumen de todos los resultados sobre multifunciones que necesitaremos en los capítulos subsiguientes. La referencia por excelencia sobre este material es el libro clásico [1]. También usaremos la referencia [8].

Sean  $X$  e  $Y$  subconjuntos de  $\mathbb{R}^n$  y  $\mathbb{R}^m$  respectivamente. Una **multifunción** u **operador multivaluado**  $\Omega : X \rightrightarrows Y$  es una aplicación que asocia a cada  $x \in X$  con un subconjunto de  $Y$  denotado por  $\Omega(x)$ . Los conjuntos  $\Omega(x)$  son las **imágenes** de  $\Omega$  y los conjuntos

$$\begin{aligned} \text{dom}(\Omega) &:= \{x \in X \mid \Omega(x) \neq \emptyset\}, \\ G(\Omega) &:= \{(x, y) \in X \times Y \mid y \in \Omega(x)\}, \\ \text{im}(\Omega) &:= \{y \in Y \mid \exists x \in X : y \in \Omega(x)\}, \end{aligned}$$

son el **dominio**, el **gráfico** y la **imagen** de  $\Omega$  respectivamente. Dado  $M \subset X$  el conjunto  $\Omega(M) := \bigcup_{x \in M} \Omega(x)$  es la **imagen** de  $M$ .

Decimos que  $\Omega$  tiene imágenes abiertas (cerradas, compactas, etc.) si cada imagen  $\Omega(x)$  es abierta (cerrada, compacta, etc.) con respecto a la topología relativa heredada de  $Y$ .

Asociados a  $\Omega$ , definimos el **operador multivaluado inverso**  $\Omega^{-1} : Y \rightrightarrows X$  y el **operador multivaluado dual**  $\Omega^* : Y \rightrightarrows X$  del siguiente modo

$$\begin{aligned} \Omega^{-1}(y) &= \{x \in X \mid y \in \Omega(x)\}, \\ \Omega^*(y) &= X \setminus \Omega^{-1}(y). \end{aligned}$$

Las imágenes de  $\Omega^{-1}$  son las **fibras** de  $\Omega$  y las imágenes de  $\Omega^*$  son las **cofibras** de  $\Omega$ . Evidentemente,  $X$  es la unión disjunta de  $\Omega^{-1}(y)$  y  $\Omega^*(y)$  para cada  $y \in Y$ .

**Ejemplo 4.1.** Toda función  $f : X \rightarrow Y$  define de modo natural una multifunción. En efecto, basta con considerar  $F : X \rightrightarrows Y$  como  $F(x) = \{f(x)\}$ . La Figura 4.1 muestra el gráfico del operador  $\Omega : \mathbb{R} \rightrightarrows \mathbb{R}$  definido por  $\Omega(x) = (x, \infty)$  si  $x > 0$ ,  $\Omega(x) = \emptyset$  si  $x < 0$  y  $\Omega(0) = [0, \infty)$ . El dominio y la imagen de  $\Omega$  son iguales a  $[0, \infty)$ .

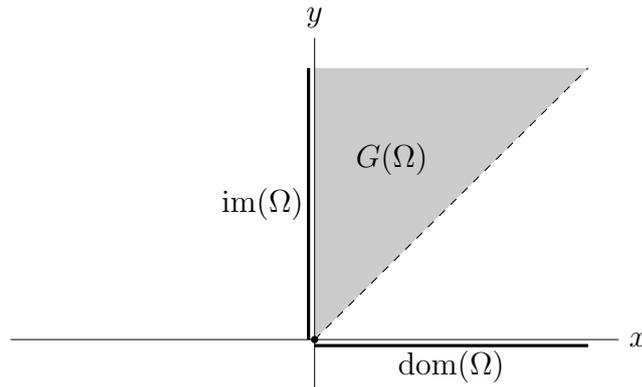


Figura 4.1: Gráfico de la multifunción del Ejemplo 4.1

Si  $Z$  es un subconjunto de  $X$ , de manera natural definimos la **restricción** de  $\Omega$  a  $Z$  como  $\Omega|_Z : Z \rightrightarrows Y$ ,  $\Omega|_Z(z) = \Omega(z)$ .

Notemos que  $\Omega$  y  $\Omega^{-1}$  están totalmente determinadas por el gráfico de  $\Omega$  pues

$$y \in \Omega(x) \quad \Leftrightarrow \quad x \in \Omega^{-1}(y) \quad \Leftrightarrow \quad (x, y) \in G(\Omega).$$

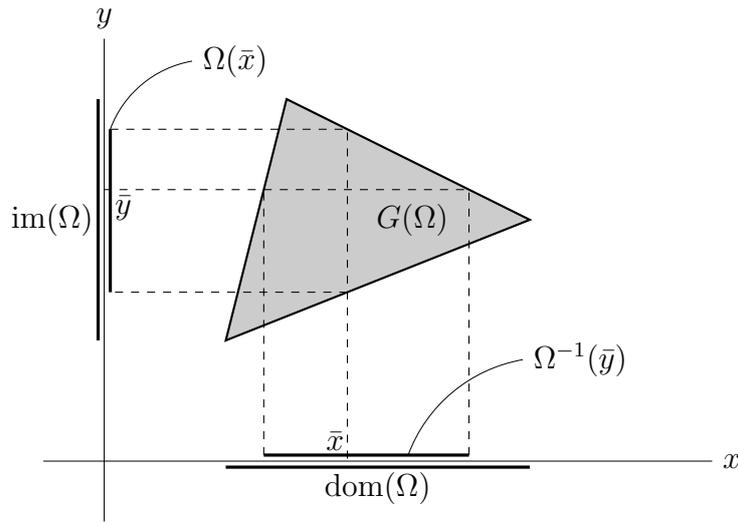


Figura 4.2: Gráfico de un operador multivaluado  $\Omega$ . Dados  $\bar{x}, \bar{y} \in \mathbb{R}$ , hallamos  $\Omega(\bar{x})$  y  $\Omega^{-1}(\bar{y})$  a partir de  $G(\Omega)$ .

Así, dados  $\bar{x} \in X$  e  $\bar{y} \in Y$ ,  $\Omega(\bar{x})$  es la proyección sobre  $Y$  del conjunto  $\{(x, y) \in G(\Omega) \mid x = \bar{x}\}$ , y análogamente,  $\Omega^{-1}(\bar{y})$  es la proyección sobre  $X$  del conjunto  $\{(x, \bar{y}) \in G(\Omega) \mid x = \bar{x}\}$  (véase la Figura 4.2). De hecho existe una biyección natural entre las multifunciones  $\Omega : X \rightrightarrows Y$  y los subconjuntos de  $X \times Y$ .

El **producto** de los operadores multivaluados  $\Omega : X \rightrightarrows Y$  y  $\Omega' : X \rightrightarrows Y'$ , donde  $Y' \subset \mathbb{R}^p$ , es el operador multivaluado  $\Omega \times \Omega' : X \rightrightarrows Y \times Y'$  dado por  $\Omega \times \Omega'(x) = \Omega(x) \times \Omega'(x)$ .

**Definición 4.2.** Sea  $\Omega : X \rightrightarrows Y$  una multifunción. Dado  $\bar{x} \in X$ , decimos que

- (1)  $\Omega$  es **semicontinua inferiormente (sci)** en  $\bar{x}$  si para todo  $V$  abierto en  $Y$  con  $V \cap \Omega(\bar{x}) \neq \emptyset$ , existe un abierto  $U$  en  $X$  que contiene a  $\bar{x}$  tal que  $V \cap \Omega(x) \neq \emptyset$  para todo  $x \in U$ .
- (2)  $\Omega$  es **semicontinua superiormente (scs)** en  $\bar{x} \in X$  si para todo  $V$  abierto en  $Y$  que contiene a  $\Omega(\bar{x})$  existe  $U$  abierto en  $X$  tal que  $\Omega(x) \subset V$  para todo  $x \in U$ .
- (3)  $\Omega$  es **cerrado** en  $\bar{x}$  si para toda sucesión  $(x_k, y_k)$  en  $G(\Omega)$  que converge a  $(\bar{x}, \bar{y}) \in X \times Y$  se cumple que  $\bar{y} \in \Omega(\bar{x})$ .
- (4)  $\Omega$  es **acotado (no vacío) cerca de  $\bar{x}$**  si existe un abierto  $U$  en  $X$  que contiene a  $\bar{x}$  y tal que  $\Omega(U)$  es acotado ( $\Omega(x) \neq \emptyset$  para cada  $x \in U$ ).
- (5)  $\bar{x} \in X$  es un **punto fijo** de  $\Omega$  si  $\bar{x} \in \Omega(\bar{x})$ . (En este caso,  $m = n$ .)

Cuando  $\Omega$  es sci (scs, cerrado) en cada uno de los puntos de  $X$ , decimos simplemente que  $\Omega$  es **sci (scs, cerrado)**.

**Ejemplo 4.3.** El operador  $F : X \rightrightarrows Y$ ,  $F(x) = \{f(x)\}$  asociado a una función continua  $f : X \rightarrow Y$  es sci, scs y cerrado. El operador  $\Omega$  del Ejemplo 4.1 es sci en todo  $\mathbb{R}$  excepto en 0, es scs (cerrado) solo en 0, y tiene como único punto fijo a 0.

**Ejemplo 4.4** (La composición con funciones continuas preserva la semicontinuidad). Si  $\Omega : X \rightrightarrows Y$  es un operador multivaluado sci (scs) y  $f : W \rightarrow X$  y  $g : Y \rightarrow Z$  son funciones continuas, entonces las multifunciones  $\Omega \circ f : W \rightrightarrows Y$  y  $g \circ \Omega : X \rightrightarrows Z$ , dadas por  $w \mapsto \Omega(f(w))$  y  $x \mapsto g(\Omega(x))$  respectivamente, también son sci (scs).

**Ejemplo 4.5** (El producto de operadores sci es sci). Si  $\Omega : X \rightrightarrows Y$  y  $\Omega' : X \rightrightarrows Y'$  son operadores sci, entonces el producto  $\Omega \times \Omega' : X \rightrightarrows Y \times Y'$  también es sci. En efecto, si  $\bar{x} \in X$  y el abierto  $V \subset Y \times Y'$  se interseca con  $\Omega(\bar{x}) \times \Omega'(\bar{x})$ , entonces existen abiertos  $W \subset Y$  y  $W' \subset Y'$  tales que  $W \times W' \subset V$  también se interseca con  $\Omega(\bar{x}) \times \Omega'(\bar{x})$ , es decir,  $W \cap \Omega(\bar{x}) \neq \emptyset$  y  $W' \cap \Omega'(\bar{x}) \neq \emptyset$ , por lo cual existe una vecindad abierta  $U$  de  $\bar{x}$  tal que  $W \cap \Omega(x) \neq \emptyset$  y  $W' \cap \Omega'(x) \neq \emptyset$  para todo  $x \in U$ .

**Ejemplo 4.6** (El producto de operadores scs con imágenes compactas es scs). Si  $\Omega : X \rightrightarrows Y$  y  $\Omega' : X \rightrightarrows Y'$  son operadores scs cuyas imágenes son compactas, entonces el producto  $\Omega \times \Omega' : X \rightrightarrows Y \times Y'$  también es scs (con imágenes compactas). En efecto, consideremos  $\bar{x} \in X$  y un abierto  $V \subset Y \times Y'$  que contiene a  $\Omega(\bar{x}) \times \Omega'(\bar{x})$ , y veamos que existe una vecindad abierta  $U$  de  $\bar{x}$  tal que  $\Omega(x) \times \Omega'(x) \subset V$  para todo  $x \in U$ . Como el conjunto compacto  $\Omega(\bar{x}) \times \Omega'(\bar{x})$  y el conjunto cerrado  $(Y \times Y') \setminus V$  son disjuntos, la distancia entre ellos es positiva (véase la prueba del ítem (2) del Teorema 2.13). Luego, si  $\varepsilon$  es suficientemente pequeño el producto  $W \times W'$  de las vecindades abiertas  $W := B_Y(\Omega(\bar{x}), \varepsilon)$  y  $W' := B_{Y'}(\Omega'(\bar{x}), \varepsilon)$  de  $\Omega(\bar{x})$  y  $\Omega'(\bar{x})$  está contenida en  $V$ . Usando la semicontinuidad superior de  $\Omega$  y  $\Omega'$  con  $W$  y  $W'$ , conseguimos la vecindad  $U$  requerida.

De los tres ejemplos anteriores deducimos que  $\Omega : X \rightrightarrows Y$  y  $\Omega' : X \rightrightarrows Y'$  son sci (scs con imágenes compactas) si y solo si su producto es sci (scs con imágenes compactas).

**Observación 4.7.** Si  $V \subset Y$  contiene la imagen  $\text{im}(\Omega)$  de  $\Omega$  entonces el operador  $\hat{\Omega} : X \rightrightarrows V$ ,  $\hat{\Omega}(x) = \Omega(x)$  tiene las mismas propiedades que  $\Omega$ , esto es,  $\hat{\Omega}$  es sci (scs, cerrado, acotado cerca de, no vacío cerca de)  $\bar{x} \in X$  si y solo si  $\Omega$  lo es en  $\bar{x}$ . Esto quiere decir que  $\Omega$  y  $\hat{\Omega}$  son esencialmente iguales y que el conjunto de llegada  $Y$  no importa. Así, por ejemplo, da lo mismo trabajar con operadores  $X \rightrightarrows Y$  que con operadores  $X \rightrightarrows \mathbb{R}^m$  cuyas imágenes están contenidas en  $Y$ .

Notemos que  $\Omega$  es cerrado si y solo si para cada  $\bar{x} \in X$  los vectores  $(\bar{x}, \bar{y}) \in X \times Y$  que son puntos de acumulación de  $G(\Omega)$  están en  $G(\Omega)$ , esto es,

$$\Omega \text{ es cerrado si y solo si } G(\Omega) \text{ es cerrado en } X \times Y;$$

en particular,  $\Omega$  es cerrado si y solo si  $\Omega^{-1}$  es cerrado. Luego, si  $\Omega$  es cerrado en  $\bar{x}$  entonces  $\Omega(\bar{x})$  es cerrado en  $Y$ ; por consiguiente, todo operador cerrado  $\Omega$  tiene imágenes cerradas en  $Y$ . Asimismo, si  $\Omega$  tiene imágenes vacías en un abierto  $U$  de  $X$  entonces  $\Omega$  es cerrado en los puntos de  $U$ .

Un vector  $\bar{x} \in X$  es punto fijo de  $\Omega$  si y solo si lo es también de  $\Omega^{-1}$ , y todo esto se cumple si y solo si  $(\bar{x}, \bar{x}) \in G(\Omega)$ .

Si  $\Omega$  es sci (scs) y  $Z \subset X$  entonces la restricción  $\Omega|_Z$  también es sci (scs).

Si  $\Omega(\bar{x}) = \emptyset$  entonces  $\Omega$  es automáticamente sci en  $\bar{x}$ . Cuando  $\Omega(\bar{x}) \neq \emptyset$ , el hecho de que  $\Omega$  sea sci en  $\bar{x}$  equivale a la siguiente implicancia

$$\left. \begin{array}{l} (x_k) \subset X, x_k \rightarrow \bar{x} \\ \bar{y} \in \Omega(\bar{x}) \end{array} \right\} \Rightarrow \exists (y_k) \subset Y, \exists N \in \mathbb{N} : y_k \rightarrow \bar{y}, y_k \in \Omega(x_k) \text{ para } k \geq N, \quad (4.1)$$

que también es válida si  $\Omega(\bar{x}) = \emptyset$ . Así, hemos probado el siguiente lema.

**Lema 4.8.** *Una multifunción  $\Omega : X \rightrightarrows Y$  es sci en  $\bar{x} \in X$  si y solo si la implicancia (4.1) se satisface en  $\bar{x}$ . En particular, cuando  $\Omega$  es sci en  $\bar{x}$ , si  $x_k \rightarrow \bar{x}$  y  $\bar{x} \in \text{dom}(\Omega)$  entonces  $x_k \in \text{dom}(\Omega)$  para  $k$  suficientemente grande, es decir,  $\bar{x}$  está en el interior de  $\text{dom}(\Omega)$ .*

De la definición (o del Lema 4.8) se sigue que si  $\Omega : X \rightrightarrows Y$  es sci entonces su dominio es abierto en  $X$ .

Dado  $B \subset Y$ , definimos el conjunto

$$\Omega^-B = \{x \in X \mid \Omega(x) \cap B \neq \emptyset\}.$$

Luego,  $\Omega$  es sci en  $\bar{x} \in X$  si y solo si para todo abierto  $V$  en  $Y$  con  $\bar{x} \in \Omega^-V$ , existe un abierto  $U$  en  $X$  tal que  $\bar{x} \in U \subset \Omega^-V$ , esto es, si y solo si  $\Omega^-V$  es una vecindad de  $\bar{x}$  para todo abierto  $V \subset Y$  tal que  $\bar{x} \in \Omega^-V$ . De esto se sigue que  $\Omega$  es sci si y solo si para todo abierto  $V \subset Y$  se cumple que  $\Omega^-V$  es una vecindad de cada uno de sus puntos, es decir, si es abierto. Resumimos esto en la siguiente proposición.

**Proposición 4.9.** *El operador  $\Omega : X \rightrightarrows Y$  es sci en  $\bar{x} \in X$  si y solo si  $\Omega^-V$  es una vecindad de  $\bar{x}$  para cada abierto  $V \subset Y$  tal que  $\Omega(\bar{x}) \cap V \neq \emptyset$ . Además,  $\Omega$  es sci si y solo si  $\Omega^-V$  es abierto en  $X$  para cada abierto  $V \subset Y$ .*

El hecho de que  $\Omega$  sea scs en  $\bar{x} \in X$  equivale a la siguiente implicancia

$$V \supset \Omega(\bar{x}) \text{ abierto, } (x_k) \subset X, x_k \rightarrow \bar{x} \Rightarrow \exists N \in \mathbb{N} : \Omega(x_k) \subset V, \forall k \geq N,$$

y por tanto obtenemos el siguiente lema.

**Lema 4.10.** *El operador  $\Omega : X \rightrightarrows Y$  no es scs en  $\bar{x} \in X$  si y solo si existen una sucesión  $(x_k)$  convergente a  $\bar{x}$ , un abierto  $V$  en  $Y$  que contiene a  $\Omega(\bar{x})$  y una sucesión  $(y_k)$  en  $Y$  tales que*

$$y_k \in \Omega(x_k) \text{ e } y_k \notin V \text{ para todo } k.$$

Dado un subconjunto  $M$  de  $Y$ , definimos

$$\Omega^+M = \{x \in X \mid \Omega(x) \subset M\}.$$

Luego,  $\Omega$  es scs en  $\bar{x} \in X$  si y solo si para todo  $V$  abierto en  $Y$  con  $\Omega(\bar{x}) \subset V$ , existe un abierto  $U$  en  $X$  con  $\bar{x} \in U \subset \Omega^+V$ , esto es, si y solo si  $\Omega^+V$  es una vecindad de  $\bar{x}$  para cada abierto  $V \subset Y$  tal que  $\bar{x} \in \Omega^+V$ . De esto se sigue que  $\Omega$  es scs si y solo si para todo abierto  $V \subset Y$  se cumple que  $\Omega^+V$  es una vecindad de cada uno de sus puntos, es decir, si es abierto. Resumimos estas observaciones en la siguiente proposición.

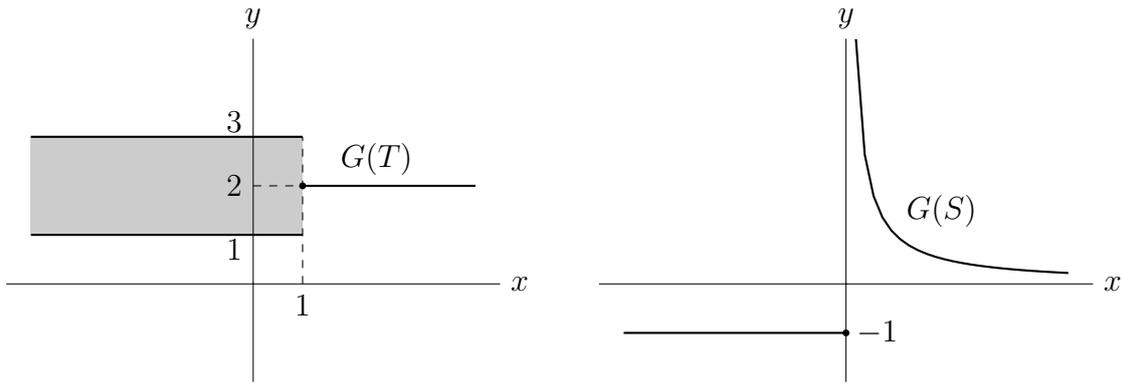


Figura 4.3: Ilustración del gráfico de los operadores  $T$  y  $S$  del Ejemplo 4.13.

**Proposición 4.11.** *El operador  $\Omega : X \rightrightarrows Y$  es scs en  $\bar{x} \in X$  si y solo si  $\Omega^+(V)$  es una vecindad de  $\bar{x}$  para cada abierto  $V \subset Y$  tal que  $\Omega(\bar{x}) \subset V$ . Además,  $\Omega$  es scs si y solo si  $\Omega^+(V)$  es abierto en  $X$  para cada abierto  $V \subset Y$ .*

**Observación 4.12.** Las Proposiciones 4.9 y 4.11 nos dan condiciones necesarias y suficientes para que la semicontinuidad inferior y superior de operadores multivaluados ocurra. También, nos recuerda un poco la continuidad de funciones, a saber, que una función es continua si la preimagen de todo abierto es abierto. En realidad, toda función es esencialmente una multifunción con imágenes unitarias, y el hecho de que ella sea sci (scs) como multifunción equivale a que sea continua como función.

**Ejemplo 4.13.** Un ejemplo de operador que no es scs es la multifunción  $T : \mathbb{R} \rightrightarrows \mathbb{R}$  definida por  $T(x) = [1, 3]$  si  $x < 1$  y  $T(x) = \{2\}$  si  $x \geq 1$ , que no es scs en 1. Mas el operador cuyo gráfico es  $\overline{G(T)}$  sí es scs. Esto podría sugerir que si el gráfico de un operador  $\Omega$  es cerrado entonces es scs, pero esto es falso: la multifunción  $S : \mathbb{R} \rightrightarrows \mathbb{R}$  definida mediante  $S(x) = \{-1\}$  si  $x \leq 0$  y  $S(x) = \{1/x\}$  si  $x > 0$  tiene gráfico cerrado pero no es scs en 0. No obstante, cuando la imagen de  $\Omega$  es acotada, se verifica que si el gráfico de  $\Omega$  es cerrado entonces  $\Omega$  es scs.

**Proposición 4.14.** *Supongamos que  $Y \subset \mathbb{R}^m$  es cerrado y que la imagen  $\Omega(X)$  del operador  $\Omega : X \rightrightarrows Y$  es acotada. Si  $\Omega$  es cerrado, entonces también es scs.*

PRUEBA. Supongamos que  $\Omega$  no es scs en  $x \in X$ . Por el Lema 4.10 existen una sucesión  $(x_k)$  convergente a  $x$ , un abierto  $V$  en  $Y$  que contiene a  $\Omega(x)$  y una sucesión  $(y_k)$  en  $Y$  tales que

$$y_k \in \Omega(x_k), \quad y_k \notin V, \quad \forall k \geq 1.$$

Dado que la sucesión  $(y_k)$  es acotada ( $\Omega$  tiene imagen acotada) y el conjunto  $Y$  es cerrado, podemos suponer sin pérdida de generalidad que  $(y_k)$  converge a algún  $y \in Y$ , que no está en  $V$  porque  $V$  es abierto. Pero como  $(x_k, y_k)$  es una sucesión en  $G(\Omega)$  que converge a  $(x, y)$ , también tenemos que  $(x, y) \in G(\Omega)$  pues  $G(\Omega)$  es cerrado, es decir,  $y \in \Omega(x) \subset V$ . Esta contradicción muestra que  $\Omega$  es scs.  $\square$

Recíprocamente, la semicontinuidad superior de  $\Omega$  implica que  $\Omega$  es cerrado en algunos casos.

**Proposición 4.15.** *Si el operador multivaluado  $\Omega : X \rightrightarrows Y$  es scs y tiene imágenes cerradas, entonces es cerrado.*

PRUEBA. Tomemos una sucesión  $(x_k, y_k)$  en  $G(\Omega)$  que converge a  $(\bar{x}, \bar{y})$ . Debemos mostrar que  $(\bar{x}, \bar{y}) \in G(\Omega)$ , esto es, que  $\bar{y} \in \Omega(\bar{x})$ . Supongamos que  $\bar{y} \notin \Omega(\bar{x})$ . Como  $\Omega(\bar{x})$  es cerrado, existe una vecindad abierta  $V \subset \mathbb{R}^m$  de  $\bar{y}$  tal que  $\bar{V} \subset \mathbb{R}^m \setminus \Omega(\bar{x})$ . Entonces  $\mathbb{R}^m \setminus V$  es una vecindad de  $\Omega(\bar{x})$  y por la semicontinuidad superior de  $\Omega$  tenemos  $y_k \in \Omega(x_k) \subset \mathbb{R}^m \setminus V$  para  $k$  suficientemente grande (pues  $x_k \rightarrow \bar{x}$ ), lo cual, considerando que  $V$  es abierto y que  $y_k \rightarrow \bar{y}$ , implica  $\bar{y} \notin V$ , una evidente contradicción. Así,  $\bar{y} \in \Omega(\bar{x})$  y concluimos que  $\Omega$  es cerrado.  $\square$

Ahora veamos otros resultados relacionados con la semicontinuidad superior de operadores multivaluados.

**Proposición 4.16.** *Sea el conjunto compacto  $K \subset \mathbb{R}^n$ . Si el operador multivaluado  $\Omega : K \rightrightarrows \mathbb{R}^m$  es scs y tiene imágenes compactas entonces  $\Omega(K) \subset \mathbb{R}^m$  es compacto.*

PRUEBA. Sea  $\{U_\lambda\}_{\lambda \in \Lambda}$  un cubrimiento abierto de  $\Omega(K)$ . Entonces  $\{U_\lambda\}_{\lambda \in \Lambda}$  es un cubrimiento abierto de cada imagen  $\Omega(x)$  de  $\Omega$ , la cual, por ser compacta, admite un subcubrimiento finito  $\{U_\lambda\}_{\lambda \in \Lambda_x}$ . Cada conjunto

$$V_x := \Omega^{+1}\left(\bigcup_{\lambda \in \Lambda_x} U_\lambda\right) = \left\{z \in K \mid \Omega(z) \subset \bigcup_{\lambda \in \Lambda_x} U_\lambda\right\}$$

es abierto por la Proposición 4.11, y claramente contiene a  $x$ . Puesto que la familia  $\{V_x\}_{x \in K}$  es un cubrimiento abierto del conjunto compacto  $K$ , existen  $x_1, \dots, x_r \in K$  tales que  $K \subset \bigcup_{i=1}^r V_{x_i}$ , y en consecuencia

$$\Omega(K) \subset \bigcup_{i=1}^r \Omega(V_{x_i}) \subset \bigcup_{i=1}^r \bigcup_{\lambda \in \Lambda_{x_i}} U_\lambda = \bigcup_{\lambda \in \Lambda'} U_\lambda,$$

donde  $\Lambda' = \Lambda_{x_1} \cup \dots \cup \Lambda_{x_r}$  es finito. Así, hemos obtenido un subcubrimiento finito de  $\Omega(K)$ , de donde concluimos que  $\Omega(K)$  es compacto.  $\square$

Una **selección continua** de una multifunción  $\Omega : X \rightrightarrows Y$  es una función continua  $f : X \rightarrow Y$  con la propiedad de que  $f(x) \in \Omega(x)$  para todo  $x \in X$ . Equivalentemente, una función continua  $f : X \rightarrow Y$  es una selección de  $\Omega$  si su gráfico está contenido en el gráfico de  $\Omega$ .

Es claro que no toda multifunción  $\Omega$  tiene una selección continua. Pero si imponemos ciertas condiciones en  $\Omega$  entonces tal selección existe. Por ejemplo, el teorema de selección de Michael [8, Teorema 8.2.8] nos asegura que si  $K \subset \mathbb{R}^n$  es cerrado y  $\Omega : K \rightrightarrows \mathbb{R}^m$  es sci y tiene imágenes convexas, cerradas y no vacías entonces, para todo punto  $(\bar{x}, \bar{y})$  del gráfico de  $\Omega$  existe una selección continua  $f : K \rightarrow \mathbb{R}^m$  cuyo gráfico pasa por  $(\bar{x}, \bar{y})$ , esto es, tal que  $f(\bar{x}) = \bar{y}$ .

Si cambiamos la hipótesis de semicontinuidad inferior en el teorema de Michael por la de semicontinuidad superior entonces la conclusión es falsa, incluso si  $K$  es compacto: el operador  $\Omega : [-1, 1] \rightrightarrows [-1, 1]$  dado por  $\Omega(x) = \{\text{sgn}(x)\}$  si  $x \neq 0$  y  $\Omega(0) = [-1, 1]$  no admite una selección continua a pesar de que es scs y tiene imágenes convexas, compactas y no vacías. No obstante, aunque en este caso no existe una selección continua, existe una selección “aproximada”.

**Teorema 4.17** (Cellina). *Dado el conjunto compacto  $K \subset \mathbb{R}^n$ , supongamos que el operador  $\Omega : K \rightrightarrows \mathbb{R}^m$  es scs y tiene imágenes convexas y no vacías. Entonces, para todo  $\varepsilon > 0$  existe una función continua  $f : K \rightarrow \mathbb{R}^m$  tal que*

$$d_{G(\Omega)}(x, f(x)) < \varepsilon, \quad \forall x \in K.$$

*Además, la imagen de  $f$  está contenida en la cápsula convexa de la imagen de  $\Omega$ , esto es,  $f(K) \subset \text{co}(\Omega(K))$ .*

PRUEBA. Sea  $\varepsilon > 0$ . Como las normas en  $\mathbb{R}^n \times \mathbb{R}^m$  son equivalentes, no hay pérdida de generalidad si trabajamos con la norma

$$\|(x, y)\|_* = \|x\|_1 + \|y\|_2, \quad (x, y) \in \mathbb{R}^n \times \mathbb{R}^m,$$

donde  $\|x\|_1$  es la norma de  $x$  en  $\mathbb{R}^n$  y  $\|y\|_2$  es la norma de  $y$  en  $\mathbb{R}^m$ . Para cada  $x \in K$  el conjunto  $B(\Omega(x), \varepsilon/2)$  es abierto, y como  $\Omega$  scs existe  $\delta_x < \varepsilon$  positivo tal que

$$\Omega(z) \subset B(\Omega(x), \varepsilon/2) \quad \text{para todo } z \in K \text{ tal que } \|z - x\|_1 < \delta_x. \quad (4.2)$$

Como la familia  $\{B(x, \delta_x/2)\}_{x \in K}$  constituye un cubrimiento abierto del conjunto compacto  $K$ , existen  $x_1, \dots, x_r \in K$  tales que  $K \subset \bigcup_{i=1}^r B(x_i, \delta_i/2)$ , donde  $\delta_i := \delta_{x_i}$ . Aplicando el Teorema 1.24 al cubrimiento finito  $\{B(x_i, \delta_i/2)\}_i$  de  $K$ , obtenemos funciones continuas  $\kappa_1, \dots, \kappa_r : K \rightarrow [0, 1]$  tales que  $\sum_{i=1}^r \kappa_i(x) = 1$  si  $x \in K$  y  $\kappa_i(x) = 0$  si  $x \in K \setminus B(x_i, \delta_i/2)$ . Escogiendo para cada  $i$  un elemento  $y_i \in \Omega(x_i)$ , definimos

$$f(x) := \sum_{i=1}^r \kappa_i(x) y_i, \quad x \in K.$$

Veamos que la función continua  $f : K \rightarrow \mathbb{R}^m$  satisface la propiedad deseada. En primer lugar, de la definición de  $f$  es claro que  $f(K) \subset \text{co}(\Omega(K))$  pues cada  $y_i$  está en  $\Omega(x_i) \subset \Omega(K)$ . Por otro lado, dado  $x \in K$  arbitrario, para probar que  $d_{G(\Omega)}(x, f(x)) < \varepsilon$  basta con mostrar que

$$f(x) \in B(\Omega(x_j), \varepsilon/2)$$

para algún índice  $j$  en  $I := \{i \mid x \in B(x_i, \delta_i/2)\}$ . De hecho, en este caso existe  $y_j \in \Omega(x_j)$  tal que  $\|y_j - f(x)\|_2 < \varepsilon/2$ , o sea que  $(x_j, y_j) \in G(\Omega)$  verifica

$$\|(x, f(x)) - (x_j, y_j)\|_* = \|x - x_j\|_1 + \|f(x) - y_j\|_2 < \delta_j/2 + \varepsilon/2 < \varepsilon.$$

Ahora bien,  $f(x)$  es una combinación convexa de los puntos  $\{y_i\}_{i \in I}$ , ya que  $\kappa_i(x) = 0$  cuando  $i \notin I$ . Escogiendo  $j \in I$  tal que  $\delta_j = \max\{\delta_i \mid i \in I\}$ , para cada  $i \in I$  tenemos

$$\|x_i - x_j\|_1 \leq \|x_i - x\|_1 + \|x - x_j\|_1 < \delta_i/2 + \delta_j/2 \leq \delta_j,$$

lo cual implica  $\Omega(x_i) \subset B(\Omega(x_j), \varepsilon/2)$  por (4.2). Así, los vectores  $\{y_i\}_{i \in I}$  están en el conjunto convexo  $B(\Omega(x_j), \varepsilon/2)$ , de donde concluimos que  $f(x) \in B(\Omega(x_j), \varepsilon/2)$ .  $\square$

Antes de continuar enunciamos el teorema de punto fijo de Brouwer.

**Teorema 4.18** (Brouwer). *Si  $C \subset \mathbb{R}^n$  es un conjunto convexo, compacto y no vacío, entonces toda función continua  $f : C \rightarrow C$  posee algún punto fijo (es decir, existe  $\bar{x} \in C$  tal que  $f(\bar{x}) = \bar{x}$ ).*

Este teorema tuvo un papel fundamental en el siglo XX. Por ejemplo, se le puede utilizar en ecuaciones diferenciales (para probar el teorema de Hartman-Grobman) o en teoría de juegos (John Nash la usó para probar la existencia de solución del NEP). Se trata de uno de los primeros *teoremas de punto fijo*, que motivó el subsecuente desarrollo de una nueva área, llamada Teoría del Punto Fijo [14].

Como ya señalamos al inicio de este capítulo, no vamos a probar el teorema de punto fijo de Brouwer, pues requiere de herramientas avanzadas, que están fuera del alcance de esta tesis<sup>9</sup>. No obstante, con el teorema de Cellina podemos generalizarlo al caso de multifunciones.

**Teorema 4.19** (Kakutani-Fan). *Si el conjunto  $C \subset \mathbb{R}^n$  es convexo, compacto y no vacío, entonces todo operador scs  $\Omega : C \rightrightarrows C$  que tiene imágenes convexas, compactas y no vacías posee un punto fijo.*

PRUEBA. El operador  $\Omega$  es cerrado en virtud de la Proposición 4.15. Para cada  $k \in \mathbb{N}$  existe por el Teorema 4.17 una función continua  $f_k : C \rightarrow \mathbb{R}^n$  tal que

$$d_{G(\Omega)}(x, f_k(x)) < \frac{1}{k}, \quad \forall x \in C$$

y con la propiedad de que  $f_k(C) \subset \text{co}(\Omega(C)) \subset \text{co}(C) = C$ . Del teorema del punto fijo de Brouwer se sigue que cada  $f_k : C \rightarrow C$  tiene un punto fijo  $x_k$  que satisface

$$d_{G(\Omega)}(x_k, x_k) = d_{G(\Omega)}(x_k, f_k(x_k)) < \frac{1}{k}. \quad (4.3)$$

Como  $C$  es compacto, la sucesión  $(x_k)$  tiene una subsucesión  $(x_j)$  que converge a algún  $\bar{x} \in C$ . Luego  $(x_j, x_j) \rightarrow (\bar{x}, \bar{x}) \in C \times C$ , y tomando límite en (4.3) obtenemos  $d_{G(\Omega)}(\bar{x}, \bar{x}) = 0$ , esto es,  $(\bar{x}, \bar{x}) \in G(\Omega)$  pues  $\Omega$  es cerrado. Así,  $\bar{x}$  es punto fijo de  $\Omega$ .  $\square$

**Teorema 4.20** (Kakutani). *Si el conjunto  $C \subset \mathbb{R}^n$  es convexo, compacto y no vacío, entonces todo operador cerrado  $\Omega : C \rightrightarrows C$  que tiene imágenes convexas y no vacías posee un punto fijo.*

PRUEBA. Mostraremos que la multifunción  $\Omega$  satisface las hipótesis del Teorema 4.19. Al ser cada  $\Omega(x)$  la proyección sobre  $C$  del conjunto  $(\{x\} \times C) \cap G(\Omega) \subset C \times C$ , que es cerrado en  $C \times C$ , y por tanto compacto, concluimos que  $\Omega(x)$  es compacto, esto es,  $\Omega$  tiene imágenes compactas.

Ahora solo nos falta notar que como la imagen  $\Omega(C) \subset C$  de  $\Omega$  es acotada y  $C \subset \mathbb{R}^n$  es cerrado, la Proposición 4.14 nos asegura que  $\Omega$  es scs.  $\square$

**Observación 4.21.** De forma similar, usando la Proposición 4.15 podemos verificar que toda multifunción que satisface las hipótesis del Teorema de Kakutani-Fan también satisface las hipótesis del Teorema de Kakutani. Así, los Teoremas 4.19 y 4.20 son dos versiones levemente diferentes del mismo resultado.

<sup>9</sup>Para la demostración de este teorema referimos al lector a [8, Theorem 8.1.3], [24, Teorema 11, pág. 447] o a [14, Theorem (7.2), pág. 95].

Notemos que el teorema del punto fijo de Brouwer es equivalente al teorema de Kakutani (-Fan). De hecho, si  $C$  es un subconjunto convexo, compacto y no vacío de  $\mathbb{R}^n$  entonces toda función continua  $f : C \rightarrow C$  es esencialmente una multifunción scs ( $f$  es continua) de  $C$  en  $C$  con imágenes convexas, compactas y no vacías, y que por tanto tiene un punto fijo, es decir,  $f$  tiene un punto fijo.

## El teorema de von Neumann

En lo que resta de esta sección nos ocupamos de probar el siguiente teorema de von Neumann.

**Teorema 4.22** (Teorema Minimax de von Neumann). *Sean  $X$  e  $Y$  subconjuntos convexos y compactos de  $\mathbb{R}^m$  y  $\mathbb{R}^n$  respectivamente, y sea  $f : X \times Y \rightarrow \mathbb{R}$  una función continua tal que  $x \mapsto f(x, y)$  es cóncava para cada  $y \in Y$  y tal que  $y \mapsto f(x, y)$  es convexa para cada  $x \in X$ . Entonces existe  $(\bar{x}, \bar{y}) \in X \times Y$  tal que*

$$\min_{y \in Y} f(\bar{x}, y) = f(\bar{x}, \bar{y}) = \max_{x \in X} f(x, \bar{y}).$$

En particular,

$$\max_{x \in X} \min_{y \in Y} f(x, y) = \min_{y \in Y} \max_{x \in X} f(x, y).$$

Este teorema, muy importante en la Teoría de Juegos (véase el Ejemplo 9.3), será esencial en el desarrollo del presente trabajo pues nos ayudará a establecer la condición de optimalidad del Teorema 8.4, que relaciona los subdiferenciales con el problema de desigualdad variacional.

Existen varias pruebas del teorema de von Neumann. La que aquí presentamos está basada en [14] y hace uso de los llamados mapeos KKM.

Sea  $X$  un subconjunto de  $\mathbb{R}^n$ . Un operador multivaluado  $\Omega : X \rightrightarrows \mathbb{R}^n$  es llamado **mapeo Knaster-Kuratowski-Mazurkiewicz** o simplemente **mapeo KKM** si

$$\text{co}(A) = \text{co}\{x_1, \dots, x_p\} \subset \Omega(A) = \bigcup_{i=1}^p \Omega(x_i),$$

para todo subconjunto finito  $A = \{x_1, \dots, x_p\}$  de  $X$ . Notemos que si  $\Omega$  es un mapeo KKM entonces tiene imágenes no vacías pues  $x \in \Omega(x)$  para cada  $x$ .

En general no es fácil decidir si una multifunción es un mapeo KKM. Existe, sin embargo, un criterio muy útil.

**Lema 4.23.** *Sea  $X$  un subconjunto convexo y no vacío de  $\mathbb{R}^n$ . Si el operador multivaluado  $\Omega : X \rightrightarrows \mathbb{R}^n$  tiene cofibras  $\Omega^*(y) = X \setminus \Omega^{-1}(y)$  convexas y es tal que  $x \in \Omega(x)$  para todo  $x \in X$ , entonces es un mapeo KKM.*

PRUEBA. Dados  $A = \{x_1, \dots, x_p\} \subset X$  y  $y_0 \in \text{co}(A)$  debemos mostrar que  $y_0$  está en  $\bigcup_{i=1}^p \Omega(x_i)$ . El conjunto  $\text{co}(A)$  no está contenido en  $\Omega^*(y_0)$  ya que  $y_0 \in \Omega(y_0)$ , esto es,  $y_0 \notin \Omega^*(y_0)$ . Luego, como  $\Omega^*(y_0)$  es convexo, existe  $x_i$  que no pertenece a  $\Omega^*(y_0)$  y que por tanto está en  $\Omega^{-1}(y_0)$ , es decir,  $y_0 \in \Omega(x_i)$ .  $\square$

**Ejemplo 4.24.** Sean  $X$  e  $Y$  subconjuntos convexos de  $\mathbb{R}^m$  y  $\mathbb{R}^n$  respectivamente, y sea  $f : X \times Y \rightarrow \mathbb{R}$  una función tal que  $x \mapsto f(x, y)$  es cóncava para cada  $y \in Y$  y  $y \mapsto f(x, y)$  es convexa para cada  $x \in X$ . Entonces el operador multivaluado  $\Omega : X \times Y \rightrightarrows X \times Y$  definido por

$$\Omega(x, y) = \{(x', y') \in \mathbb{R}^m \times \mathbb{R}^n \mid f(x, y') - f(x', y) \leq 0\}$$

es un mapeo KKM. En efecto, veamos que  $\Omega$  satisface las hipótesis del Lema 4.23. Es claro que  $X \times Y$  es convexo y que  $(x, y) \in \Omega(x, y)$  para cualquier  $(x, y) \in X \times Y$ . Además, las cofibras

$$\Omega^*(x', y') = \{(x, y) \in X \times Y \mid f(x', y) - f(x, y') < 0\}$$

de  $\Omega$  son convexas pues cada aplicación  $(x, y) \mapsto f(x', y) - f(x, y')$  es convexa (las aplicaciones  $(x, y) \mapsto f(x', y)$  y  $(x, y) \mapsto -f(x, y')$  son convexas).

El siguiente teorema ilustra una propiedad geométrica importante de los mapeos KKM. Antes de enunciarla, recordemos que una familia  $\{C_\lambda\}_{\lambda \in \Lambda}$  de subconjuntos de  $\mathbb{R}^n$  tiene la propiedad de la intersección finita si  $\bigcap_{\lambda \in \Lambda'} C_\lambda \neq \emptyset$  para cada subconjunto finito  $\Lambda'$  de  $\Lambda$ .

**Teorema 4.25.** Sean  $X$  un subconjunto de  $\mathbb{R}^n$  y  $\Omega : X \rightrightarrows \mathbb{R}^n$  un mapeo KKM con imágenes convexas y cerradas. Entonces la familia  $\{\Omega(x)\}_{x \in X}$  tiene la propiedad de la intersección finita.

PRUEBA. Dado  $A = \{x_1, \dots, x_p\} \subset X$  mostraremos que

$$\text{co}(A) \cap \bigcap_{i=1}^p \Omega(x_i) \neq \emptyset, \quad (4.4)$$

lo que es claramente suficiente para obtener el resultado deseado. La prueba se hará por inducción sobre la cantidad de elementos de  $A$ . Antes de comenzar notemos que al ser  $\Omega$  un mapeo KKM tenemos

$$\text{co}(A) \subset \bigcup_{i=1}^p \Omega(x_i).$$

Si  $A = \{x\}$  entonces (4.4) se cumple pues  $x \in \Omega(x)$  ( $\Omega$  es un mapeo KKM). Supongamos que la afirmación es cierta para todo subconjunto de  $X$  con  $p-1$  elementos, donde  $p \geq 2$ . Podemos entonces escoger para cada  $i = 1, \dots, p$  un vector

$$y_i \in \text{co}(A \setminus \{x_i\}) \cap \bigcap_{j \neq i} \Omega(x_j), \quad (4.5)$$

y considerar el conjunto

$$Y := \text{co}(\{y_1, \dots, y_p\}) \subset \text{co}(A) \subset \bigcup_{i=1}^p \Omega(x_i),$$

que es compacto por el Corolario 2.5. Luego, para probar (4.4) basta con mostrar que

$$Y \cap \bigcap_{i=1}^p \Omega(x_i) \neq \emptyset.$$

Como  $p \geq 2$ , de (4.5) se sigue que los conjuntos  $\Omega(x_j) \cap Y$  son no vacíos. Como para cada  $j$  la función  $\varphi_j : Y \rightarrow \mathbb{R}$  dada por  $\varphi_j(y) = d(y, \Omega(x_j) \cap Y)$  es convexa y continua (cada  $\Omega(x_j) \cap Y$  es convexo y cerrado), la función  $\varphi : Y \rightarrow \mathbb{R}$  definida por  $\varphi(y) = \max\{\varphi_1(y), \dots, \varphi_p(y)\}$  también es convexa y continua. La compacidad de  $Y$  implica en particular que  $\varphi$  alcanza su mínimo en algún punto  $\hat{y} \in Y \subset \bigcup_{i=1}^p \Omega(x_i)$ ; sin pérdida de generalidad suponemos  $\hat{y} \in \Omega(x_p)$ .

Queremos ver que  $\varphi(\hat{y}) = 0$ , lo cual nos da  $\hat{y} \in Y \cap \bigcap_{i=1}^p \Omega(x_i)$ . En efecto, si  $\varphi(\hat{y}) > 0$  podemos evaluar las funciones convexas  $\varphi_i$  en los puntos  $z_t := t\hat{y} + (1-t)y_p$  del conjunto  $[\hat{y}, y_p] \subset Y$  para obtener

$$\varphi_i(z_t) \leq t\varphi_i(\hat{y}) + (1-t)\varphi_i(y_p);$$

como  $\varphi_p(\hat{y}) = 0$  (pues  $\hat{y} \in \Omega(x_p)$ ) y  $\varphi_i(y_p) = 0$  cuando  $i \neq p$  (pues  $y_p \in \Omega(x_i)$ ), lo anterior significa que para cada  $t < 1$  tenemos

$$\varphi_p(z_t) \leq (1-t)\varphi_p(y_p) \quad \text{y} \quad \varphi_i(z_t) \leq t\varphi_i(\hat{y}) < \varphi(\hat{y}) \text{ si } i \neq p,$$

de donde se sigue que  $\varphi(z_t) = \varphi_p(z_t)$  pues  $\varphi(\hat{y}) \leq \varphi(z_t) = \max\{\varphi_1(z_t), \dots, \varphi_p(z_t)\}$ , es decir,  $\varphi(\hat{y}) \leq \varphi(z_t) \leq (1-t)\varphi_p(y_p)$  para todo  $t < 1$ , lo cual contradice  $\varphi(\hat{y}) > 0$ . Así,  $\varphi(\hat{y}) = 0$ , lo que completa la prueba del teorema.  $\square$

**Corolario 4.26** (Principio Geométrico KKM). *Sean  $X$  un subconjunto de  $\mathbb{R}^n$  y  $\Omega : X \rightrightarrows \mathbb{R}^n$  un mapeo KKM con imágenes convexas y cerradas. Si  $\Omega(\bar{x})$  es compacto para algún  $\bar{x} \in X$  entonces la intersección  $\bigcap\{\Omega(x) \mid x \in X\}$  es no vacía.*

PRUEBA. Como la familia  $\{\Omega(x) \cap \Omega(\bar{x})\}_{x \in X}$  de subconjuntos cerrados del conjunto compacto  $\Omega(\bar{x})$  posee la propiedad de la intersección finita, la Proposición 1.13 nos asegura que su intersección  $\bigcap_{x \in X} \Omega(x)$  es no vacía.  $\square$

Ahora ya podemos demostrar el teorema de von Neumann.

PRUEBA DEL TEOREMA DE VON NEUMANN. Afirmamos que existe  $(\bar{x}, \bar{y}) \in X \times Y$  tal que

$$f(x, \bar{y}) \leq f(\bar{x}, y) \text{ para todo } (x, y) \in X \times Y. \quad (4.6)$$

En efecto, observemos primero que  $X \times Y$  es un conjunto convexo y compacto. Por lo visto en el Ejemplo 4.24, la multifunción  $\Omega : X \times Y \rightrightarrows \mathbb{R}^m \times \mathbb{R}^n$  definida por

$$\Omega(x, y) = \{(x', y') \in X \times Y \mid f(x, y') - f(x', y) \leq 0\}$$

es un mapeo KKM, cuyas imágenes  $\Omega(x, y)$  son convexas y compactas. Para ver esto último, basta con notar que las aplicaciones  $(x', y') \mapsto f(x, y') - f(x', y)$  son convexas y que, siendo  $f$  continua,  $\Omega(x, y)$  es cerrado en  $X \times Y$ . Utilizando el Corolario 4.26 deducimos que existe  $(\bar{x}, \bar{y}) \in X \times Y$  tal que  $(\bar{x}, \bar{y}) \in \Omega(x, y)$  para todo  $(x, y) \in X \times Y$ , esto es, tal que (4.6) se verifica. Esto prueba la afirmación.

Ahora vemos fácilmente que

$$\min_{y \in Y} f(\bar{x}, y) = f(\bar{x}, \bar{y}) = \max_{x \in X} f(x, \bar{y}),$$

lo cual implica

$$\max_{x \in X} \min_{y \in Y} f(x, y) = f(\bar{x}, \bar{y}) = \min_{y \in Y} \max_{x \in X} f(x, y),$$

ya que para cada  $x' \in X$  e  $y' \in Y$  tenemos  $\min_{y \in Y} f(x', y) \leq \max_{x \in X} f(x, y')$ .  $\square$

## 5. Análisis cuasiconvexo

Recordemos que los subniveles (estrictos) de toda función convexa son conjuntos convexos. La noción de cuasiconvexidad, que tiene su origen en esta observación, generaliza la noción de convexidad. El interés por estudiar las funciones cuasiconvexas tal vez provenga de la importancia que tienen en economía<sup>10</sup>.

En esta sección realizamos un breve estudio de las funciones cuasiconvexas, haciendo hincapié en la recientemente introducida noción de subniveles ajustados y sus operadores normales [4]. Históricamente, no ha sido fácil estudiar los mínimos de las funciones cuasiconvexas pues muchas de las herramientas que se usan para analizar las funciones convexas ya no se pueden aplicar a las funciones cuasiconvexas (por ejemplo, los subdiferenciales). Es allí donde intervendrán los operadores normales (véase la Subsección 8.3).

### 5.1. Funciones cuasiconvexas

Una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es **cuasiconvexa** si para todo  $x, y \in \mathbb{R}^n$  y  $t \in [0, 1]$  se cumple

$$f(tx + (1-t)y) \leq \max\{f(x), f(y)\}.$$

Decimos que  $f$  es **cuasicóncava** si  $-f$  es cuasiconvexa.

Así como la convexidad del epígrafo caracteriza la convexidad de funciones, la convexidad de los subniveles (estrictos) caracteriza la cuasiconvexidad de funciones.

**Proposición 5.1.** *Dada  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  las siguientes afirmaciones son equivalentes*

- (1)  $f$  es cuasiconvexa;
- (2)  $S_\lambda^<$  es convexo para todo  $\lambda \in \mathbb{R}$ ;
- (3)  $S_\lambda$  es convexo para todo  $\lambda \in \mathbb{R}$ .

**PRUEBA.** Supongamos que  $f$  es cuasiconvexa y sea  $\lambda \in \mathbb{R}$ . Dados  $x, y \in S_\lambda^<$  y  $t \in [0, 1]$  tenemos  $f(x), f(y) < \lambda$ , de modo que  $f(tx + (1-t)y) \leq \max\{f(x), f(y)\} < \lambda$ , es decir,  $tx + (1-t)y \in S_\lambda^<$ . Luego  $S_\lambda^<$  es convexo. Esto muestra que (1) implica (2).

Es fácil ver que (2) implica (3), ya que  $S_\lambda = \bigcap_{\lambda < \mu} S_\mu^<$ , donde cada conjunto  $S_\mu^<$  es convexo por hipótesis.

<sup>10</sup>En Microeconomía, por ejemplo, si las funciones de utilidad son cuasicóncavas entonces las preferencias de los consumidores son convexas.

Supongamos finalmente que se cumple (3). Sean  $x, y \in \mathbb{R}^n$  y  $t \in [0, 1]$ . Para  $\lambda := \max\{f(x), f(y)\} \in \mathbb{R}$  tenemos que  $x$  e  $y$  están en  $S_\lambda$ , que es convexo por hipótesis, de modo que  $tx + (1-t)y \in S_\lambda$ , es decir,  $f(tx + (1-t)y) \leq \max\{f(x), f(y)\}$ . Así,  $f$  es cuasiconvexa y por tanto (3) implica (1).  $\square$

**Observación 5.2.** De la prueba de la proposición anterior deducimos que para que  $f$  sea cuasiconvexa es suficiente que  $S_{f(x)}$  sea convexo para todo  $x \in \mathbb{R}^n$ .

**Ejemplo 5.3.** Toda función convexa es cuasiconvexa. Por ejemplo, las funciones  $f, g : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = x$ ,  $g(x) = x^2$  son cuasiconvexas. La función  $\gamma : \mathbb{R} \rightarrow \mathbb{R}$ ,  $\gamma(x) = \sqrt{|x|}$  es cuasiconvexa pero no convexa.

Una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es denominada **cuasiconvexa semiestricta** si es cuasiconvexa y si para todo  $x, y \in \mathbb{R}^n$  con  $f(x) \neq f(y)$ , y todo  $t \in (0, 1)$  se cumple

$$f(tx + (1-t)y) < \max\{f(x), f(y)\}. \quad (5.1)$$

Notemos que la condición dada por la ecuación anterior no implica la cuasiconvexidad de la función  $f$  pues, por ejemplo, la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  definida mediante  $f(x) = 0$  si  $x \neq 0$  y  $f(0) = 1$  satisface (5.1) pero no es cuasiconvexa.

Intuitivamente, una función cuasiconvexa semiestricta es aquella cuyo gráfico no posee partes planas, excepto posiblemente en  $\arg \min f$ . En otras palabras, una función cuasiconvexa  $f$  es cuasiconvexa semiestricta cuando para todo  $\lambda \in \mathbb{R} \setminus \{\min f\}$ , el conjunto  $f^{-1}(\lambda)$  no posee segmentos  $[x, y]$  no degenerados<sup>11</sup>. Esta propiedad tiene consecuencias interesantes. Por ejemplo, los subniveles usuales y los subniveles estrictos de toda función cuasiconvexa semiestricta tienen la misma clausura (véase la Proposición 5.8).

**Ejemplo 5.4** (Convexidad implica cuasiconvexidad semiestricta). Toda función convexa  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es cuasiconvexa semiestricta. En efecto, dados  $t \in (0, 1)$  y  $x, y \in \mathbb{R}^n$  con  $f(x) \neq f(y)$  tenemos

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) < \max\{f(x), f(y)\}.$$

La función  $\gamma$  del Ejemplo 5.3 es cuasiconvexa semiestricta pero no es convexa.

En resumen,

$$\text{convexidad} \Rightarrow \text{cuasiconvexidad semiestricta} \Rightarrow \text{cuasiconvexidad},$$

donde el recíproco de cada implicancia no vale.

## 5.2. Subniveles ajustados y operadores normales

Los subniveles ajustados de una función y sus correspondientes operadores normales fueron introducidos por D. Aussel y N. Hadjisavvas en [3]. Al igual que los subniveles (estrictos) usuales, la convexidad de los subniveles ajustados caracteriza la cuasiconvexidad de la función  $f$  en cuestión. A diferencia de los operadores normales usuales (esto es, los operadores normales con respecto a los subniveles usuales),

<sup>11</sup>Un segmento  $[x, y]$  es no degenerado si  $x \neq y$ .

los operadores normales obtenidos a partir de los subniveles ajustados tienen buenas propiedades (cuasimonotonía, semicontinuidad superior, etc.) incluso cuando la función cuasiconvexa  $f$  considerada tiene “partes planas”<sup>12</sup>. Esta es una característica muy importante pues en artículos previos (véase por ejemplo [7]) se habían estudiado los operadores normales usuales y cómo estos se relacionaban con la cuasiconvexidad de  $f$ , obteniéndose que ellos tenían buenas propiedades solo cuando  $f$  no tiene partes planas. Así, a diferencia de los subniveles usuales, los subniveles ajustados permiten estudiar todos los tipos de funciones cuasiconvexas.

En esta subsección definimos y damos las principales propiedades de los subniveles ajustados y sus operadores normales. Trabajaremos en  $\mathbb{R}^n$ , si bien estas nociones fueron dadas en el contexto más general de espacios de Banach. Solo probaremos aquello que necesitaremos, y remitimos al lector interesado en mayor información al artículo [3].

Sean la función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  y el vector  $x \in \mathbb{R}^n$ . El **subnivel ajustado** de  $f$  en  $x$  es el conjunto dado por

$$S_f^a(x) = \begin{cases} S_{f(x)} \cap \overline{B(S_{f(x)}^<, \rho_x)}, & \text{si } x \notin \arg \min f, \\ S_{f(x)}, & \text{si } x \in \arg \min f, \end{cases}$$

donde  $\rho_x = d(x, S_{f(x)}^<)$  y

$$B(S_{f(x)}^<, \rho_x) = S_{f(x)}^< + \rho_x B(0, 1).$$

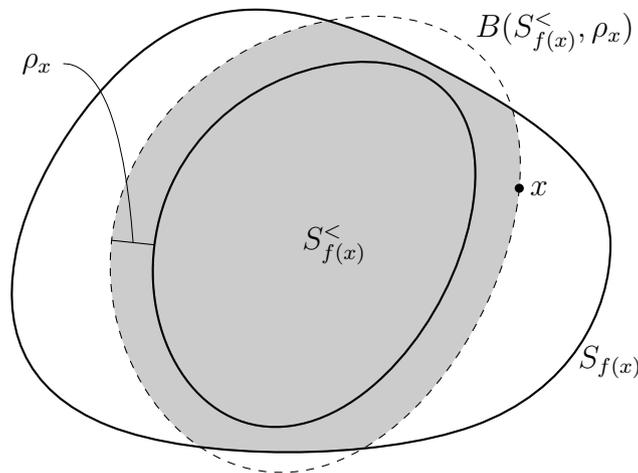


Figura 5.1: Ilustración gráfica de los subniveles ajustados. El punto  $x$  no está en  $\arg \min f$ . El conjunto del centro del gráfico es  $S_{f(x)}^<$ , el conjunto de fuera es  $S_{f(x)}$ , y el conjunto punteado es  $B(S_{f(x)}^<, \rho_x)$ . La región sombreada representa a  $S_f^a(x)$ .

<sup>12</sup>Una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  tiene “partes planas” si existe  $\lambda > \inf_{x \in \mathbb{R}^n} f(x)$  tal que  $S_\lambda \setminus \overline{S_\lambda^<} \neq \emptyset$ . Por ejemplo, las funciones cuasiconvexas semiestrictas no tienen partes planas (Proposición 5.8).

**Observación 5.5.** Si  $x \notin \arg \min f$  entonces

$$\overline{B(S_{f(x)}^<, \rho_x)} = \{y \in \mathbb{R}^n \mid d(y, S_{f(x)}^<) \leq \rho_x\}, \quad (5.2)$$

y por tanto

$$x \in S_f^a(x), \quad \forall x \in \mathbb{R}^n.$$

En efecto, (5.2) claramente se cumple si  $\rho_x = 0$ . Cuando  $\rho_x > 0$  tenemos

$$B(S_{f(x)}^<, \rho_x) = \{y \in \mathbb{R}^n \mid d(y, S_{f(x)}^<) < \rho_x\},$$

y la continuidad de  $d(\cdot, S_{f(x)}^<)$  implica que  $\{y \in \mathbb{R}^n \mid d(y, S_{f(x)}^<) \leq \rho_x\}$  contiene a  $\overline{B(S_{f(x)}^<, \rho_x)}$ . Recíprocamente, si  $y \in \mathbb{R}^n$  es tal que  $d(y, S_{f(x)}^<) \leq \rho_x$ , tomando una sucesión  $(a_k)$  de números positivos con  $a_k \rightarrow 0$  vemos que existe una sucesión  $(x_k)$  en  $S_{f(x)}^<$  tal que  $\|y - x_k\| \leq \rho_x + a_k$  para todo  $k$ . Suponiendo sin pérdida de generalidad que  $\rho_x > a_k$  para cada  $k$ , podemos escoger  $y_k \in [y, x_k]$  tal que

$$\|y_k - x_k\| = \rho_x - a_k < \rho_x \quad \text{y} \quad \|y_k - y\| \leq a_k,$$

de donde  $y \in \overline{B(S_{f(x)}^<, \rho_x)}$  pues  $(y_k)$  es una sucesión en  $B(S_{f(x)}^<, \rho_x)$  que converge a  $y$ . Esto prueba (5.2).

Claramente, el subnivel estricto siempre está contenido en el subnivel ajustado, y este a su vez está contenido en el subnivel usual, esto es,

$$S_{f(x)}^< \subset S_f^a(x) \subset S_{f(x)}$$

para todo  $x \in \mathbb{R}^n$ . Veremos que los subniveles ajustados, al igual que los subniveles que ya conocemos, caracterizan la cuasiconvexidad de una función.

**Proposición 5.6.** Una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es cuasiconvexa si y solo si todos sus subniveles ajustados son convexos.

PRUEBA. Si  $f$  es cuasiconvexa, los conjuntos  $S_\lambda$  y  $S_\lambda^<$  son convexos para todo  $\lambda \in \mathbb{R}$ . En particular, dado  $x \in \mathbb{R}^n$ ,  $S_{f(x)}$  y  $S_{f(x)}^<$  son convexos; luego  $B(S_{f(x)}^<, \rho_x)$  también es convexo, y concluimos que  $S_f^a(x)$  es convexo.

Recíprocamente, supongamos que todos los subniveles ajustados de  $f$  son convexos, y probemos que  $S_{f(x)}$  es convexo para cada  $x \in \mathbb{R}^n$ , lo que implicará que  $f$  es cuasiconvexa (Observación 5.2). Si  $x \in \arg \min f$  entonces  $S_{f(x)} = S_f^a(x)$  es convexo por hipótesis. Supongamos ahora que  $x \notin \arg \min f$ . Sean  $y, z \in S_{f(x)}$ , y veamos que  $[y, z] \subset S_{f(x)}$ . Sabemos que  $S_{f(x)} \cap \overline{B(S_{f(x)}^<, \rho_x)} = S_f^a(x)$  y que  $f(y), f(z) \leq f(x)$ . Consideraremos tres casos. Si  $y, z \in \overline{B(S_{f(x)}^<, \rho_x)}$ , entonces  $y, z \in S_f^a(x)$  y por la convexidad de este conjunto,  $[y, z] \subset S_f^a(x) \subset S_{f(x)}$ .

Si  $y \notin \overline{B(S_{f(x)}^<, \rho_x)}$  y  $z \in \overline{B(S_{f(x)}^<, \rho_x)}$  entonces  $y \notin S_{f(x)}^<$  y por tanto  $f(x) = f(y)$ ; luego  $S_{f(x)}^< = S_{f(y)}^<$ ,  $\rho_y > \rho_x$  y  $z \in \overline{B(S_{f(y)}^<, \rho_x)} \subset \overline{B(S_{f(y)}^<, \rho_y)}$ , de donde obtenemos que  $z \in S_f^a(y)$ . Como el conjunto convexo  $S_f^a(y)$  contiene a  $z$  e  $y$  se sigue que  $[y, z] \subset S_f^a(y) \subset S_{f(y)} = S_{f(x)}$ .

Si  $y, z \notin \overline{B(S_{f(x)}^<, \rho_x)}$ , como en el caso anterior deducimos  $f(x) = f(y) = f(z)$ ; luego,  $S_{f(x)}^< = S_{f(y)}^< = S_{f(z)}^<$  y  $\rho_y, \rho_z > \rho_x$ . Suponiendo sin pérdida de generalidad que  $\rho_y \geq \rho_z$  tenemos que  $z \in S_f^a(z) \subset S_f^a(y)$ , y como  $y \in S_f^a(y)$ , la convexidad de  $S_f^a(y)$  implica  $[y, z] \subset S_f^a(y) \subset S_{f(y)} = S_{f(x)}$ .  $\square$

Por la Observación 5.5 sabemos que  $x \in S_f^a(x)$  para todo  $x$ . Más aun, puede mostrarse que  $x \in \partial S_f^a(x)$  cuando  $x \notin \arg \min f$ .

**Proposición 5.7.** *Supongamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función cuasiconvexa y consideremos  $x \in \mathbb{R}^n \setminus \arg \min f$ . Entonces  $x \in \partial S_f^a(x)$ .*

Este resultado nos será útil luego para ver que el operador normal asociado a los subniveles ajustados tiene imágenes no nulas en cada  $x \notin \arg \min f$  (Proposición 5.10).

PRUEBA. Mostramos primero un resultado preliminar: si  $A$  y  $B$  son subconjuntos de  $\mathbb{R}^n$  y  $z \in A \cap B$  es tal que  $z \in \partial B$ , entonces  $z \in \partial(A \cap B)$ . En efecto, basta con notar que como  $z \in \partial B$ , toda bola  $B(z, \varepsilon)$  con centro en  $z$  contiene algún punto que no está en  $B$ , y que por consiguiente tampoco está en  $A \cap B$ .

Definamos  $A := S_{f(x)}$  y  $B := \overline{B(S_{f(x)}^<, \rho_x)}$ . Aplicaremos nuestro resultado preliminar para ver que  $x \in \partial(A \cap B) = \partial S_f^a(x)$ . Como  $x \in B$  debido a la Observación 5.5, o sea  $x \in A \cap B$ , tan solo nos falta mostrar que  $x \in \partial B$ .

Tenemos que  $x \notin B(S_{f(x)}^<, \rho_x)$ , ya que si  $x \in B(S_{f(x)}^<, \rho_x)$  entonces  $d(x, S_{f(x)}^<) < \rho_x$  (cuando  $\rho_x > 0$ ) o  $x \in S_{f(x)}^<$  (cuando  $\rho_x = 0$ ), una contradicción (en ambos casos). Por otro lado, siendo  $B(S_{f(x)}^<, \rho_x)$  convexo tenemos por la Proposición 2.9 que

$$\text{int}(B) = \text{int}(\overline{B(S_{f(x)}^<, \rho_x)}) = \text{int}(B(S_{f(x)}^<, \rho_x)) \subset B(S_{f(x)}^<, \rho_x),$$

y en consecuencia  $x \notin \text{int}(B)$ , es decir,  $x \in \partial B$ . □

Antes de pasar a estudiar los operadores normales asociados a los subniveles ajustados, mostraremos algunas propiedades interesantes de las funciones cuasiconvexas semiestrictas. Por ejemplo, sus subniveles ajustados no son otra cosa que sus subniveles usuales.

**Proposición 5.8.** *Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función cuasiconvexa semiestricta. Entonces:*

- (1) *Para todo  $x \in \mathbb{R}^n \setminus \arg \min f$  se cumple que  $\overline{S_{f(x)}^<} = \overline{S_{f(x)}}$ .*
- (2) *Para todo  $x \in \mathbb{R}^n$  se cumple que  $S_f^a(x) = S_{f(x)}$ .*
- (3) *Si  $f$  es continua, para todo  $x \in \mathbb{R}^n$ , el conjunto  $S_f^a(x)$  es cerrado; si además  $x \notin \arg \min f$  también se cumple que  $S_f^a(x) = \overline{S_{f(x)}^<}$  y que  $\text{int}(S_f^a(x)) = S_{f(x)}^<$ .*

PRUEBA. (1) Como la inclusión  $\overline{S_{f(x)}^<} \subset \overline{S_{f(x)}}$  es clara bastará con mostrar que dados  $z \in S_{f(x)}$  y  $\varepsilon > 0$ , el conjunto  $B(z, \varepsilon) \cap S_{f(x)}^<$  es no vacío. Sabemos que existe  $y \in S_{f(x)}^<$ , ya que  $x \notin \arg \min f$ . Obviamente podemos suponer que  $z \neq y$ . Como  $f(y) < f(x)$  y  $f(z) \leq f(x)$ , si  $w \in (z, y)$  entonces la cuasiconvexidad de  $f$  implica que

$$f(w) \leq \max\{f(z), f(y)\} = f(y) < f(x) \quad \text{cuando } f(z) = f(y),$$

y la cuasiconvexidad semiestricta de  $f$  implica que

$$f(w) < \max\{f(z), f(y)\} \leq f(x) \quad \text{cuando } f(z) \neq f(y).$$

Luego  $(z, y] \subset S_{f(x)}^<$ , lo cual prueba que  $B(z, \varepsilon) \cap S_{f(x)}^<$  es no vacío.

(2) Podemos suponer que  $x \notin \arg \min f$ . De la parte (1) vemos que  $x$  está en  $\overline{S_{f(x)}^<}$ , lo cual significa  $\rho_x = d(x, S_{f(x)}^<) = 0$ . Tenemos así que  $\overline{B(S_{f(x)}^<, \rho_x)} = \overline{S_{f(x)}^<} = \overline{S_{f(x)}}$ , y por consiguiente  $S_f^a(x) = S_{f(x)} \cap \overline{S_{f(x)}^<} = S_{f(x)}$ .

(3) Utilizamos los ítems (1) y (2). Fijemos  $x \in \mathbb{R}^n$ . Al ser  $f$  continua, el conjunto  $S_f^a(x) = S_{f(x)}$  es cerrado. Si  $x \notin \arg \min f$ , entonces  $S_f^a(x) = S_{f(x)} = \overline{S_{f(x)}^<} = \overline{S_{f(x)}^<}$ ; como además  $S_{f(x)}^<$  es abierto, convexo y no vacío ( $f$  es cuasiconvexa y continua), del Corolario 2.8 concluimos

$$\text{int}(S_f^a(x)) = \text{int}(\overline{S_{f(x)}^<}) = \text{int}(S_{f(x)}^<) = S_{f(x)}^<. \quad \square$$

Volvamos a los operadores normales. Dada una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ , el **operador normal** asociado a  $f$  es el operador multivaluado  $N_f^a : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  que asocia a cada  $x \in \mathbb{R}^n$  con el cono normal de  $S_f^a(x)$  en  $x$ , es decir,

$$N_f^a(x) = N(x, S_f^a(x)) = \{v \in \mathbb{R}^n \mid \langle v, y - x \rangle \leq 0, \forall y \in S_f^a(x)\}.$$

**Observación 5.9.** De igual modo podemos definir operadores normales teniendo en consideración los conjuntos  $S_{f(x)}^<$  y  $S_{f(x)}$ . Concretamente,

$$\begin{aligned} N_f^<(x) &= N(x, S_{f(x)}^<) = \{v \in \mathbb{R}^n \mid \langle v, y - x \rangle \leq 0, \forall y \in S_{f(x)}^<\}, \\ N_f(x) &= N(x, S_{f(x)}) = \{v \in \mathbb{R}^n \mid \langle v, y - x \rangle \leq 0, \forall y \in S_{f(x)}\}. \end{aligned}$$

Debido a que  $S_{f(x)}^< \subset S_f^a(x) \subset S_{f(x)}$ , siempre se cumple que

$$N_f(x) \subset N_f^a(x) \subset N_f^<(x).$$

Cuando  $f$  es cuasiconvexa semiestricta, la Proposición 5.8 nos garantiza que  $S_{f(x)}$  y  $S_f^a(x)$  siempre son iguales, y que  $\overline{S_{f(x)}^<} = \overline{S_f^a(x)}$  cuando  $x \notin \arg \min f$ ; luego,

$$\begin{aligned} N_f(x) &= N_f^a(x), & \text{si } x \in \mathbb{R}^n \\ N_f(x) &= N_f^a(x) = N_f^<(x), & \text{si } x \in \mathbb{R}^n \setminus \arg \min f. \end{aligned}$$

**Proposición 5.10.** *Supongamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es cuasiconvexa y consideremos  $x \in \mathbb{R}^n \setminus \arg \min f$ . Entonces el conjunto  $N_f^a(x) \setminus \{0\}$  es no vacío.*

PRUEBA. Como  $x$  pertenece a la frontera del conjunto convexo  $S_f^a(x)$ , por el teorema de separación (Teorema 2.12) existe un vector no nulo  $v \in \mathbb{R}^n$  tal que  $\langle v, x \rangle \geq \langle v, y \rangle$  para todo  $y \in S_f^a(x)$ , es decir,  $v \in N_f^a(x) \setminus \{0\}$ .  $\square$

La Proposición 5.10 obviamente se satisface también para el operador normal  $N_f^<$ . No obstante, ella no se cumple para el operador  $N_f$  pues existen ejemplos simples donde  $x \in \text{int}(S_{f(x)})$ , en cuyo caso  $N_f(x) = \{0\}$ .

## 6. Funciones definidas a partir de supremos

El propósito de esta sección es probar dos resultados técnicos que necesitaremos en la Sección 15, a saber, un teorema de sensibilidad estándar (Teorema 6.6) y el teorema de Danskin (Teorema 6.7). La demostración del Teorema 6.6 está basada en el artículo [20], mientras que la prueba del teorema de Danskin ha sido obtenida de [13, Theorem 10.2.1]. Utilizaremos la terminología introducida en la Sección 4.

Sea  $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\} = [-\infty, +\infty]$ .

Decimos que una función  $f : X \rightarrow \overline{\mathbb{R}}$  definida sobre un subconjunto  $X$  de  $\mathbb{R}^n$  es **semicontinua inferior (sci)** en  $\bar{x} \in \mathbb{R}^n$  si para toda sucesión  $(x_k)$  en  $X$  que converge a  $\bar{x}$  se cumple que  $f(\bar{x}) \leq \liminf_k f(x_k)$ . Cuando  $f$  es sci en todos los puntos de  $X$  decimos simplemente que  $f$  es **sci**. De modo análogo,  $f$  es **semicontinua superior (scs)** en  $\bar{x} \in X$  si  $-f$  es sci en  $\bar{x}$ ; y  $f$  es **scs** si  $-f$  es sci.

**Ejemplo 6.1.** Toda función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  continua (en el sentido usual) es sci y scs. La función  $g : \mathbb{R} \rightarrow \mathbb{R}$  dada por  $g(x) = 0$  si  $x \neq \pm 1$  y  $g(x) = x$  si  $x = \pm 1$  no es sci en  $x = 1$  ni tampoco es scs en  $x = -1$ , pero sí es sci y scs en  $\mathbb{R} \setminus \{\pm 1\}$ .

Sean el operador  $\Omega : X \rightrightarrows Y$  y la función  $f : X \times Y \rightarrow \mathbb{R}$ , donde  $X \subset \mathbb{R}^n$  e  $Y \subset \mathbb{R}^m$ . Para cada  $x \in X$  definimos

$$v(x) := \sup_{y \in \Omega(x)} f(x, y),$$

y el conjunto

$$M(x) := \{y \in \Omega(x) \mid v(x) = f(x, y)\}.$$

Obtenemos así una función

$$v : X \rightarrow \overline{\mathbb{R}}$$

y una multifunción

$$M : X \rightrightarrows Y.$$

Queremos ver cómo se relacionan las propiedades de  $\Omega$  y  $f$  con las de  $v$  y  $M$ .

**Proposición 6.2.** Si  $\Omega$  es sci en  $\bar{x} \in X$  y  $f$  es sci en  $\bar{x} \times \Omega(\bar{x})$ , entonces  $v$  es sci en  $\bar{x}$ .

PRUEBA. Si  $\Omega(\bar{x}) = \emptyset$  no hay nada que probar pues  $v(\bar{x}) = -\infty$ . Si  $\Omega(\bar{x}) \neq \emptyset$ , consideramos dos casos.

Si  $v(\bar{x}) < \infty$ , tomemos una sucesión  $x_k \rightarrow \bar{x}$  y fijemos  $\varepsilon > 0$ . Sea  $\bar{y} \in \Omega(\bar{x})$  tal que  $v(\bar{x}) \leq f(\bar{x}, \bar{y}) + \varepsilon$ . Como  $\Omega$  es sci en  $\bar{x}$ , existe  $(y_k) \subset Y$  tal que  $y_k \in \Omega(x_k)$  para  $k$  suficientemente grande y tal que  $y_k \rightarrow \bar{y}$ . Luego, puesto que  $f$  es sci en  $\bar{x} \times \Omega(\bar{x})$ ,

$$v(\bar{x}) - \varepsilon \leq f(\bar{x}, \bar{y}) \leq \liminf f(x_k, y_k) \leq \liminf v(x_k),$$

esto es,  $v(\bar{x}) \leq \liminf v(x_k)$ , pues  $\varepsilon$  fue arbitrario.

Si  $v(\bar{x}) = +\infty$ , existe una sucesión  $(\bar{y}^j)_j \subset \Omega(\bar{x})$  tal que  $f(\bar{x}, \bar{y}^j) \rightarrow v(\bar{x})$ . Tomemos  $x_k \rightarrow \bar{x}$ . Como  $\Omega$  es sci en  $\bar{x}$ , para cada  $j$  existe una sucesión  $(y_k^j)_k$  con  $y_k^j \in \Omega(x_k)$  y  $\lim_k y_k^j = \bar{y}^j$ . Así, de la semicontinuidad inferior de  $f$  en  $\bar{x} \times \Omega(\bar{x})$  se sigue que  $f(\bar{x}, \bar{y}^j) \leq \liminf_k f(x_k, y_k^j) \leq \liminf_k v(x_k)$  para cada  $j$  y por tanto

$$v(\bar{x}) = +\infty = \lim f(\bar{x}, \bar{y}^j) \leq \liminf v(x_k).$$

Como las sucesiones  $(x_k)$  fueron arbitrarias, lo anterior prueba que  $v$  es sci en  $\bar{x}$ .  $\square$

**Proposición 6.3.** *Si  $\Omega$  es sci y cerrado en  $\bar{x} \in X$ , y si  $f$  es continua en  $\bar{x} \times \Omega(\bar{x})$ , entonces  $M$  es cerrado en  $\bar{x}$ .*

PRUEBA. Supongamos que  $x_k \rightarrow \bar{x}$ ,  $y_k \in M(x_k)$  y  $y_k \rightarrow \bar{y}$ . Entonces  $f(x_k, y_k) = v(x_k)$  para cada  $k$  y  $\bar{y} \in \Omega(\bar{x})$  (pues  $\Omega$  es cerrado en  $\bar{x}$ ), de modo que  $f(\bar{x}, \bar{y}) \leq v(\bar{x})$ . Luego, por la Proposición 6.2,  $v$  es sci en  $\bar{x}$  y así,

$$f(\bar{x}, \bar{y}) = \lim f(x_k, y_k) = \lim v(x_k) \geq v(\bar{x}),$$

de donde se sigue que  $\bar{y} \in M(\bar{x})$ . □

**Corolario 6.4.** *Supongamos que  $\Omega$  es sci y cerrado en  $\bar{x} \in X$ , y que  $f$  es continua en  $\bar{x} \times \Omega(\bar{x})$ . Supongamos además que  $M$  es no vacío y acotado cerca de  $\bar{x}$ , y que  $M(\bar{x})$  es unitario. Entonces  $M$  es cerrado y sci en  $\bar{x}$ .*

PRUEBA. Sabemos que  $M$  es cerrado en  $\bar{x}$  (Proposición 6.3) y para ver que es sci usamos el Lema 4.8. Sean  $x_k \rightarrow \bar{x}$  y  $M(\bar{x}) = \bar{y}$ . Tomemos  $(y_k) \subset Y$  tal que  $y_k \in M(x_k)$  para  $k$  suficientemente grande (puede hacerse ya que  $M$  es no vacío cerca de  $\bar{x}$ ) y supongamos que  $y_k \not\rightarrow \bar{y}$ , esto es, que existe un abierto  $B$  que contiene a  $\bar{y}$  tal que  $y_k \notin B$  para infinitos  $k$ . Esta subsucesión estaría acotada ( $M$  es acotada cerca de  $\bar{x}$ ) y tendría una subsucesión convergente a algún  $\hat{y} \notin B$ . Luego  $\hat{y} \in M(\bar{x}) = \{\bar{y}\}$  pues  $M$  es cerrado en  $\bar{x}$ , una contradicción dado que  $\bar{y} \in B$ . Se sigue que  $y_k \rightarrow \bar{y}$  y concluimos que  $M$  es sci en  $\bar{x}$ . □

**Teorema 6.5.** *Supongamos que  $f$  es continua y que  $f(x, \cdot)$  es cuasicóncava para cada  $x \in X$  fijo. Supongamos también que  $\Omega$  es cerrado en una vecindad de  $\bar{x} \in X$  y sci en  $\bar{x}$ , y que  $\Omega(x)$  es convexo para todo  $x$  en algún entorno de  $\bar{x}$ . Entonces, si  $M(\bar{x})$  es acotado y no vacío, se cumple que  $M$  es acotado y no vacío cerca de  $\bar{x}$ .*

PRUEBA. Si la conclusión del teorema fuera falsa, habría una sucesión  $x_k \rightarrow \bar{x}$  tal que o bien  $M(x_k)$  es vacío, o bien  $\bigcup_{j>k} M(x_j)$  es no acotado, o ambas condiciones se cumplen, para una cantidad infinita de  $k$ . Sea  $\bar{y} \in M(\bar{x})$ . Como  $\Omega$  es sci en  $\bar{x}$ , existe una sucesión  $y_k^1 \in \Omega(x_k)$  tal que  $y_k^1 \rightarrow \bar{y}$ .

Supongamos que  $M(x_k) = \emptyset$ . Si  $v(x_k) < \infty$ , como  $f$  es continua y  $\Omega(x_k)$  es cerrado, existe  $(z_j) \subset \Omega(x_k)$  con  $\|z_j\| \rightarrow \infty$  tal que  $v(x_k) \leq f(x_k, z_j) + 1/j$  para cada  $j$ . Si  $v(x_k) = \infty$ , nuevamente por la continuidad de  $f$  y por el hecho de que  $\Omega(x_k)$  es cerrado, podemos tomar  $(z_j) \subset \Omega(x_k)$  con  $\|z_j\| \rightarrow \infty$  y tal que  $f(x_k, y_k^1) \leq f(x_k, z_j)$  para todo  $j$ . En cualquier caso, escogemos  $y_k := z_j$  con  $\|z_j\| > k$ .

Si  $\bigcup_{j>k} M(x_j)$  fuera no acotado para infinitos  $k$ , habría subsucesiones  $(x_j)$  e  $(y_j)$  tales que  $y_j \in M(x_j)$  y  $\|y_j\| \rightarrow \infty$ .

De esta forma, si la conclusión del teorema fuera falsa, habría una sucesión  $y_k \in \Omega(x_k)$  con  $\|y_k\| \rightarrow \infty$  y tal que

$$f(x_k, y_k^1) - 1/k \leq f(x_k, y_k).$$

Definamos la sucesión  $y_k^2 = \lambda_k y_k^1 + (1 - \lambda_k) y_k$  (donde  $\lambda_k \in [0, 1]$  son números por determinar). Entonces  $y_k^2 \in \Omega(x_k)$  para todo  $k$  pues estos conjuntos son convexos. De la cuasiconcavidad de  $f(x_k, \cdot)$  obtenemos

$$f(x_k, y_k^2) \geq \min\{f(x_k, y_k^1), f(x_k, y_k)\} \geq f(x_k, y_k^1) - 1/k.$$

Luego, todo punto de acumulación  $\hat{y}$  de  $(y_k^2)_k$  pertenecerá a  $\Omega(\bar{x})$  (pues  $\Omega$  es cerrado en  $\bar{x}$ ) y satisfará  $f(\bar{x}, \hat{y}) \geq f(\bar{x}, \bar{y}) = v(\bar{x})$  (pues  $f$  es continua), es decir,  $\hat{y} \in M(\bar{x})$ . Afirmamos que  $(y_k^2)_k$  tiene puntos de acumulación arbitrariamente grandes (para  $\lambda_k$  adecuados), lo que nos daría la contradicción deseada ya que  $M(\bar{x})$  es acotado.

Dado un número  $N > \|\bar{y}\|$  supongamos que ningún punto del conjunto compacto  $K := \{x \in \mathbb{R}^n \mid N \leq \|x\| \leq N+1\}$  puede expresarse como punto de acumulación de una sucesión del tipo  $(y_k^2)$ . Entonces podríamos cubrir  $K$  con una familia de abiertos tal que cada uno de ellos interseca a una cantidad finita de las rectas  $[y_k^1, y_k]$ , de donde obtendríamos un subcubrimiento finito, lo que implicaría que solo una cantidad finita de rectas  $[y_k^1, y_k]$  intersecan  $K$ , absurdo pues  $\|y_k\| \rightarrow \infty$  y  $y_k^1 \rightarrow \bar{y}$ . Así, las sucesiones del tipo  $(y_k^2)$  tienen puntos de acumulación arbitrariamente grandes.  $\square$

Como resultado del Teorema 6.5 y del Corolario 6.4, obtenemos el siguiente teorema de sensibilidad estándar.

**Teorema 6.6.** *Supongamos que  $f$  es continua y que  $f(x, \cdot)$  es cóncava para cada  $x \in X$  fijo. Supongamos también que  $\Omega$  es cerrado en una vecindad de  $\bar{x} \in \text{dom}(\Omega)$  y sci en  $\bar{x}$ , y que  $\Omega(x)$  es convexo para todo  $x$  en algún entorno de  $\bar{x}$ . Entonces, si  $M$  es univaluado en  $\text{dom}(\Omega)$  (esto es, si  $M$  es una función en  $\text{dom}(\Omega)$ ), se cumple que  $M$  es continua (como función) en  $\bar{x}$ .*

PRUEBA. Como  $f$ ,  $\Omega$  y  $M$  satisfacen las hipótesis del Teorema 6.5,  $M$  es acotado y no vacío cerca de  $\bar{x}$ . Luego,  $f$ ,  $\Omega$  y  $M$  satisfacen las hipótesis del Corolario 6.4, y por eso  $M$  es cerrado y sci en  $\bar{x}$ . Todo esto significa que existe  $U \subset X$  abierto que contiene a  $\bar{x}$  y tal que la restricción de  $M$  a  $U$  es una multifunción cerrada y sci en  $\bar{x}$ , y con valores unitarios. Luego, el hecho de que  $M$  sea sci en  $\bar{x}$  como multifunción implica que es continua en  $\bar{x}$  como función (Observación 4.12).  $\square$

Para terminar esta sección mostraremos otro resultado que relaciona la diferenciabilidad de  $f$  con la de  $v$  bajo la hipótesis de que  $\Omega$  tiene una única imagen convexa, cerrada y no vacía  $K$ , o sea  $\Omega(x) = K$  para todo  $x$ .

**Teorema 6.7** (Danskin). *Sea  $K \subset \mathbb{R}^m$  un conjunto convexo, cerrado y no vacío, y sea  $U \subset \mathbb{R}^n$  un conjunto abierto y no vacío. Asumamos que la función  $f : U \times K \rightarrow \mathbb{R}$  es continua en  $U \times K$  y que  $\nabla_x f(x, y)$  existe y es continua en  $U \times K$ . Para la función  $v : U \rightarrow \bar{\mathbb{R}}$  dada por*

$$v(x) = \sup_{y \in K} f(x, y)$$

se define

$$M(x) = \{y \in K \mid v(x) = f(x, y)\}.$$

*Supongamos que  $M$  es acotado y no vacío cerca de  $\bar{x} \in U$ , esto es, existe una vecindad  $V \subset U$  de  $\bar{x}$  tal que  $M(x)$  es no vacío para todo  $x \in V$  y el conjunto*

$$\bigcup_{x \in U} M(x) \tag{6.1}$$

*es acotado. Entonces se cumplen las siguientes afirmaciones:*

(1) Existen las derivadas direccionales de  $v$  en  $\bar{x}$  y están dadas por

$$\frac{\partial v}{\partial d}(\bar{x}) = \sup_{y \in M(\bar{x})} \langle \nabla_x f(\bar{x}, y), d \rangle.$$

(2) Si  $M(\bar{x})$  es unitario, digamos  $M(\bar{x}) = \{y(\bar{x})\}$ , entonces  $v$  es diferenciable en  $\bar{x}$  con

$$\nabla v(\bar{x}) = \nabla_x f(\bar{x}, y(\bar{x})).$$

PRUEBA. Sea  $d \in \mathbb{R}^n$  y escojamos  $\bar{t} > 0$  tal que  $\bar{x} + td \in V$  para todo  $t \in (0, \bar{t}]$ . Para cada  $y \in M(\bar{x})$  tenemos por la definición de  $v$  que

$$v(\bar{x} + td) - v(\bar{x}) \geq f(\bar{x} + td, y) - f(\bar{x}, y),$$

de donde

$$\liminf_{t \downarrow 0} \frac{v(\bar{x} + td) - v(\bar{x})}{t} \geq \liminf_{t \downarrow 0} \frac{f(\bar{x} + td, y) - f(\bar{x}, y)}{t} = \langle \nabla_x f(\bar{x}, y), d \rangle$$

y por tanto

$$\liminf_{t \downarrow 0} \frac{v(\bar{x} + td) - v(\bar{x})}{t} \geq \sup_{y \in M(\bar{x})} \langle \nabla_x f(\bar{x}, y), d \rangle. \quad (6.2)$$

Análogamente, dado  $y_t \in M(\bar{x} + td)$  tenemos

$$v(\bar{x} + td) - v(\bar{x}) \leq f(\bar{x} + td, y_t) - f(\bar{x}, y_t) = t \cdot \langle \nabla_x f(\bar{x}_t, y_t), d \rangle$$

para algún  $\bar{x}_t \in (\bar{x}, \bar{x} + td)$  por el teorema del valor medio, de forma que

$$\frac{v(\bar{x} + td) - v(\bar{x})}{t} \leq \langle \nabla_x f(\bar{x}_t, y_t), d \rangle, \quad \forall y_t \in M(\bar{x} + td). \quad (6.3)$$

Ahora bien, si  $(t_k)$  es una sucesión de números positivos que converge a 0 y  $y^k \in M(\bar{x} + t_k d)$  para cada  $k$ , es decir,  $f(\bar{x} + t_k d, y^k) \geq f(\bar{x} + t_k d, y)$  para todo  $y \in K$ , entonces  $(y^k)$  es una sucesión acotada por la condición de acotación (6.1), y de la continuidad de  $f$  se sigue que

$$f(\bar{x}, \bar{y}) \geq f(\bar{x}, y), \quad \forall y \in K$$

para cualquier punto de acumulación  $\bar{y}$  de  $(y^k)$ , esto es,  $\bar{y} \in M(\bar{x})$ . Así pues, si  $t_k \rightarrow 0$  es tal que  $\frac{v(\bar{x} + t_k d) - v(\bar{x})}{t_k}$  converge a  $\limsup_{t \downarrow 0} \frac{v(\bar{x} + td) - v(\bar{x})}{t}$  podemos suponer que la sucesión  $(y_{t_k})$  converge a algún punto  $\bar{y} \in M(\bar{x})$  y la continuidad de  $\nabla_x f(x, y)$  implica  $\nabla_x f(\bar{x}_{t_k}, y_{t_k}) \rightarrow \nabla_x f(\bar{x}, \bar{y})$ , de donde aplicando (6.3) conseguimos

$$\limsup_{t \downarrow 0} \frac{v(\bar{x} + td) - v(\bar{x})}{t} \leq \sup_{y \in M(\bar{x})} \langle \nabla_x f(\bar{x}, y), d \rangle.$$

Combinando esto con (6.2) obtenemos la parte (1) del teorema. Para ver la parte (2) basta con notar que si  $M(\bar{x}) = \{y(\bar{x})\}$  entonces  $\frac{\partial v}{\partial d}(\bar{x}) = \langle \nabla_x f(\bar{x}, y(\bar{x})), d \rangle$  para cada  $d \in \mathbb{R}^n$ , lo cual significa  $\nabla v(\bar{x}) = \nabla_x f(\bar{x}, y(\bar{x}))$ .  $\square$

## 7. Desigualdades variacionales

Uno de los objetivos de esta tesis es presentar reformulaciones del Problema de Equilibrio de Nash Generalizado como problemas de desigualdad variacional. ¿Por qué serían útiles tales reformulaciones? Pues porque la teoría de las desigualdades variacionales está muy desarrollada<sup>13</sup>, tanto teórica como algorítmicamente, de modo que existe la posibilidad de que el problema reformulado sea más manejable que el problema original.

En esta sección estudiamos brevemente las desigualdades variacionales, estableciendo las notaciones y la terminología a usar luego. Además del libro de Facchinei y Pang [13], hemos usado la referencia [8], de la cual, por ejemplo, hemos obtenido la prueba del Teorema 7.6. La parte sobre condiciones KKT está basada en [13].

### 7.1. El problema de desigualdad variacional

Una consecuencia inmediata de las Proposiciones 3.1 y 3.3 es que las soluciones del problema de minimización

$$\min f(x), \quad x \in X, \quad (7.1)$$

donde  $X \subset \mathbb{R}^n$  es convexo y  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es convexa y diferenciable, son precisamente los puntos factibles  $\bar{x} \in X$  tales que

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \text{para todo } x \in X. \quad (7.2)$$

Los problemas que consisten en encontrar puntos  $\bar{x} \in X$  que satisfagan (7.2) son en general llamados *problemas de desigualdad variacional*.

Dada una función  $F : X \rightarrow \mathbb{R}^n$ , donde  $X$  es un subconjunto de  $\mathbb{R}^n$ , el **problema de desigualdad variacional** asociado a  $F$ , denotado por  $VI(F, X)$ , consiste en hallar puntos  $\bar{x} \in X$  tales que

$$\langle F(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \text{para todo } x \in X.$$

Decimos entonces que  $\bar{x}$  (si existe) es una **solución** del problema  $VI(F, X)$ . Evidentemente, todo punto  $\bar{x}$  tal que  $F(\bar{x}) = 0$  es solución de  $VI(F, X)$ .

**Ejemplo 7.1.** La observación hecha al inicio de esta subsección nos dice que las soluciones del problema de optimización (7.1) coinciden con las soluciones del problema de desigualdad variacional  $VI(F, X)$ , donde  $F(x) = \nabla f(x)$ . Esta fue una de las primeras conexiones que se establecieron entre la teoría de las desigualdades variacionales y la Optimización matemática.

<sup>13</sup>En dos volúmenes [13], que conforman más de mil páginas, Facchinei y Pang hacen un estudio pormenorizado del problema de desigualdad variacional y del problema de complementariedad. Entre muchas cosas, este libro muestra que existen una teoría matemática muy amplia al respecto (existencia y unicidad de soluciones), una infinidad de algoritmos y numerosas conexiones con otras disciplinas. He ahí la importancia de las desigualdades variacionales.

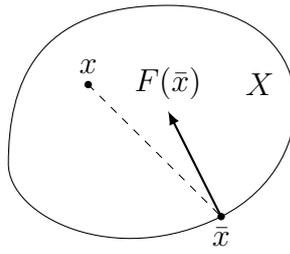


Figura 7.1: Geométricamente, el problema de desigualdad variacional  $VI(F, X)$  consiste en encontrar un punto  $\bar{x} \in X$  tal que el vector  $F(\bar{x})$  forme un ángulo agudo con todos los vectores de la forma  $x - \bar{x}$  para cada  $x \in X$ .

La siguiente definición extiende el problema  $VI(F, X)$  al caso en que la función  $F$  no es una función sino una multifunción.

Dado el operador multivaluado  $\Omega : X \rightrightarrows \mathbb{R}^n$ , donde  $X$  es un subconjunto de  $\mathbb{R}^n$ , el **problema de desigualdad variacional** asociado a  $\Omega$ , denotado por  $VI(\Omega, X)$ , consiste en hallar puntos  $\bar{x} \in X$  tales que exista  $v \in \Omega(\bar{x})$  con la propiedad de que

$$\langle v, x - \bar{x} \rangle \geq 0 \quad \text{para todo } x \in X$$

Decimos entonces que  $\bar{x}$  (si existe) es una **solución** del problema  $VI(\Omega, X)$ .

De nuevo, es claro que si  $0 \in \Omega(\bar{x})$  entonces  $\bar{x}$  es solución de  $VI(\Omega, X)$ .

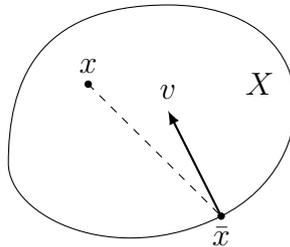


Figura 7.2: Geométricamente, el problema de desigualdad variacional consiste en encontrar un punto  $\bar{x} \in X$  y un vector  $v \in \Omega(\bar{x})$  que forme un ángulo agudo con todos los vectores de la forma  $x - \bar{x}$  para cada  $x \in X$ .

## Condiciones KKT

Como hicimos en la Sección 3, ahora averiguamos qué condiciones (necesarias y/o suficientes) deben satisfacer las soluciones del problema de desigualdad variacional  $VI(F, X)$  cuando el conjunto  $X$  tiene la forma

$$X = \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \forall i = 1, \dots, r\},$$

donde  $g_1, \dots, g_r : \mathbb{R}^n \rightarrow \mathbb{R}$  son funciones diferenciables. De nuevo, veremos que esta cuestión va a estar relacionada con los *multiplicadores de Lagrange*. Escribamos  $g := (g_1, \dots, g_r) : \mathbb{R}^n \rightarrow \mathbb{R}^r$ .

Dado un punto factible  $\bar{x} \in X$ , el **conjunto activo** de  $\bar{x}$  está dado por

$$I(\bar{x}) = \{i \mid g_i(\bar{x}) = 0\}.$$

Un vector  $\lambda = (\lambda_1, \dots, \lambda_r)$  es llamado **vector de multiplicadores de Lagrange** en  $\bar{x}$  si las siguientes condiciones se satisfacen

- (1)  $F(\bar{x}) + \sum_i \lambda_i \nabla g_i(\bar{x}) = 0$ ;
- (2)  $\lambda_i \geq 0$  para todo  $i$ ;
- (3)  $\lambda_i = 0$  para todo  $i \notin I(\bar{x})$ .

Abreviadamente, las condiciones (2) y (3) pueden escribirse así

$$0 \leq \lambda \perp g(\bar{x}) \leq 0.$$

Notemos que esta noción de multiplicadores de Lagrange coincide con la que vimos en la Sección 3 cuando  $F(x) = \nabla f(x)$ .

Una **condición de calificación (CQ)** para el problema  $VI(F, X)$  es una condición suficiente para que en una solución  $\bar{x} \in X$  del problema exista un vector de multiplicadores de Lagrange.

Para cada  $x \in X$  sea  $K(x) := \text{con}(\{\nabla g_i(x) \mid i \in I(x)\})$  el cono convexo generado por los vectores  $\{\nabla g_i(x)\}_{i \in I(x)}$ , que a veces es llamado la *linealización* de  $X$  en  $x$ . Sabemos por la Proposición 2.20 que cada  $K(x)$  es un cono convexo y cerrado tal que  $K(x)^{**} = K(x)$ . Afirmamos que

$$T(x, X) \subset K(x)^*, \tag{7.3}$$

donde  $T(x, X)$  es el cono tangente de  $X$  en  $x$ , definido en la página 29. En efecto, sea  $v$  un vector no nulo en el cono tangente  $T(x, X)$ , esto es, existe una sucesión  $(x_k)$  en  $X$  tal que  $\lim \frac{x_k - x}{\|x_k - x\|} = \frac{v}{\|v\|}$ . Fijemos  $i \in I(x)$ . De la diferenciabilidad de  $g_i$  tenemos que

$$g_i(x_k) = g_i(x) + \langle \nabla g_i(x), x_k - x \rangle + r(x_k - x)$$

para todo  $k$ , donde  $\lim_{h \rightarrow 0} \frac{r(h)}{\|h\|} = 0$ . Como  $g_i(x_k) \leq 0 = g_i(x)$ , tomando límite en

$$\left\langle \nabla g_i(x), \frac{x_k - x}{\|x_k - x\|} \right\rangle + \frac{r(x_k - x)}{\|x_k - x\|} \leq 0$$

obtenemos  $\langle \nabla g_i(x), \frac{v}{\|v\|} \rangle \leq 0$ , es decir,  $\langle \nabla g_i(x), v \rangle \leq 0$ . Esto prueba (7.3).

Decimos que la **condición de calificación de Abadie (CQ de Abadie)** se satisface en  $\bar{x} \in X$  cuando la inclusión (7.3) es una igualdad, es decir, cuando

$$T(\bar{x}, X) = \{v \in \mathbb{R}^n \mid \langle \nabla g_i(\bar{x}), v \rangle \leq 0, \forall i \in I(\bar{x})\}.$$

Notemos que en todo punto  $\bar{x} \in X$  en el que la CQ de Abadie se satisface tenemos que

$$T(\bar{x}, X)^* = \text{con}(\{\nabla g_i(\bar{x})\}_{i \in I(\bar{x})}),$$

es decir,

$$-v \in T(\bar{x}, X)^* \Leftrightarrow \exists \lambda = (\lambda_1, \dots, \lambda_r) : \begin{aligned} v + \sum_i \lambda_i \nabla g(\bar{x}) &= 0 \\ 0 \leq \lambda \perp g(\bar{x}) &\leq 0 \end{aligned},$$

lo cual implica en particular que  $-F(\bar{x}) \in T(\bar{x}, X)^*$  si y solo si existe un vector de multiplicadores de Lagrange en  $\bar{x}$ .

**Teorema 7.2** (KKT). *Supongamos que  $\bar{x} \in X$  es una solución del problema  $VI(F, X)$  y que la CQ de Abadie se satisface en  $\bar{x}$ . Entonces existe un vector de multiplicadores de Lagrange en  $\bar{x}$ .*

PRUEBA. Sabemos que  $\langle F(\bar{x}), x - \bar{x} \rangle \leq 0$  para todo  $x \in X$ . Dado un vector no nulo  $v \in T(\bar{x}, X)$ , existen sucesiones  $(x_k) \subset X$  y  $(t_k) \subset (0, \infty)$  con  $x_k \rightarrow \bar{x}$ ,  $t_k \rightarrow 0$  y  $\frac{x_k - \bar{x}}{t_k} \rightarrow v$ , de modo que

$$\langle F(\bar{x}), v \rangle = \lim_{k \rightarrow \infty} \frac{1}{t_k} \langle F(\bar{x}), x_k - \bar{x} \rangle \geq 0.$$

Luego  $-F(\bar{x}) \in T(\bar{x}, X)^*$ , lo cual, por el comentario que hicimos antes de enunciar el teorema, significa que existe un vector de multiplicadores de Lagrange en  $\bar{x}$ .  $\square$

**Teorema 7.3.** *Supongamos que las funciones  $g_i$  son convexas. Si en un punto factible dado  $\bar{x} \in X$  existen los multiplicadores de Lagrange, entonces  $\bar{x}$  es solución del problema  $VI(F, X)$ .*

PRUEBA. Este teorema es consecuencia de la Observación 3.10. En efecto, como las funciones  $g_i$  son convexas y como existe  $\lambda = (\lambda_1, \dots, \lambda_r)$  tal que  $0 \leq \lambda \perp g(\bar{x}) \leq 0$  y

$$F(\bar{x}) = - \sum_i \lambda_i \nabla g_i(\bar{x}),$$

obtenemos  $\langle F(\bar{x}), x - \bar{x} \rangle \leq 0$  para todo  $x \in X$ , esto es,  $\bar{x}$  es solución de  $VI(F, X)$ .  $\square$

En vista de la equivalencia entre el problema de minimización (7.1) y el problema de desigualdad variacional  $VI(F, X)$  cuando  $F(x) = \nabla f(x)$ , el Teorema 7.3 puede considerarse como una extensión del Teorema 3.9.

## Existencia de solución

Habiendo analizado las condiciones KKT para el problema  $VI(F, X)$ , ahora nos ocupamos de una cuestión fundamental: ¿en qué casos el problema de desigualdad variacional posee soluciones? El teorema de punto fijo de Brouwer (Teorema 4.18) nos ayuda a dar una primera respuesta.

**Teorema 7.4.** *Sea  $X$  un subconjunto convexo, compacto y no vacío de  $\mathbb{R}^n$ . Si la función  $F : X \rightarrow \mathbb{R}^n$  es continua, entonces el problema  $VI(F, X)$  posee solución.*

Antes de comenzar la demostración, escribamos  $I$  para denotar la función identidad de  $\mathbb{R}^n$ , y recordemos que todo subconjunto convexo y cerrado  $X$  de  $\mathbb{R}^n$  tiene asociada una función proyección  $P_X : \mathbb{R}^n \rightarrow X$  (véase la Subsección 2.1).

PRUEBA. Debido a que el conjunto  $X - F(X)$  es compacto ( $X$  es compacto y  $F$  es continua), gracias al Corolario 2.5 el conjunto  $C := \text{co}(X - F(X))$  es convexo y compacto. Como la función continua  $(I - F) \circ P_X : \mathbb{R}^n \rightarrow C$  aplica  $C$  en  $C$ , posee un punto fijo por el teorema de punto fijo de Brouwer, es decir, existe  $y \in C$  tal que  $y = (I - F) \circ P_X(y) = \bar{x} - F(\bar{x})$ , donde  $\bar{x} := P_X(y) \in X$ . Luego, de la Proposición 2.11 obtenemos

$$-\langle F(\bar{x}), x - \bar{x} \rangle = \langle y - \bar{x}, x - \bar{x} \rangle \leq 0 \quad \text{para todo } x \in X,$$

esto es,  $\bar{x}$  es solución del problema  $VI(F, X)$ .  $\square$

Vamos a extender el teorema anterior al caso de  $VI(\Omega, X)$ , o sea cuando la función  $F$  es reemplazada por una multifunción  $\Omega : X \rightrightarrows \mathbb{R}^n$ . Veremos que este proceso será una generalización de la demostración de dicho teorema. Así, en vez del teorema de Brouwer, emplearemos el teorema de Kakutani-Fan (Teorema 4.19).

**Lema 7.5.** *Supongamos que el conjunto  $X$  es convexo y cerrado. Entonces resolver el problema  $VI(\Omega, X)$  es equivalente a encontrar un punto fijo del operador multivaluado  $(I - \Omega) \circ P_X : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ .*

Naturalmente,  $(I - \Omega) \circ P_X(z) = P_X(z) - \Omega(P_X(z))$ .

PRUEBA. Supongamos que  $\bar{x} \in X$  es solución del problema  $VI(\Omega, X)$ , esto es, existe  $v \in \Omega(\bar{x})$  tal que  $\langle v, x - \bar{x} \rangle \geq 0$  para todo  $x \in X$ . Escribiendo  $y := \bar{x} - v$  tenemos

$$\langle y - \bar{x}, x - \bar{x} \rangle \leq 0 \quad \text{para todo } x \in X,$$

lo cual en virtud de la Proposición 2.11 significa que  $\bar{x} = P_X(y)$ , y por consiguiente

$$y = P_X(y) - v \in P_X(y) - \Omega(P_X(y)) = (I - \Omega) \circ P_X(y),$$

es decir,  $y$  es un punto fijo de  $(I - \Omega) \circ P_X$ .

Recíprocamente, si  $y \in \mathbb{R}^n$  es un punto fijo de  $(I - \Omega) \circ P_X$ , escribiendo  $\bar{x} := P_X(y)$  vemos que el vector  $v := \bar{x} - y$  está en  $\Omega(\bar{x})$  pues

$$y \in (I - \Omega) \circ P_X(y) = P_X(y) - \Omega(P_X(y)) = \bar{x} - \Omega(\bar{x});$$

como  $\bar{x} = P_X(y)$ , también tenemos que  $\langle y - \bar{x}, x - \bar{x} \rangle \leq 0$  para todo  $x \in X$ , de donde se sigue que

$$\langle v, x - \bar{x} \rangle \geq 0 \quad \text{para todo } x \in X,$$

esto es,  $\bar{x} \in X$  es solución del problema  $VI(\Omega, X)$ .  $\square$

Notemos que  $X - \Omega(X)$  contiene a la imagen del operador  $(I - \Omega) \circ P_X$  que aparece en el lema anterior: si  $z \in \mathbb{R}^n$  entonces  $P_X(z) \in X$  y

$$(I - \Omega) \circ P_X(z) = P_X(z) - \Omega(z) \in X - \Omega(X).$$

**Teorema 7.6.** *Sea  $X$  un subconjunto convexo, compacto y no vacío de  $\mathbb{R}^n$ . Si el operador multivaluado  $\Omega : X \rightrightarrows \mathbb{R}^n$  es scs y tiene imágenes compactas, convexas y no vacías, entonces el problema  $VI(\Omega, X)$  tiene solución.*

PRUEBA. Mostraremos que  $(I - \Omega) \circ P_X$  tiene algún punto fijo utilizando el teorema de Kakutani-Fan (Teorema 4.19). Como ya fue observado, el conjunto  $X - \Omega(X)$  contiene a la imagen de  $(I - \Omega) \circ P_X$ . Por la Proposición 4.16, las condiciones del operador  $\Omega$  nos garantizan que la imagen  $\Omega(X)$  de  $\Omega$  es compacta; luego,  $X - \Omega(X)$  es compacto y, por el Corolario 2.5, el conjunto  $K := \text{co}(X - \Omega(X))$  también es compacto.

Por otro lado, al ser  $\Omega$  scs con imágenes convexas, compactas y no vacías, el operador  $(I - \Omega) \circ P_X$  también es scs con imágenes convexas, compactas y no vacías (véase el Ejemplo 4.4).

Así, la multifunción  $(I - \Omega) \circ P_X : K \rightrightarrows K$  satisface las hipótesis del teorema de Kakutani-Fan y por tanto posee un punto fijo.  $\square$

## 7.2. El problema de desigualdad cuasivariacional

El problema de desigualdad cuasivariacional es la extensión natural del problema de desigualdad variacional, cuando el conjunto del problema no permanece fijo sino que cambia con la variable.

Sea la multifunción  $\Lambda : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ , la cual va a reemplazar al conjunto  $X$  de la definición del problema de desigualdad variacional.

Dada la función  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , el **problema de desigualdad cuasivariacional** asociado a  $F$ , denotado por  $QVI(F, \Lambda)$ , consiste en hallar puntos  $\bar{x} \in \Lambda(\bar{x})$  tales que

$$\langle F(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \text{para todo } x \in \Lambda(\bar{x}).$$

Decimos entonces que  $\bar{x}$  (si existe) es una **solución** del problema  $QVI(F, \Lambda)$ . Obviamente, si  $\bar{x} \in \Lambda(\bar{x})$  y  $F(\bar{x}) = 0$  entonces  $\bar{x}$  es solución del problema  $QVI(F, \Lambda)$ .

Similarmente, dado el operador multivaluado  $\Omega : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ , el **problema de desigualdad cuasivariacional** asociado a  $\Omega$ , denotado por  $QVI(\Omega, \Lambda)$ , consiste en hallar puntos  $\bar{x} \in \Lambda(\bar{x})$  tales que exista algún  $v \in \Omega(\bar{x})$  con la propiedad de que

$$\langle v, x - \bar{x} \rangle \geq 0 \quad \text{para todo } x \in \Lambda(\bar{x}).$$

Entonces decimos que  $\bar{x}$  (si existe) es una **solución** del problema  $QVI(\Omega, \Lambda)$ . Es claro que si  $\bar{x} \in \Lambda(\bar{x})$  y  $0 \in \Omega(\bar{x})$  entonces  $\bar{x}$  es solución de  $QVI(\Omega, \Lambda)$ .

**Ejemplo 7.7.** El problema de encontrar los puntos fijos de un operador multivaluado  $\Lambda : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  es un problema de desigualdad cuasivariacional. En efecto, dicho problema equivale al problema de desigualdad cuasivariacional  $QVI(0, \Lambda)$ , donde  $0 : \mathbb{R}^n \rightarrow \mathbb{R}^n$  es la función idénticamente nula  $0(x) = 0$ .

A diferencia del problema de desigualdad variacional, existen muy pocas herramientas para abordar el problema de desigualdad cuasivariacional. Esto se debe a que esta teoría todavía no está muy desarrollada<sup>14</sup>. Es por ello que, por lo general, si para un problema dado existe una reformulación como problema de desigualdad cuasivariacional, en la mayoría de casos ella es útil casi únicamente desde un punto de vista teórico.

<sup>14</sup>El hecho de que el conjunto  $\Lambda(x)$  dependa de  $x$  hace que el problema de desigualdad cuasivariacional tenga un alto grado de complejidad.

## 8. Condiciones de optimalidad

Dados una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  y un subconjunto convexo  $X$  de  $\mathbb{R}^n$ , consideremos el problema de minimización

$$\min f(x), \quad x \in X. \quad (8.1)$$

Recordemos de la Sección 3 que una condición de optimalidad para este problema es una condición necesaria y/o suficiente para que un vector factible  $x \in X$  sea una solución óptima.

Habiendo estudiado las desigualdades variacionales en la sección anterior, en esta sección estamos interesados en reformular el problema (8.1) como problemas de desigualdad variacional, esto es, queremos encontrar condiciones de optimalidad en términos de problemas de desigualdad variacional. En general, una reformulación de un problema de optimización como un problema de desigualdad variacional es útil porque nos permite abordar el problema de optimización utilizando herramientas de la teoría de desigualdades variacionales, que, como ya sabemos, es amplia.

Vamos a considerar varios casos, y, conforme avancemos, iremos debilitando nuestras hipótesis. Las condiciones de optimalidad de las Subsecciones 8.1 y 8.2 son resultados clásicos. La Subsección 8.3 está basada en [3].

### 8.1. Caso convexo y diferenciable

El siguiente resultado es consecuencia de las Proposiciones 3.1 y 3.3.

**Proposición 8.1.** *Supongamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función convexa y diferenciable. Sea  $\bar{x}$  un elemento del conjunto convexo  $X \subset \mathbb{R}^n$ . Entonces  $\bar{x}$  minimiza  $f$  sobre  $X$  si y solo si*

$$\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0 \quad \text{para todo } x \in X.$$

En otras palabras,  $\bar{x} \in X$  es solución del problema (8.1) si y solo si es solución del problema de desigualdad variacional  $VI(\nabla f, X)$ , donde  $\nabla f$  denota la función  $x \mapsto \nabla f(x)$ . Es decir, la proposición nos proporciona una reformulación del problema (8.1) como un problema de desigualdad variacional para el caso en que  $f$  es convexa y diferenciable.

**Observación 8.2** (Caso pseudoconvexo). Puesto que las Proposiciones 3.1 y 3.3 valen cuando la función  $f$  es pseudoconvexa (en vez de convexa y diferenciable), existe una obvia generalización de la Proposición 8.1 al contexto más general en que  $f$  es solo pseudoconvexa. Esto significa que la equivalencia entre los problemas (8.1) y  $VI(\nabla f, X)$  se mantiene bajo la hipótesis de pseudoconvexidad.

Así, la equivalencia clásica entre el problema de minimización (8.1) y el problema de desigualdad variacional  $VI(\nabla f, X)$ , que se da cuando  $f$  es convexa y diferenciable, puede generalizarse cuando debilitamos la hipótesis de convexidad. En la siguiente sección debilitamos la hipótesis de diferenciable.

## 8.2. Caso convexo

Supongamos que  $f$  es convexa, mas no necesariamente diferenciable. Entonces no podemos usar el gradiente de  $f$  pero sí sus subdiferenciales.

De hecho, la multifunción  $\partial f : x \mapsto \partial f(x)$  nos da una condición de optimalidad natural para el problema de minimización global

$$\min f(x), \quad x \in \mathbb{R}^n.$$

**Proposición 8.3.** *Supongamos que  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función convexa y sea  $\bar{x} \in \mathbb{R}^n$ . Entonces  $\bar{x} \in \arg \min f$  si y solo si  $0 \in \partial f(\bar{x})$ .*

PRUEBA. En efecto,  $0 \in \partial f(\bar{x})$  si y solo si  $f(x) \geq f(\bar{x}) + \langle 0, x - \bar{x} \rangle$  para todo  $x \in \mathbb{R}^n$ , esto es, si y solo si  $f(\bar{x}) \leq f(x)$  para todo  $x \in \mathbb{R}^n$ .  $\square$

Como  $\partial f(x) = \{\nabla f(x)\}$  cuando  $f$  es convexa y diferenciable (véase el Ejemplo 2.33), la Proposición 8.1 nos dice que los problemas (8.1) y  $VI(\partial f, X)$  son equivalentes cuando  $f$  es convexa y diferenciable. En realidad, esto se sigue cumpliendo cuando  $f$  es solo convexa.

**Teorema 8.4.** *Sean  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  una función convexa y  $X \subset \mathbb{R}^n$  convexo. Entonces  $\bar{x} \in X$  minimiza  $f$  sobre  $X$  si y solo si  $\bar{x}$  es solución del problema de desigualdad variacional  $VI(\partial f, X)$ .*

PRUEBA. Si  $\bar{x}$  es solución del problema de desigualdad variacional  $VI(\partial f, X)$  entonces existe  $v \in \partial f(\bar{x})$  tal que  $\langle v, x - \bar{x} \rangle \geq 0$  para todo  $x \in X$ , lo cual implica

$$f(x) \geq f(\bar{x}) + \langle v, x - \bar{x} \rangle \geq f(\bar{x})$$

para todo  $x \in X$ , esto es,  $\bar{x}$  minimiza  $f$  sobre  $X$ .

Recíprocamente, supongamos que  $\bar{x}$  minimiza  $f$  sobre  $X$ . Como  $f$  es continua,  $\bar{x}$  minimiza  $f$  sobre el conjunto convexo y cerrado  $\bar{X}$ . Como  $\bar{x}$  minimiza la función convexa  $f$  sobre el conjunto convexo  $K := \bar{B}(\bar{x}, 1) \cap \bar{X}$ , del Corolario 3.2 y de la Proposición 2.35 obtenemos

$$0 \leq f'(\bar{x}; x - \bar{x}) = \max_{v \in \partial f(\bar{x})} \langle v, x - \bar{x} \rangle \quad \text{para todo } x \in K,$$

de donde se sigue que

$$\min_{x \in K} \max_{v \in \partial f(\bar{x})} \langle v, x - \bar{x} \rangle = 0,$$

ya que  $\bar{x} \in K$  y  $\max_{v \in \partial f(\bar{x})} \langle v, \bar{x} - \bar{x} \rangle = 0$ .

Ahora bien, el conjunto  $\partial f(\bar{x})$  es convexo y compacto por la Proposición 2.36, al igual que el conjunto  $K$ . Como la función  $\varphi : \partial f(\bar{x}) \times K \rightarrow \mathbb{R}$ ,  $\varphi(v, x) = \langle v, x - \bar{x} \rangle$  verifica las hipótesis del teorema de von Neumann (Teorema 4.22), se tiene que

$$\max_{v \in \partial f(\bar{x})} \min_{x \in K} \langle v, x - \bar{x} \rangle = \min_{x \in K} \max_{v \in \partial f(\bar{x})} \langle v, x - \bar{x} \rangle = 0,$$

y por consiguiente existe  $v \in \partial f(\bar{x})$  tal que  $0 = \min_{x \in K} \langle v, x - \bar{x} \rangle$ , esto es,

$$0 \leq \langle v, x - \bar{x} \rangle \quad \text{para todo } x \in K.$$

Finalmente, dado  $x \in X$ , tomando  $t \in (0, 1)$  suficientemente pequeño tal que el vector  $x_t := \bar{x} + t(x - \bar{x})$  esté en  $\overline{B}(\bar{x}, 1) \cap \overline{X} = K$  obtenemos

$$0 \leq \langle v, x_t - \bar{x} \rangle = t \langle v, x - \bar{x} \rangle,$$

es decir,  $0 \leq \langle v, x - \bar{x} \rangle$ . Esto muestra que  $\bar{x}$  es solución del problema  $VI(\partial f, X)$ .  $\square$

Concluimos que el problema (8.1) es equivalente al problema de desigualdad variacional  $VI(\partial f, X)$  en el caso en que  $f$  es convexa.

Debilitando la hipótesis de convexidad, en la siguiente sección vamos a obtener una nueva reformulación del problema (8.1).

### 8.3. Caso cuasiconvexo

Analicemos ahora el caso en que  $f$  ya no es convexa sino tan solo cuasiconvexa. Al igual que en el caso convexo queremos reformular el problema (8.1) como un problema de desigualdad variacional. Aquí no tenemos la multifunción  $\partial f$  pues  $f$  no es convexa, pero podemos utilizar los subniveles ajustados y sus operadores normales (véase la Subsección 5.2).

Dada una función cuasiconvexa  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  definimos el operador multivaluado  $F_f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  mediante

$$F_f(x) = \begin{cases} \overline{B}(0, 1), & \text{si } x \in \arg \min f, \\ \text{co}(N_f^a(x) \cap S(0, 1)), & \text{si } x \notin \arg \min f, \end{cases} \quad (8.2)$$

donde  $\overline{B}(0, 1)$  y  $S(0, 1)$  son la bola unitaria cerrada y la esfera de  $\mathbb{R}^n$  respectivamente.

Veremos que el operador  $F : f \mapsto F_f$  tiene propiedades similares a las del operador subdiferencial  $\partial : f \mapsto \partial f$ . De hecho, la Proposición 2.36 también se cumple en este contexto.

**Proposición 8.5.** *Si  $f$  es cuasiconvexa entonces el operador  $F_f$  tiene imágenes convexas, compactas y no vacías.*

PRUEBA. Dado  $x \in \mathbb{R}^n$ , es claro que si  $x \in \arg \min f$  el conjunto  $F_f(x) = \overline{B}(0, 1)$  es convexo, compacto y no vacío. Si  $x \notin \arg \min f$  entonces, como  $f$  es cuasiconvexa, la Proposición 5.10 nos asegura que  $N_f^a(x)$  es un cono convexo y cerrado que posee elementos no nulos. Luego,  $N_f^a(x) \cap S(0, 1)$  es no vacío y compacto, y, por el Corolario 2.5, el conjunto  $F_f(x) = \text{co}(N_f^a(x) \cap S(0, 1))$  es no vacío, convexo y compacto. Así,  $F_f$  tiene imágenes convexas, compactas y no vacías.  $\square$

La condición de optimalidad de la Proposición 8.3 también se satisface para  $F_f$ .

**Proposición 8.6.** *Supongamos que  $f$  es cuasiconvexa y continua y sea  $\bar{x} \in \mathbb{R}^n$ . Entonces  $\bar{x} \in \arg \min f$  si y solo si  $0 \in F_f(\bar{x})$ .*

PRUEBA. Supongamos que  $\bar{x} \in \arg \min f$ . Entonces  $0 \in \overline{B}(0, 1) = F_f(\bar{x})$ . Recíprocamente, supongamos que  $0 \in F_f(\bar{x})$  y que  $\bar{x} \notin \arg \min f$ . Como  $f$  es continua,  $S_{f(\bar{x})}^<$  es un conjunto abierto y no vacío que está contenido en  $S_f^a(\bar{x})$ , de modo que  $\text{int}(S_f^a(\bar{x})) \neq \emptyset$ . Luego, la Proposición 2.22 nos da  $N_f^a(\bar{x}) \cap (-N_f^a(\bar{x})) = \{0\}$ .

Por el teorema de Carathéodory, de  $0 \in F_f(\bar{x}) = \text{co}(N_f^a(\bar{x}) \cap S(0, 1))$  se sigue que existen vectores  $v_1, \dots, v_{n+1} \in N_f^a(\bar{x}) \cap S(0, 1)$  y escalares  $\lambda_1, \dots, \lambda_{n+1} \in [0, 1]$  tales que

$$\sum_{i=1}^{n+1} \lambda_i = 1 \quad \text{y} \quad \sum_{i=1}^{n+1} \lambda_i v_i = 0.$$

Ahora bien, cada  $v_i$  es no nulo. Además, existe algún  $\lambda_k > 0$ , de forma que

$$-v_k = \sum_{i=1, i \neq k}^{n+1} \frac{\lambda_i}{\lambda_k} v_i,$$

lo cual, considerando que  $N_f^a(\bar{x})$  es un cono convexo, implica que  $-v_k \in N_f^a(\bar{x})$ , es decir,  $v_k \in -N_f^a(\bar{x})$ . Así,  $0 \neq v_k \in N_f^a(\bar{x}) \cap (-N_f^a(\bar{x})) = \{0\}$ , una contradicción. Esto concluye la prueba de la proposición.  $\square$

El Teorema 8.4 también se generaliza a este caso para  $F_f$  mas debemos imponer algunas restricciones adicionales sobre  $f$ . Cuando  $f$  es cuasiconvexa tenemos una condición suficiente de optimalidad para el problema (8.1).

**Teorema 8.7** (Condición suficiente de optimalidad). *Sea  $X$  un subconjunto no vacío de  $\mathbb{R}^n$  y supongamos que  $f$  es cuasiconvexa y continua. Si  $\bar{x} \in X$  es solución del problema de desigualdad variacional  $VI(F_f, X)$  entonces  $\bar{x}$  minimiza  $f$  sobre  $X$ .*

PRUEBA. Como  $\bar{x}$  es solución del problema  $VI(F_f, X)$ , existe  $v \in F_f(\bar{x})$  tal que

$$\langle v, x - \bar{x} \rangle \geq 0, \quad \forall x \in X.$$

Si  $\bar{x} \in \arg \min f$  no hay nada que probar. Si  $\bar{x} \notin \arg \min f$  tenemos  $0 \notin F_f(\bar{x})$  por la Proposición 8.6; también tenemos  $v \in F_f(\bar{x}) = \text{co}(N_f^a(\bar{x}) \cap S(0, 1)) \subset N_f^a(\bar{x})$ , de donde  $v \in N_f^a(\bar{x}) \setminus \{0\}$ . Entonces  $v$  es un vector no nulo tal que

$$\langle v, y - \bar{x} \rangle \leq 0, \quad \forall y \in S_f^a(\bar{x}),$$

y combinando esto con la ecuación anterior vemos que

$$\langle v, y \rangle \leq \langle v, \bar{x} \rangle \leq \langle v, x \rangle, \quad \forall y \in S_f^a(\bar{x}), \forall x \in X.$$

De esto tenemos que  $S_f^a(\bar{x})$  está contenido en el conjunto  $\{y \in \mathbb{R}^n \mid \langle v, y \rangle \leq \langle v, \bar{x} \rangle\}$ , cuyo interior es  $\{y \in \mathbb{R}^n \mid \langle v, y \rangle < \langle v, \bar{x} \rangle\}$ . (Aquí hemos usado que  $v \neq 0$ .) Además,  $S_{f(\bar{x})}^<$  es abierto (pues  $f$  es continua) y está contenido en  $S_f^a(\bar{x})$ . Obtenemos así que

$$S_{f(\bar{x})}^< \subset \text{int}(S_f^a(\bar{x})) \subset \{y \in \mathbb{R}^n \mid \langle v, y \rangle < \langle v, \bar{x} \rangle\},$$

y como también  $X \subset \{z \in \mathbb{R}^n \mid \langle v, z \rangle \geq \langle v, \bar{x} \rangle\}$ , concluimos que  $X \cap S_{f(\bar{x})}^< = \emptyset$ . Esto muestra que  $\bar{x}$  minimiza  $f$  en  $X$ .  $\square$

El siguiente ejemplo muestra que el recíproco del teorema anterior no se verifica.

**Ejemplo 8.8.** La función  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  definida por  $f(x) = \min\{1, \|x\|\}$  es cuasiconvexa y tiene la propiedad de que el punto  $\bar{x} := (2, 0)$  minimiza  $f$  sobre el conjunto convexo  $X := \overline{B}(\bar{x}, 1)$ . Pero  $\bar{x}$  no es solución del problema  $VI(F_f, X)$ . En efecto, por un lado  $S_{f(\bar{x})} = \mathbb{R}^2$  implica

$$S_f^a(\bar{x}) = S_{f(\bar{x})} \cap \overline{B(S_{f(\bar{x})}^<, 1)} = \{x \in \mathbb{R}^2 \mid \|x\| \leq 2\},$$

de donde se sigue que  $N_f^a(\bar{x}) = \{(t, 0) \mid t \geq 0\}$ , y por tanto

$$F_f(\bar{x}) = \text{co}(N_f^a(\bar{x}) \cap S(0, 1)) = \{v\},$$

donde  $v := (1, 0)$ . Por otro lado,  $x := (1, 0)$  está en  $X$  y  $\langle v, x - \bar{x} \rangle = -1 < 0$ .

No obstante, cuando  $f$  es cuasiconvexa semiestricta el recíproco del Teorema 8.7 se cumple. Conseguimos así una reformulación completa del problema (8.1) como un problema de desigualdad variacional.

**Teorema 8.9** (Condición necesaria y suficiente de optimalidad). *Sea  $X$  un subconjunto convexo y no vacío de  $\mathbb{R}^n$  y supongamos que  $f$  es cuasiconvexa semiestricta y continua. Entonces  $\bar{x} \in X$  es solución del problema  $VI(F_f, X)$  si y solo si  $\bar{x}$  minimiza  $f$  sobre  $X$ .*

PRUEBA. Supongamos que  $\bar{x}$  minimiza  $f$  sobre  $X$ . Afirmamos que  $\bar{x}$  es solución de  $VI(F_f, X)$ . En efecto, esto es claramente cierto si  $\bar{x} \in \arg \min f$  pues en este caso se cumple  $0 \in \overline{B}(0, 1) = F(\bar{x})$ .

Si  $\bar{x} \notin \arg \min f$ , tenemos  $F(\bar{x}) = \text{co}(N_f^a(\bar{x}) \cap S(0, 1))$ . Entonces, por la Proposición 5.8, el conjunto  $S_f^a(\bar{x})$  es convexo y cerrado y satisface

$$\text{int}(S_f^a(\bar{x})) = S_{f(\bar{x})}^< ,$$

donde el conjunto de la derecha de la igualdad anterior es no vacío y no interseca a  $X$  (pues  $\bar{x}$  minimiza  $f$  en  $X$ ). Luego, por el teorema de separación existe  $v \in \mathbb{R}^n$  no nulo tal que

$$\langle v, y \rangle \leq \langle v, \bar{x} \rangle \leq \langle v, x \rangle, \quad \forall x \in X, \forall y \in S_f^a(\bar{x}).$$

Tomando  $\frac{v}{\|v\|}$  en vez de  $v$  podemos suponer sin pérdida de generalidad que  $\|v\| = 1$ , esto es, que  $v \in S(0, 1)$ . Además,

$$\begin{aligned} \langle v, y - \bar{x} \rangle &\leq 0, & \forall y \in S_f^a(\bar{x}), \\ \langle v, x - \bar{x} \rangle &\geq 0, & \forall x \in X, \end{aligned}$$

y de esto se sigue que  $v \in N_f^a(\bar{x}) \cap S(0, 1) \subset F_f(\bar{x})$  y que  $\bar{x}$  es solución del problema  $VI(F_f, X)$ . Esto muestra una de las implicancias. La otra es consecuencia directa del Teorema 8.7.  $\square$

Para terminar esta sección probaremos el análogo de la Proposición 2.37 para el operador  $F$ .

**Teorema 8.10.** *Sea  $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$  una función continua tal que la aplicación  $f^y : \mathbb{R}^m \rightarrow \mathbb{R}$ ,  $f^y(x) = f(x, y)$  es cuasiconvexa semiestricta para cada  $y \in \mathbb{R}^n$ . Sea también la multifunción  $\Omega : \mathbb{R}^m \times \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  dada por  $\Omega(x, y) = F_{f^y}(x)$ ; explícitamente,*

$$\Omega(x, y) = \begin{cases} \overline{B}_m(0, 1), & \text{si } x \in \arg \min f^y, \\ \text{co}(N_{f^y}^a(x) \cap S_m(0, 1)), & \text{si } x \notin \arg \min f^y, \end{cases}$$

donde  $\overline{B}_m(0, 1)$  y  $S_m(0, 1)$  son la bola unitaria cerrada y la esfera unitaria cerrada de  $\mathbb{R}^m$ . Entonces  $\Omega$  es scs.

PRUEBA. Si  $\Omega$  fuera un operador cerrado, esto es,  $G(\Omega)$  es cerrado en  $\mathbb{R}^m \times \mathbb{R}^n$ , entonces la semicontinuidad superior de  $\Omega$  sería consecuencia directa de la Proposición 4.14 ya que  $\Omega(\mathbb{R}^m \times \mathbb{R}^n) \subset \overline{B}_m(0, 1)$  es acotado. Esto significa que basta con mostrar que  $G(\Omega)$  es cerrado.

Sea  $(x_k, y_k, v_k)$  una sucesión en  $G(\Omega)$  que converge a  $(x, y, v)$ . Tenemos que  $v_k \in \Omega(x_k, y_k)$  para cada  $k$  y debemos probar que  $v \in \Omega(x, y)$ .

Si  $x \in \arg \min f^y$  entonces  $\Omega(x, y) = \overline{B}_m(0, 1)$ , y puesto que cada  $v_k$  está en  $\Omega(x_k, y_k) \subset \overline{B}_m(0, 1)$ , el límite  $v$  también está en  $\Omega(x, y)$ .

Si  $x \notin \arg \min f^y$  entonces  $\Omega(x, y) = \text{co}(N_{f^y}^a(x) \cap S_m(0, 1))$  y además existe  $x' \in \mathbb{R}^m$  tal que  $f(x', y) < f(x, y)$ . Como  $f$  es continua, de  $f(x', y_k) \rightarrow f(x', y)$  y  $f(x_k, y_k) \rightarrow f(x, y)$  se sigue que  $f(x', y_k) < f(x_k, y_k)$  para  $k$  suficientemente grande, es decir,  $x_k \notin \arg \min f^{y_k}$  para  $k$  grande. Luego, sin pérdida de generalidad podemos suponer que  $x_k \notin \arg \min f^{y_k}$  para todo  $k$  y tenemos

$$v_k \in \Omega(x_k, y_k) = \text{co}(N_{f^{y_k}}^a(x_k) \cap S_m(0, 1)), \quad \forall k \in \mathbb{N}.$$

Por el teorema de Carathéodory (Teorema 2.4), para cada  $k$  existen  $u_k^1, \dots, u_k^{m+1}$  en  $N_{f^{y_k}}^a(x_k) \cap S_m(0, 1)$  y  $\lambda_k^1, \dots, \lambda_k^{m+1}$  en  $[0, 1]$  tales que

$$v_k = \sum_{i=1}^{m+1} \lambda_k^i u_k^i \quad \text{y} \quad \sum_{i=1}^{m+1} \lambda_k^i = 1.$$

Como las sucesiones  $(u_k^i)_k$  y  $(\lambda_k^i)_k$  para  $i = 1, \dots, m+1$  están acotadas podemos suponer que convergen, digamos a  $u^i \in S_m(0, 1)$  y  $\lambda^i \in [0, 1]$  respectivamente. Tomando límite en la ecuación anterior deducimos que

$$v = \sum_{i=1}^{m+1} \lambda^i u^i \quad \text{y} \quad \sum_{i=1}^{m+1} \lambda^i = 1. \quad (8.3)$$

Veamos ahora que los vectores  $u^i$  están en  $N_{f^y}^a(x)$ . Dado  $(x', y') \in \mathbb{R}^m \times \mathbb{R}^n$  con  $x' \notin \arg \min f^{y'}$ , como  $f^{y'}$  es continua y cuasiconvexa semiestricta, de la Observación 5.9 se sigue que  $N_{f^{y'}}^a(x') = N_{f^{y'}}^<(x')$ , es decir,

$$N_{f^{y'}}^a(x') = \{u \in \mathbb{R}^m \mid \langle u, z - x' \rangle \leq 0, \forall z \in S_{f^{y'}}^<(x')\}. \quad (8.4)$$

En particular, esto se cumple cuando  $(x', y') = (x, y)$  y cuando  $(x', y') = (x_k, y_k)$ . Sea ahora  $z \in S_{f^y}^<(x)$ . Entonces  $f(z, y) < f(x, y)$  y de la continuidad de  $f$  se sigue

que  $f(z, y_k) < f(x_k, y_k)$  para  $k$  suficientemente grande. Así, podemos suponer sin pérdida de generalidad que

$$f(z, y_k) < f(x_k, y_k), \quad \forall k \in \mathbb{N},$$

es decir, que  $z$  está en cada conjunto  $S_{f^{y_k}}^<(x_k)$ . Además, cada  $u_k^i$  está en  $N_{f^{y_k}}^a(x_k)$ , de modo que

$$\langle u_k^i, z - x_k \rangle \leq 0.$$

Haciendo  $k \rightarrow \infty$  obtenemos  $\langle u^i, z - x \rangle \leq 0$ . Esto muestra que  $\langle u^i, z - x \rangle \leq 0$  para cualquier  $z \in S_{f^y}^<(x)$ , lo que en virtud de (8.4) equivale a

$$u^i \in N_{f^y}^a(x).$$

Así, todos los vectores  $u^i$  están en  $N_{f^y}^a(x) \cap S_m(0, 1)$  y de la ecuación (8.3) se sigue que  $v$  está en  $\text{co}(N_{f^y}^a(x) \cap S_m(0, 1)) = \Omega(x, y)$ .  $\square$

**Observación 8.11.** Si en la definición (8.2) del operador  $F_f$  cambiamos  $\overline{B}(0, 1)$  por  $\{0\}$ , es decir, si definimos  $F_f(x) = \{0\}$  cuando  $x \in \arg \min f$ , entonces no es difícil ver que las Propositiones 8.5 y 8.6 y los Teoremas 8.7 y 8.9 siguen siendo válidos. Sin embargo, esta restricción es necesaria pues al trabajar con los equilibrios de Nash en el Capítulo 3 necesitaremos mostrar que un operador similar al operador  $\Omega$  del Teorema 8.10, definido a partir de multifunciones del tipo  $F_f$ , es scs (véase el Teorema 13.6). Entonces será preciso usar el Teorema 8.10, que no vale cuando imponemos

$$\Omega(x, y) = F_{f^y}(x) = \{0\} \text{ si } x \in \arg \min f^y$$

pues la condición

$$\Omega(x, y) = \overline{B}_m(0, 1) \text{ si } x \in \arg \min f^y$$

fue usada en su prueba (en el tercer párrafo). En realidad, lo esencial de la definición de  $F_f$  cuando  $x \in \arg \min f$  es que  $F_f(x)$ , además de ser un conjunto convexo, compacto y no vacío, contiene tanto al vector 0 como a la imagen de  $F_f$ . Así, habríamos obtenido los mismos resultados si en vez de trabajar con  $\overline{B}(0, 1)$  hubiésemos trabajado con cualquier conjunto convexo, compacto y no vacío  $K$  que contenga a  $\overline{B}(0, 1)$ .

# Capítulo 2

## Equilibrios de Nash

El Problema de Equilibrio de Nash (Generalizado) es un modelo matemático que sirve para modelar situaciones de la vida real como la que hemos descrito en el segundo párrafo de la Introducción. La idea es la siguiente:  $N$  agentes o jugadores interactúan en un “juego”, intentando reducir sus pérdidas (o aumentar sus beneficios)<sup>1</sup> a través del uso de una “estrategia”. Como cada jugador tiene control sobre su estrategia propia (pero no sobre las de sus rivales), y como sus pérdidas (o beneficios) dependen de las estrategias de *todos* los agentes, a cada momento él trata de escoger una estrategia que, con respecto a las estrategias empleadas por sus rivales, le reporte las menores pérdidas (o los mayores beneficios). Es decir, cada jugador conoce las estrategias de sus rivales y, sobre la base de ese conocimiento, escoge la suya propia. Naturalmente, si un jugador modifica su estrategia, otros también podrían hacerlo, y luego otros, de modo que no es fácil que el juego llegue a un *equilibrio* (o sea, a un punto en que las estrategias empleadas se mantengan estables porque cada jugador está empleando su mejor estrategia respecto de las de sus rivales); incluso, tal equilibrio podría no existir.

El objetivo de este capítulo es presentar el Problema de Equilibrio de Nash Generalizado. Primero exploramos el Problema de Equilibrio de Nash, dando algunos ejemplos y probando el teorema de Nash, que garantiza la existencia de solución para el problema bajo ciertas circunstancias. Luego introducimos el problema generalizado, y establecemos la terminología y notación que utilizaremos en los siguientes capítulos para reformular dicho problema.

### 9. EL NEP

Sean  $n, n_1, \dots, n_N$  números naturales tales que  $n = n_1 + \dots + n_N$ , de forma que podamos escribir

$$\mathbb{R}^n = \mathbb{R}^{n_1} \oplus \dots \oplus \mathbb{R}^{n_N}. \quad (9.1)$$

Por simplicidad, escribimos también

$$\mathbb{R}^{n-n_\nu} := \mathbb{R}^{n_1} \oplus \dots \oplus \mathbb{R}^{n_{\nu-1}} \oplus \mathbb{R}^{n_{\nu+1}} \oplus \dots \oplus \mathbb{R}^{n_N}, \quad (\nu = 1, \dots, N).$$

---

<sup>1</sup>Véase la Observación 9.5

Para cada  $\nu \in \{1, \dots, N\}$  fijemos un conjunto no vacío

$$X_\nu \subset \mathbb{R}^{n_\nu}, \quad (9.2)$$

y una función

$$\theta_\nu : \mathbb{R}^n \rightarrow \mathbb{R}.$$

Dado un vector  $x \in \mathbb{R}^n$ , denotamos a su  $\nu$ -ésima componente en la suma directa (9.1) mediante el símbolo  $x^\nu$ , y a sus componentes restantes mediante el símbolo  $x^{-\nu}$ . Es decir,

$$x = (x^1, \dots, x^N) \quad \text{y} \quad x^{-\nu} = (x^1, \dots, x^{\nu-1}, x^{\nu+1}, \dots, x^N).$$

A fin de destacar  $x^\nu$  en  $x$ , a veces escribiremos  $(x^\nu, x^{-\nu})$  en lugar de  $x$ , no significando esto que las componentes de  $x$  hayan sido reordenadas<sup>2</sup>. Por extensión, los elementos de  $\mathbb{R}^{n_\nu}$  serán genéricamente denotados por  $x^\nu$  (sin que necesariamente exista un  $x \in \mathbb{R}^n$  cuya  $\nu$ -ésima componente sea  $x^\nu$ ), de la misma manera en que usamos  $x$  para denotar los elementos de  $\mathbb{R}^n$ . Igualmente, los elementos de  $\mathbb{R}^{n-n_\nu}$  serán denotados genéricamente por  $x^{-\nu}$ .

El **Problema de Equilibrio de Nash** (**NEP** por sus siglas en inglés) consiste en encontrar puntos  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N)$  en  $X_1 \times \dots \times X_N$  que satisfagan la siguiente propiedad

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu}) \quad \text{para cualesquiera } \nu \in \{1, \dots, N\} \text{ y } x^\nu \in X_\nu.$$

Estos puntos (si existen) son llamados **equilibrios (de Nash)** o **soluciones del NEP**. Naturalmente, si existen tales puntos decimos que el NEP **posee solución**; de lo contrario, decimos que **no posee solución**.

Debido al origen económico del NEP, cada uno de sus elementos posee una denominación sugestiva. Así,  $N$  agentes o jugadores interactúan en un *juego*, intentando reducir sus pérdidas a través del uso de una *estrategia*. El  $\nu$ -ésimo jugador controla las  $n_\nu$  variables

$$x^\nu \in \mathbb{R}^{n_\nu}$$

de  $\mathbb{R}^n$ , que son conocidas como sus **variables de decisión** o como su **estrategia** en el juego; la función

$$\theta_\nu : \mathbb{R}^n \rightarrow \mathbb{R},$$

cuyos valores dependen de *todas* las variables de decisión, mide sus **pérdidas** (o **beneficios**), y por ello es denominada su **función objetivo**, su **función de coste** o su **función de pérdidas**<sup>3</sup>; el conjunto

$$X_\nu \subset \mathbb{R}^{n_\nu},$$

<sup>2</sup>La igualdad  $x = (x^\nu, x^{-\nu})$  es un abuso de notación muy difundido en la literatura científica, cuya razón de ser responde a la necesidad de una notación concisa. En efecto, dados dos vectores  $y = (y^1, \dots, y^N)$  y  $z = (z^1, \dots, z^N)$  en  $\mathbb{R}^n$ , es más fácil escribir  $(z^\nu, y^{-\nu})$  en vez de  $(y^1, \dots, y^{\nu-1}, z^\nu, y^{\nu+1}, \dots, y^N)$ .

<sup>3</sup>Véase la Observación 9.5.

conocido como su **conjunto de estrategias (factibles)** o como su **región factible**, representa las estrategias que él puede utilizar durante el juego. Cada vector

$$x = (x^1, \dots, x^N) \in \mathbb{R}^n$$

nos da un conjunto de estrategias del juego, una por cada jugador, donde  $x^{-\nu}$  representa las estrategias de los rivales del  $\nu$ -ésimo jugador. Cuando cada estrategia  $x^\nu$  sea **factible**, esto es  $x^\nu \in X_\nu$ , diremos que

$$x = (x^1, \dots, x^N) \in X_1 \times \dots \times X_N$$

es un vector **factible**. Así, para un conjunto dado de estrategias  $x^{-\nu}$  de sus rivales, el  $\nu$ -ésimo jugador intenta encontrar una estrategia factible  $x^\nu \in X_\nu$  que minimice sus pérdidas, es decir, trata de resolver el problema de minimización<sup>4</sup>

$$\min_{x^\nu} \theta_\nu(x^\nu, x^{-\nu}), \quad x^\nu \in X_\nu. \quad (9.3)$$

Los *equilibrios* (como los definimos en el párrafo anterior) son justamente aquellos puntos factibles en que cada jugador ha logrado su objetivo (o sea, resolver su respectivo problema de optimización). Si  $\mathcal{S}_\nu(x^{-\nu}) \subset X_\nu$  representa al conjunto solución del problema (9.3) y

$$\mathcal{S}(x) = \mathcal{S}(x^1, \dots, x^N) := \mathcal{S}_1(x^{-1}) \times \dots \times \mathcal{S}_N(x^{-N}),$$

entonces  $\mathcal{S} : X_1 \times \dots \times X_N \rightrightarrows X_1 \times \dots \times X_N$  define un operador multivaluado cuyos puntos fijos son precisamente los equilibrios del NEP.

Veamos un par de ejemplos simples que ilustren estos conceptos.

**Ejemplo 9.1.** Consideremos  $N = 2$  jugadores interactuando en  $\mathbb{R}^4$  con  $n_1 = n_2 = 2$ , cuyos problemas de minimización asociados son los siguientes

$$\begin{array}{ll} \min_{x^1} x_1 + x_2 + x_1 x_3 & \min_{x^2} (x_3 - 1)^2 + x_1(x_4 - 1)^2 \\ |x_1| \leq 1 & |x_3| \leq 1 \\ |x_2| \leq 1 & |x_4| \leq 1 \end{array}$$

Aquí las variables de decisión son

$$x^1 = (x_1, x_2) \quad \text{y} \quad x^2 = (x_3, x_4),$$

de modo que cada vector  $x$  de  $\mathbb{R}^4$  se escribe  $x = (x^1, x^2)$ ; en particular,  $x^{-1} = x^2$  y  $x^{-2} = x^1$ . Las funciones de pérdidas son

$$\theta_1(x^1, x^{-1}) = x_1 + x_2 + x_1 x_3 \quad \text{y} \quad \theta_2(x^2, x^{-2}) = (x_3 - 1)^2 + x_1(x_4 - 1)^2,$$

y los conjuntos de estrategias factibles son

$$X_1 = X_2 = [-1, 1] \times [-1, 1].$$

---

<sup>4</sup>Véase la Observación 9.5.

Dada una estrategia factible  $x^{-1} = x^2 = (x_3, x_4) \in X_2$  del segundo jugador, como  $0 \leq x_3 + 1$  el problema de optimización asociado al primer jugador

$$\min_{x^1} \theta_1(x^1, x^{-1}), \quad x^1 \in X_1$$

tiene una única solución en el punto  $x^1 = (x_1, x_2) = (-1, -1)$ . Similarmente, dada una estrategia factible  $x^{-2} = x^1 = (x_1, x_2) \in X_1$  del primer jugador, el problema de optimización del segundo jugador

$$\min_{x^2} \theta_2(x^2, x^{-2}), \quad x^2 \in X_2$$

posee el siguiente conjunto solución:

$$\mathcal{S}_2(x^{-2}) = \begin{cases} \{(1, 1)\}, & \text{si } x_1 > 0, \\ \{(1, 0)\}, & \text{si } x_1 < 0, \\ \{1\} \times [-1, 1], & \text{si } x_1 = 0. \end{cases}$$

Esto implica que

$$\bar{x} = (\bar{x}^1, \bar{x}^2) := (-1, -1, 1, 0)$$

es el único equilibrio de este NEP.

**Ejemplo 9.2.** Modificamos levemente el ejemplo anterior. El NEP

$$\begin{array}{ll} \min_{x^1} x_1 + x_2 + x_1 x_3 & \min_{x^2} (x_3 - 1)^2 + x_1(x_4 - 1)^2 \\ x_1 \leq 0 & |x_3| \leq 1 \\ x_2 \leq 0 & |x_4| \leq 1 \end{array}$$

con funciones objetivo

$$\theta_1(x^1, x^{-1}) = x_1 + x_2 + x_1 x_3 \quad \text{y} \quad \theta_2(x^2, x^{-2}) = (x_3 - 1)^2 + x_1(x_4 - 1)^2,$$

y con conjuntos de estrategias factibles

$$X_1 = (-\infty, 0] \times (-\infty, 0] \quad \text{y} \quad X_2 = [-1, 1] \times [-1, 1]$$

no posee solución, ya que el problema del primer jugador

$$\min_{x^1} x_1 + x_2 + x_1 x_3, \quad x^1 \in X_1$$

no posee solución para ninguna estrategia  $x^2$  del segundo jugador.

Veamos otro ejemplo no tan simple, pero interesante.

**Ejemplo 9.3** (*Zero-sum game of two players*). Consideremos un juego de suma cero (*zero-sum game*) con  $N = 2$  participantes. Esto quiere decir que las funciones  $\theta_1$  y  $\theta_2$  del NEP suman cero, o sea

$$\theta_1(x^1, x^2) = -\theta_2(x^1, x^2) \quad \text{para todo } (x^1, x^2).$$

Supongamos que las regiones factibles  $X_\nu \subset \mathbb{R}^{n_\nu}$  son compactas. Supongamos además que las funciones  $\theta_\nu$  son continuas y que para cada  $\nu$  y  $x^{-\nu}$  fijos la aplicación

$$x^\nu \mapsto \theta_\nu(x^\nu, x^{-\nu})$$

es convexa. Entonces la función  $\theta_1 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \rightarrow \mathbb{R}$  tiene la propiedad de que  $x^1 \mapsto \theta_1(x^1, x^2)$  es convexa para cada  $x^2 \in \mathbb{R}^{n_2}$  fijo y de que  $x^2 \mapsto \theta_1(x^1, x^2)$  es cóncava para cada  $x^1 \in \mathbb{R}^{n_1}$  fijo. En virtud del teorema minimax de von Neumann (Teorema 4.22), existe  $\bar{x} := (\bar{x}^1, \bar{x}^2) \in X_1 \times X_2$  tal que

$$\min_{x^1 \in X_1} \theta_1(x^1, \bar{x}^2) = \theta_1(\bar{x}^1, \bar{x}^2) = \max_{x^2 \in X_2} \theta_1(\bar{x}^1, x^2),$$

es decir,

$$\min_{x^1 \in X_1} \theta_1(x^1, \bar{x}^2) = \theta_1(\bar{x}^1, \bar{x}^2) \quad \text{y} \quad \min_{x^2 \in X_2} \theta_2(\bar{x}^1, x^2) = \theta_2(\bar{x}^1, \bar{x}^2),$$

lo cual significa que  $\bar{x}$  es un equilibrio de Nash. Así, este NEP admite soluciones<sup>5</sup>.

John Nash introdujo el NEP en su artículo [26], donde también probó que el NEP admite soluciones bajo ciertas hipótesis. La prueba original de este resultado, mostrado en su tesis de doctorado, hacía uso del teorema de punto fijo de Brouwer. No obstante, en [26] John Nash simplificó dicha prueba original al utilizar el teorema del punto fijo de Kakutani (en lugar del teorema de Brouwer). Este resultado es importante porque nos dice que el NEP admite soluciones bajo hipótesis naturales, y que por tanto vale la pena estudiarlo.

**Teorema 9.4** (Nash, [26]). *Supongamos que las funciones  $\theta_\nu$  son continuas en  $\mathbb{R}^n$  y que cada aplicación*

$$x^\nu \mapsto \theta_\nu(x^\nu, x^{-\nu})$$

*es convexa para cualquier  $x^{-\nu}$  fijo. Supongamos también que los conjuntos  $X_\nu \subset \mathbb{R}^{n_\nu}$  son convexos, compactos y no vacíos. Entonces el NEP posee solución.*

PRUEBA. Usando el teorema de Kakutani (Teorema 4.20) probaremos que la multifunción  $\mathcal{S}$  posee algún punto fijo. Sea  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N)$  un elemento de  $X_1 \times \dots \times X_N$ . Dado  $\nu \in \{1, \dots, N\}$ , la aplicación

$$x^\nu \mapsto \theta_\nu(x^\nu, \bar{x}^{-\nu})$$

es continua y por tanto alcanza su mínimo, digamos  $\lambda \in \mathbb{R}$ , sobre el conjunto compacto  $X_\nu$ , o sea  $\mathcal{S}_\nu(\bar{x}^{-\nu})$  es no vacío; como además  $\mathcal{S}_\nu(\bar{x}^{-\nu}) = X_\nu \cap S_\lambda$ , donde

---

<sup>5</sup> La interpretación práctica de un juego de suma cero es la siguiente. Supongamos que existe una economía en la que solo participan dos agentes, que interactúan como en el NEP para obtener la mayor ganancia posible. Como solo hay dos participantes en la economía, el dinero que uno gane lo habrá perdido el otro, y viceversa, lo cual quiere decir que la suma de las funciones de beneficio de ambos agentes (las cuales dependen de sus estrategias) es igual a cero. Lo interesante del teorema minimax de von Neumann (Teorema 4.22) es que nos dice que todo juego de suma cero posee un punto de equilibrio.

$\mathcal{S}_\lambda$  es el subnivel (convexo) de dicha aplicación convexa, concluimos que  $\mathcal{S}_\nu(\bar{x}^{-\nu})$  es también convexo. Puesto que

$$\mathcal{S}(\bar{x}) = \mathcal{S}_1(\bar{x}^{-1}) \times \cdots \times \mathcal{S}_N(\bar{x}^{-N}),$$

esto muestra que  $\mathcal{S}$  tiene imágenes convexas y no vacías.

Por otra parte, la continuidad de las funciones  $\theta_\nu$  nos garantiza que el gráfico  $G(\mathcal{S})$  de  $\mathcal{S}$  es cerrado: si  $(x_k, y_k)$  es una sucesión en  $G(\mathcal{S})$  que converge a  $(\bar{x}, \bar{y})$ , entonces  $y_k^\nu \in \mathcal{S}_\nu(x_k^{-\nu})$  para cualesquiera  $k$  y  $\nu$ , es decir,

$$\theta_\nu(y_k^\nu, x_k^{-\nu}) \leq \theta_\nu(x^\nu, x_k^{-\nu}) \text{ para cualesquiera } k, \nu \text{ y } x^\nu \in X_\nu,$$

de donde, haciendo  $k \rightarrow \infty$  para  $\nu$  y  $x^\nu$  fijos, obtenemos

$$\theta_\nu(\bar{y}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu}) \text{ para cualesquiera } \nu \text{ y } x^\nu \in X_\nu,$$

o sea  $\bar{y}^\nu \in \mathcal{S}_\nu(\bar{x}^{-\nu})$  para todo  $\nu$ , lo cual significa  $(\bar{x}, \bar{y}) \in G(\mathcal{S})$ . Vemos por tanto que  $\mathcal{S}$  verifica las hipótesis del Teorema 4.20, completando así la prueba del teorema.  $\square$

Teniendo en cuenta el teorema anterior<sup>6</sup> podemos ver que una de las razones por las que en nuestro primer ejemplo existe solución, y en el segundo no, es que el conjunto  $X := X_1 \times X_2$  es compacto en el primer ejemplo, pero no en el segundo.

**Observación 9.5.** Como mencionamos en el inicio de este capítulo, la idea básica del NEP es que los jugadores desean *minimizar* sus *pérdidas*, o *maximizar* sus *beneficios*. Nosotros hemos convenido en trabajar con *pérdidas*, lo cual, matemáticamente, significa que el problema de optimización (9.3) que enfrenta cada jugador es un problema de *minimización*. Podríamos haber trabajado con *beneficios*. En ese caso, las funciones objetivo  $\theta_\nu$  serían llamadas **funciones de beneficio** o **funciones de utilidad**, el problema (9.3) sería un problema de *maximización*, y los puntos de equilibrio serían aquellos vectores

$$\bar{x} = (\bar{x}^1, \dots, \bar{x}^N) \in X_1 \times \cdots \times X_N$$

tales que

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \geq \theta_\nu(x^\nu, \bar{x}^{-\nu}) \text{ para cualesquiera } \nu \in \{1, \dots, N\} \text{ y } x^\nu \in X_\nu.$$

## 10. EL GNEP

Volvamos al ejemplo de la Introducción. Si el equipo A ataca, obviamente el equipo B no puede ponerse a atacar, pues ni siquiera tiene el balón; más bien va a defender o, si fuera posible, preparar un contraataque. Esto ilustra un hecho importante: en los “juegos” de la vida real el conjunto de las estrategias que un jugador puede usar en un momento dado va a *depender* de las estrategias de sus rivales. Matemáticamente, esto significa que el conjunto  $X_\nu \subset \mathbb{R}^{n_\nu}$  del NEP no permanece fijo, sino que depende de  $x^{-\nu}$ . La versión generalizada del NEP incorpora esta condición.

<sup>6</sup>Véanse también los resultados de existencia de solución del siguiente capítulo.

Así, en vez de los conjuntos  $X_\nu \subset \mathbb{R}^{n_\nu}$  consideramos las *multifunciones*

$$X_\nu : \mathbb{R}^{n-n_\nu} \rightrightarrows \mathbb{R}^{n_\nu}. \quad (10.1)$$

Entonces, el problema de minimización que enfrenta el  $\nu$ -ésimo jugador pasa a ser

$$\min_{x^\nu} \theta_\nu(x^\nu, x^{-\nu}), \quad x^\nu \in X_\nu(x^{-\nu}), \quad (10.2)$$

donde  $x^{-\nu}$  es un conjunto dado de estrategias de sus rivales. Hablando más sugestivamente, a cada momento el  $\nu$ -ésimo jugador trata de minimizar sus pérdidas escogiendo una estrategia adecuada  $x^\nu$  de entre un conjunto posible de estrategias  $X_\nu(x^{-\nu})$ , el cual depende de las estrategias  $x^{-\nu}$  empleadas por sus rivales.

El **Problema de Equilibrio de Nash Generalizado** (GNEP por sus siglas en inglés) consiste en encontrar los *equilibrios* del problema en cuestión, o sea los puntos en que cada jugador ha resuelto su respectivo problema de optimización. Explícitamente, el GNEP consiste en hallar vectores  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N)$  tales que  $\bar{x}^\nu \in X_\nu(\bar{x}^{-\nu})$  y

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu}) \text{ para cualesquiera } \nu \in \{1, \dots, N\} \text{ y } x^\nu \in X_\nu(\bar{x}^{-\nu}).$$

Estos puntos (si existen) son llamados **equilibrios de Nash Generalizado** o **soluciones del GNEP**. Naturalmente, si existen tales puntos decimos que el GNEP **posee solución**; de lo contrario, decimos que **no posee solución**.

Si  $\mathcal{S}_\nu(x^{-\nu}) \subset X_\nu(x^{-\nu})$  representa el conjunto solución del problema (10.2) y

$$\mathcal{S}(x) = \mathcal{S}(x^1, \dots, x^N) := \mathcal{S}_1(x^{-1}) \times \dots \times \mathcal{S}_N(x^{-N}),$$

entonces  $\mathcal{S} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  define una multifunción cuyos puntos fijos  $\bar{x} \in \mathcal{S}(\bar{x})$  son exactamente las soluciones del GNEP.

**Observación 10.1** (El GNEP generaliza el NEP). Cuando las multifunciones  $X_\nu$  introducidas en (10.1) son constantes, esto es, cuando existen conjuntos  $X_\nu \subset \mathbb{R}^{n_\nu}$  tales que  $X_\nu(x^{-\nu}) = X_\nu$  para todo  $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$ , el GNEP puede verse como un NEP. (El doble uso del símbolo  $X_\nu$  no debería causar confusión.) Así, los NEPs constituyen una subclase de los GNEPs<sup>7</sup>.

Un conjunto de estrategias  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N) \in \mathbb{R}^n$  es llamado **factible** si puede ocurrir en el juego, es decir, si cada estrategia  $\bar{x}^\nu$  es **factible** respecto de las estrategias rivales  $\bar{x}^{-\nu}$ , o sea  $\bar{x}^\nu \in X_\nu(\bar{x}^{-\nu})$ . Los puntos factibles del GNEP son exactamente los puntos fijos  $\bar{x} \in \Omega(\bar{x})$  de la multifunción  $\Omega : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  dada por

$$\Omega(x) = \Omega(x^1, \dots, x^N) = X_1(x^{-1}) \times \dots \times X_N(x^{-N}).$$

Claramente,

$$\bar{x} \in \mathcal{S}(\bar{x}) \text{ implica } \bar{x} \in \Omega(\bar{x}),$$

lo cual expresa el hecho de que toda solución del GNEP es necesariamente un vector factible.

Nótese que si el GNEP es un NEP entonces esta noción de factibilidad coincide con la previamente introducida en el contexto de los NEPs.

<sup>7</sup>No todos los GNEPs que pueden verse como NEPs tiene multifunciones asociadas  $X_\nu$  constantes (véase la Observación 10.4).

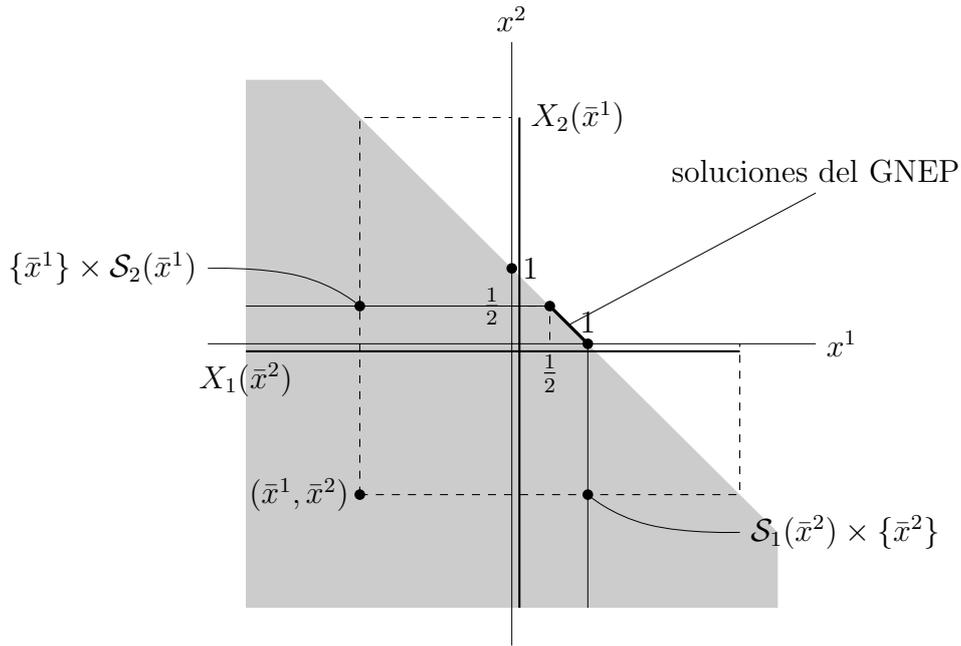


Figura 10.1: Gráfico del Ejemplo 10.2. La región sombreada  $\{(x^1, x^2) \mid x^1 + x^2 \leq 1\}$  representa al conjunto de vectores factibles. Para cada punto factible  $\bar{x} = (\bar{x}^1, \bar{x}^2)$  tenemos  $\mathcal{S}(\bar{x}) = \{(\min\{1, 1 - \bar{x}^2\}, \min\{\frac{1}{2}, 1 - \bar{x}^1\})\}$ ; además,  $\bar{x} \in \mathcal{S}(\bar{x})$  si y solo si  $\bar{x}$  está en el segmento que une  $(\frac{1}{2}, \frac{1}{2})$  con  $(1, 0)$ .

**Ejemplo 10.2** ([12, Example 1.1]). Consideremos en  $\mathbb{R}^2$  un juego con  $N = 2$  participantes, cada uno controlando una variable (o sea  $n_1 = n_2 = 1$ ). Por simplicidad, usamos el símbolo  $x^1$  para denotar a la variable  $x_1$ , que es la única variable de  $\mathbb{R}^2$  controlada por el primer jugador; hacemos lo mismo con el segundo jugador. De esta manera, los vectores de  $\mathbb{R}^2$  se escriben  $x = (x^1, x^2)$ , y además

$$x^{-1} = x^2 \quad \text{y} \quad x^{-2} = x^1.$$

Definimos los problemas de optimización de nuestro GNEP mediante

$$\begin{array}{ll} \min_{x^1} (x^1 - 1)^2 & \min_{x^2} (x^2 - \frac{1}{2})^2 \\ x^1 + x^2 \leq 1 & x^1 + x^2 \leq 1 \end{array}$$

Las funciones de pérdidas de este GNEP son

$$\theta_1(x^1, x^2) = (x^1 - 1)^2 \quad \text{y} \quad \theta_2(x^1, x^2) = (x^2 - \frac{1}{2})^2.$$

Dada una estrategia  $x^{-1} = x^2$  del segundo jugador, el conjunto de estrategias factibles del primer jugador es

$$X_1(x^{-1}) = X_1(x^2) = (-\infty, 1 - x^2],$$

de modo que su problema de optimización asociado

$$\min_{x^1} \theta_1(x^1, x^{-1}), \quad x^1 \in X_1(x^{-1})$$

posee una única solución en  $x^1 = \min\{1, 1 - x^2\}$ . Similarmente, dada una estrategia  $x^{-2} = x^1$  del primer jugador, el conjunto de estrategias factibles del segundo jugador es

$$X_2(x^{-2}) = X_2(x^1) = (-\infty, 1 - x^1],$$

de forma que su problema de optimización asociado

$$\min_{x^2} \theta_2(x^2, x^{-2}), \quad x^2 \in X_2(x^{-2})$$

tiene una única solución en  $x^2 = \min\{\frac{1}{2}, 1 - x^1\}$ . En consecuencia,  $\bar{x} = (\bar{x}^1, \bar{x}^2)$  es un punto de equilibrio de Nash generalizado si y solo si  $\bar{x}^1 = \min\{1, 1 - \bar{x}^2\}$  y  $\bar{x}^2 = \min\{\frac{1}{2}, 1 - \bar{x}^1\}$ , es decir, si y solo si  $\bar{x}^2 = 1 - \bar{x}^1$  y  $\frac{1}{2} \leq \bar{x}^1 \leq 1$ . Por tanto, el conjunto de soluciones

$$\{(t, 1 - t) \in \mathbb{R}^2 \mid \frac{1}{2} \leq t \leq 1\}$$

de este GNEP es infinito.

Antes de continuar introducimos un poco de terminología. Dados  $\nu \in \{1, \dots, N\}$  y  $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$ , el símbolo  $\theta_\nu(\cdot, x^{-\nu})$  representará la aplicación

$$x^\nu \in \mathbb{R}^{n_\nu} \mapsto \theta_\nu(x^\nu, x^{-\nu}) \in \mathbb{R}.$$

Una afirmación del tipo “ $\theta_\nu$  es continua (respectivamente convexa, diferenciable, etc.) respecto a la  $\nu$ -ésima variable” significará que  $\theta_\nu(\cdot, x^{-\nu})$  es continua (respectivamente convexa, diferenciable, etc.) para todo  $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$ .

La principal pregunta que surge al analizar un GNEP dado consiste en decidir si este admite soluciones. La respuesta a esta cuestión está sujeta a las características del GNEP considerado. Como regla general, mientras más débiles sean las hipótesis bajo las que trabajemos, menos podremos afirmar sobre nuestro problema. Por ello es usual considerar GNEPs que satisfagan hipótesis “razonables”. Una primera hipótesis razonable es la siguiente hipótesis

**H1. (Hipótesis de continuidad)** Cada  $\theta_\nu$  es continua en  $\mathbb{R}^{n_\nu}$ .

Otras hipótesis que comúnmente se toman en cuenta son las siguientes.

**H2. (Hipótesis de diferenciableidad)** Cada  $\theta_\nu$  es de clase  $C^1$  en  $\mathbb{R}^{n_\nu}$ .

**H3. (Hipótesis de convexidad)** Cada  $\theta_\nu$  es convexa respecto a la  $\nu$ -ésima variable.

**H4. (Hipótesis de cuasiconvexidad)** Cada  $\theta_\nu$  es cuasiconvexa respecto a la  $\nu$ -ésima variable.

Los GNEPs que verifican la hipótesis de diferenciableidad **H2.**, que es más fuerte que la hipótesis de continuidad **H1.**, pueden ser más fáciles de analizar debido a que en este caso las funciones objetivo de los problemas de minimización (10.2) asociados al GNEP son diferenciables. En las Secciones 11, 15 y 16 estudiaremos más a fondo este tipo de GNEPs.

Por otra parte, las hipótesis **H3.** y **H4.** son motivadas por el hecho de que el GNEP posee un origen económico. Más precisamente, como en economía surgen

naturalmente funciones convexas, o, más generalmente, funciones cuasiconvexas, los GNEPs que satisfacen estas hipótesis han sido históricamente importantes. Por ejemplo, cuando John Nash definió el NEP en [26] y probó su teorema de existencia de solución (Teorema 9.4), trabajaba bajo las hipótesis de convexidad **H3.** y continuidad **H1.** La hipótesis de cuasiconvexidad **H4.** es la generalización natural de la hipótesis de convexidad **H3.**

Cuando se trabaja bajo la hipótesis de convexidad (o cuasiconvexidad) es común requerir que los conjuntos factibles de los problemas (10.2) sean convexas, o sea que

$$\text{cada } X_\nu(x^{-\nu}) \text{ sea convexo.}$$

Los NEPs constituyen la subclase más importante de los GNEPs. En [29], J.B. Rosen introdujo una subclase de GNEPs que, por un lado, contiene a los NEPs, y que, por el otro, es lo suficientemente restrictiva como para que tenga propiedades interesantes (véase el Teorema 10.6). La hipótesis que Rosen introdujo es la siguiente.

**HR. (Hipótesis de Rosen)** Existe un conjunto no vacío y convexo (posiblemente cerrado o compacto)  $X \subset \mathbb{R}^n$  tal que cada  $X_\nu(x^{-\nu})$  está dado por

$$X_\nu(x^{-\nu}) = \{x^\nu \in \mathbb{R}^n \mid (x^\nu, x^{-\nu}) \in X\}. \quad (10.3)$$

Claramente, cuando esta hipótesis se satisface se tiene que

$$\text{cada } X_\nu(x^{-\nu}) \text{ es convexo;}$$

además, si el conjunto  $X$  es cerrado (compacto) entonces

$$\text{cada } X_\nu(x^{-\nu}) \text{ también es cerrado (compacto).}$$

Otra ventaja es que en estos casos los vectores factibles del GNEP, o sea los puntos  $\bar{x} \in \mathbb{R}^n$  tales que  $\bar{x}^\nu \in X_\nu(\bar{x}^{-\nu})$  para todo  $\nu$ , o equivalentemente  $\bar{x} \in \Omega(\bar{x})$ , son precisamente los elementos del conjunto  $X$  (véase el Lema 10.3). Una consecuencia inmediata de este hecho es que la definición de equilibrio de Nash generalizado se simplifica levemente: un vector  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N)$  es solución del GNEP si y solo si  $\bar{x} \in X$  y

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu}) \text{ para cualesquiera } \nu \in \{1, \dots, N\} \text{ y } x^\nu \in X_\nu(\bar{x}^{-\nu}).$$

**Lema 10.3.** *Supongamos que existe un conjunto  $X \subset \mathbb{R}^n$  tal que la hipótesis de Rosen es satisfecha. Entonces para todo  $x \in \mathbb{R}^n$  se cumple que  $x \in \Omega(x)$  si y solo si  $x \in X$ .*

**PRUEBA.** Si  $x \in \Omega(x)$ , entonces  $x^\nu \in X_\nu(x^{-\nu})$  para todo  $\nu \in \{1, \dots, N\}$ , esto es,  $x = (x^\nu, x^{-\nu}) \in X$ . Recíprocamente, si  $x \in X$ , entonces es claro que cada  $x^\nu$  está en  $X_\nu(x^{-\nu})$  pues  $(x^\nu, x^{-\nu}) = x \in X$ , y en consecuencia  $x \in \Omega(x)$ .  $\square$

**Observación 10.4.** Un GNEP que cumple la hipótesis de Rosen es un NEP si y solo si su conjunto asociado  $X$  se escribe de la forma

$$X = X_1 \times \dots \times X_N, \quad (10.4)$$

donde cada conjunto  $X_\nu \subset \mathbb{R}^{n_\nu}$  es convexo y no vacío. Cuando esto sucede las multifunciones  $X_\nu$  están dadas por

$$X_\nu(x^{-\nu}) = \begin{cases} X_\nu, & \text{si existe } x^\nu \in \mathbb{R}^{n_\nu} \text{ tal que } (x^\nu, x^{-\nu}) \in X, \\ \emptyset, & \text{si } (x^\nu, x^{-\nu}) \notin X \text{ para todo } x^\nu \in \mathbb{R}^{n_\nu}. \end{cases}$$

Si vemos el GNEP desde el punto de vista de los beneficios o utilidades (véase la Observación 9.5), el conjunto  $X$  de la hipótesis de Rosen tiene una interpretación interesante. Solo para ilustrar este punto supongamos que  $X$  tiene la forma

$$X = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_1^2 + \dots + x_n^2 \leq M \text{ y } 0 \leq x_i \text{ para todo } i\},$$

donde  $M > 0$ . Entonces podemos pensar en  $X$  como la totalidad de recursos conjuntos de que disponen todos los jugadores: mientras más recursos utilice uno de ellos para incrementar sus utilidades, habrá menos recursos disponibles que los demás jugadores puedan emplear.

El GNEP del Ejemplo 10.2 satisface todas las hipótesis que hemos introducido en esta sección. En este caso el conjunto  $X$  está dado por

$$X = \{(x^1, x^2) \in \mathbb{R}^2 \mid x^1 + x^2 \leq 1\}.$$

**Ejemplo 10.5.** En  $\mathbb{R}^2$  consideremos el siguiente juego de  $N = 2$  participantes

$$\begin{array}{ll} \min_{x^1} x^1 + x^2 - 1, & \min_{x^2} x^1 x^2, \\ x^1 \geq 0 & x^1 \geq 0 \\ x^2 \geq (x^1)^2 & x^2 \geq (x^1)^2 \end{array}$$

en el que el jugador 1 controla la variable  $x^1$  y el jugador 2 controla la variable  $x^2$ . Este GNEP satisface la hipótesis de Rosen. Su conjunto asociado  $X$  es

$$X = \{(x^1, x^2) \in \mathbb{R}^2 \mid x^1 \geq 0, x^2 \geq (x^1)^2\}.$$

Dada una estrategia  $x^2 \geq 0$  del segundo jugador, la región factible del primer jugador es

$$X_1(x^2) = X_1(x^2) = \{x^1 \in \mathbb{R} \mid (x^1, x^2) \in X\} = [0, \sqrt{x^2}],$$

por lo que el conjunto solución de su problema de optimización asociado es

$$\mathcal{S}_1(x^2) = \{0\}.$$

De manera semejante, dada una estrategia  $x^1 \geq 0$  del primer jugador, la región factible del segundo jugador es

$$X_2(x^1) = X_2(x^1) = \{x^2 \in \mathbb{R} \mid (x^1, x^2) \in X\} = [(x^1)^2, \infty),$$

de forma que

$$\mathcal{S}_2(x^1) = \begin{cases} [0, \infty), & \text{si } x^1 = 0, \\ \{(x^1)^2\}, & \text{si } x^1 > 0. \end{cases}$$

Luego, un vector factible  $\bar{x} = (\bar{x}^1, \bar{x}^2) \in X$  es una solución del GNEP si y solo si  $\bar{x}^1 = 0$  y  $\bar{x}^2 \geq 0$ , o sea que  $\{0\} \times [0, \infty[$  representa al conjunto de equilibrios de Nash generalizado.

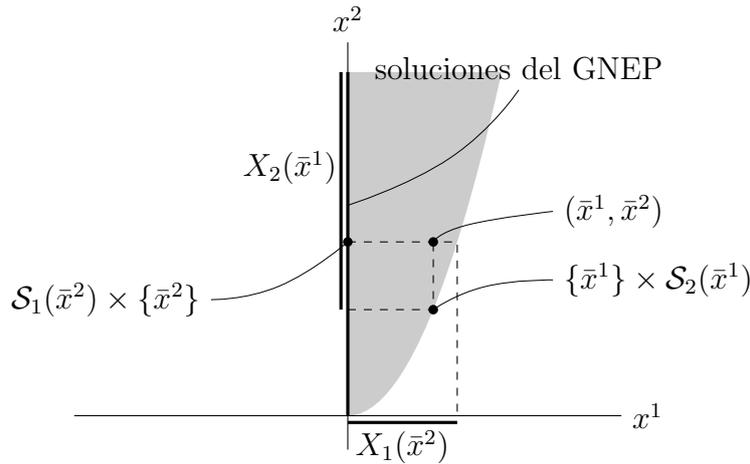


Figura 10.2: Gráfico del Ejemplo 10.5. La región sombreada representa a  $X$ . Para cada punto factible  $\bar{x} = (\bar{x}^1, \bar{x}^2) \in X$  tenemos  $\mathcal{S}(\bar{x}) = \{0\} \times [0, \infty)$  si  $x^1 = 0$  y  $\mathcal{S}(\bar{x}) = \{0\} \times \{(x^1)^2\}$  si  $x^1 > 0$ ; además,  $\bar{x} \in \mathcal{S}(\bar{x})$  si y solo si  $\bar{x}^1 = 0$ .

Los GNEPs que verifican la hipótesis de Rosen fueron estudiados en detalle en [29], donde, entre otras cosas, Rosen generalizó el teorema de existencia de solución de Nash. Debido a la importancia que estos GNEPs han tenido en las últimas décadas, a veces se les ha confundido con la versión general del GNEP.

**Teorema 10.6** (Rosen, [29, Theorem 1]). *Supongamos que se satisface la hipótesis de Rosen y que además se verifican las hipótesis de continuidad y convexidad **H1.** y **H3.** Supongamos también que el conjunto  $X$  asociado es compacto. Entonces el GNEP admite solución.*

PRUEBA. Sea la función  $\rho : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  definida por

$$\rho(x, y) = \sum_{\nu=1}^N \theta_{\nu}(y^{\nu}, x^{-\nu}).$$

Esta función es continua, y tiene la propiedad de que para cada  $x \in \mathbb{R}^n$  fijo la aplicación

$$\rho(x, \cdot) : y \mapsto \rho(x, y)$$

es convexa y continua, por lo que el conjunto

$$\Gamma(x) := \{y \in X \mid \rho(x, y) = \min_{z \in X} \rho(x, z)\} \subset X$$

es convexo y no vacío (pues  $X$  es compacto) . Esto nos da un operador  $\Gamma : X \rightrightarrows X$  cuyas imágenes son convexas y no vacías.

Afirmamos que el gráfico  $G(\Gamma)$  de  $\Gamma$  es cerrado, lo que será una fácil consecuencia de la continuidad de  $\rho$ . De hecho, si  $(x_k, y_k)$  es una sucesión en  $G(\Gamma)$  que converge a  $(\bar{x}, \bar{y})$ , entonces  $\bar{x}, \bar{y} \in X$  y

$$\rho(x_k, y_k) \leq \rho(x_k, z) \text{ para todo } z \in X,$$

de donde, tomando límite cuando  $k \rightarrow \infty$  para  $z \in X$  fijo, deducimos que

$$\rho(\bar{x}, \bar{y}) \leq \rho(\bar{x}, z) \text{ para todo } z \in X,$$

es decir,  $(\bar{x}, \bar{y}) \in G(\Gamma)$ . Así,  $\Gamma$  satisface las hipótesis del teorema de Kakutani (Teorema 4.20), y en consecuencia posee un punto fijo  $\bar{x} \in \Gamma(\bar{x})$ .

Veamos finalmente que  $\bar{x} \in X$  es una solución del GNEP. Dados  $\nu \in \{1, \dots, N\}$  y  $x^\nu \in X_\nu(\bar{x}^{-\nu})$ , tenemos que verificar que  $\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \theta_\nu(x^\nu, \bar{x}^{-\nu}) \leq 0$ . El vector

$$y := (x^\nu, \bar{x}^{-\nu}) = (\bar{x}^1, \dots, \bar{x}^{\nu-1}, x^\nu, \bar{x}^{\nu+1}, \dots, \bar{x}^N)$$

pertenece a  $X$  y tiene la propiedad de que  $y^\mu = \bar{x}^\mu$  si  $\mu \neq \nu$  y  $y^\nu = x^\nu$ , de manera que

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \theta_\nu(x^\nu, \bar{x}^{-\nu}) = \sum_{\mu=1}^N (\theta_\mu(\bar{x}^\mu, \bar{x}^{-\mu}) - \theta_\mu(y^\mu, \bar{x}^{-\mu})) = \rho(\bar{x}, \bar{x}) - \rho(\bar{x}, y) \leq 0,$$

donde la última desigualdad se debe a que  $\bar{x} \in \Gamma(\bar{x})$ . Esto completa la prueba del teorema.  $\square$

Los puntos fijos  $\bar{x} \in \Gamma(\bar{x})$  del operador  $\Gamma$  definido en la prueba del teorema anterior son llamados **equilibrios de Nash normalizados** (véase la Sección 15). Acabamos de ver que todo equilibrio de Nash normalizado es un equilibrio de Nash generalizado.

**Observación 10.7.** El teorema de Nash sigue siendo válido si reemplazamos la hipótesis de convexidad **H3.** por la hipótesis de cuasiconvexidad **H4.**, como puede verse fácilmente de su demostración. No acontece lo mismo con el teorema de Rosen pues la hipótesis **H4.** no implica necesariamente que las funciones  $\rho(x, \cdot)$  sean cuasiconvexas.

El teorema de Rosen fue uno de los primeros resultados que garantizaban que ciertos GNEPs poseían solución (el primero fue dado por Debreu [10]). Existen muchos otros resultados de este tipo, el más importante de los cuales tal vez sea el siguiente teorema, aplicable a GNEPs que no necesariamente verifican la hipótesis de Rosen.

**Teorema 10.8** ([21]). *Supongamos que se cumplen las hipótesis de continuidad y cuasiconvexidad **H1.** y **H4.** (pero no necesariamente la hipótesis de Rosen). Supongamos también que existen  $N$  conjuntos convexos, compactos y no vacíos  $K_\nu \subset \mathbb{R}^{n_\nu}$  tales que para cada  $x \in K_1 \times \dots \times K_N$  el conjunto*

$$\Omega(x) = X_1(x^{-1}) \times \dots \times X_N(x^{-N})$$

*está contenido en  $K_1 \times \dots \times K_N$  y es convexo, compacto y no vacío. Si cada multifunción*

$$X_\nu : K_1 \times \dots \times K_{\nu-1} \times K_{\nu+1} \times \dots \times K_N \rightrightarrows \mathbb{R}^{n_\nu}$$

*es sci y scs entonces el GNEP posee solución.*

Antes de comenzar la prueba observemos que la hipótesis del teorema sobre  $\Omega$  significa que si  $x \in K_1 \times \cdots \times K_N$ , entonces cada  $X_\nu(x^{-\nu})$  está contenido en  $K_\nu$  y es convexo, compacto y no vacío. En particular, cada multifunción  $X_\nu : K_{-\nu} \rightrightarrows K_\nu$ , donde

$$K_{-\nu} := K_1 \times \cdots \times K_{\nu-1} \times K_{\nu+1} \times \cdots \times K_N,$$

tiene imágenes convexas, compactas y no vacías.

Siendo cada  $X_\nu : K_{-\nu} \rightrightarrows K_\nu$  sci y scs, las multifunciones  $K \rightrightarrows K_\nu$  dadas por  $x \mapsto X_\nu(x^{-\nu})$ , donde  $K := K_1 \times \cdots \times K_N$ , son también sci y scs (véase el Ejemplo 4.4), lo cual significa que

$$\Omega : K \rightrightarrows K$$

es sci y scs (véase el comentario que sigue al Ejemplo 4.6). Recíprocamente, si las multifunciones  $K \rightrightarrows K_\nu$ ,  $x \mapsto X_\nu(x^{-\nu})$  son sci y scs, entonces cada  $X_\nu : K_{-\nu} \rightrightarrows K_\nu$  también es sci y scs (véase el Ejemplo 4.4).

Así, las hipótesis sobre  $\Omega$  y los  $X_\nu$  en el Teorema 10.8 nos dicen en esencia que  $\Omega : K \rightrightarrows K$  es sci y scs con imágenes convexas, compactas y no vacías. En particular, el gráfico  $G(\Omega)$  de  $\Omega : K \rightrightarrows K$  es cerrado por la Proposición 4.15.

**PRUEBA DEL TEOREMA 10.8.** Procedemos como en la prueba del teorema de Nash (Teorema 9.4): mostraremos que el operador  $\mathcal{S} : K \rightrightarrows K$  satisface las hipótesis del teorema de Kakutani (Teorema 4.20), lo cual implicará que  $\mathcal{S}$  posee un punto fijo. La prueba de que  $\mathcal{S}$  tiene imágenes convexas y no vacías es la misma que en la demostración del teorema de Nash, reemplazando  $X_\nu$  por  $X_\nu(\bar{x}^{-\nu})$  y usando la cuasiconvexidad de las aplicaciones  $x^\nu \mapsto \theta_\nu(x^\nu, \bar{x}^{-\nu})$ .

Solo nos falta mostrar que las semicontinuidades superior e inferior de cada  $X_\nu$  implican que  $G(\mathcal{S})$  es cerrado. Sea  $(x_k, y_k)$  una sucesión en  $G(\mathcal{S})$  que converge a  $(\bar{x}, \bar{y})$ . Entonces  $(\bar{x}, \bar{y}) \in G(\Omega)$ , pues  $G(\Omega)$  es cerrado y  $G(\mathcal{S}) \subset G(\Omega)$ . Además  $y_k^\nu \in \mathcal{S}_\nu(x_k^{-\nu})$  para cualesquiera  $k$  y  $\nu$ , es decir,

$$\theta_\nu(y_k^\nu, x_k^{-\nu}) \leq \theta_\nu(z^\nu, x_k^{-\nu}) \text{ para cualesquiera } k, \nu \text{ y } z^\nu \in X_\nu(\bar{x}^{-\nu}).$$

Queremos ver que  $(\bar{x}, \bar{y}) \in G(\mathcal{S})$ , o sea que

$$\theta_\nu(\bar{y}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu}) \text{ para cualesquiera } \nu \text{ y } x^\nu \in X_\nu(\bar{x}^{-\nu}).$$

Dados  $\nu$  y  $x^\nu \in X_\nu(\bar{x}^{-\nu})$ , como  $X_\nu : K_{-\nu} \rightrightarrows K_\nu$  es sci y  $x_k^{-\nu} \rightarrow \bar{x}^{-\nu}$ , existe por el Lema 4.8 una sucesión  $(z_k^\nu)$  en  $K_\nu$  que converge a  $x^\nu$  y tal que  $z_k^\nu \in X_\nu(x_k^{-\nu})$  para todo  $k$ . De manera que tomando límite en

$$\theta_\nu(y_k^\nu, x_k^{-\nu}) \leq \theta_\nu(z_k^\nu, x_k^{-\nu})$$

conseguimos

$$\theta_\nu(\bar{y}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu}),$$

que era lo que deseábamos probar.  $\square$

**Observación 10.9.** Pareciera que en la prueba anterior no fue usada la hipótesis de semicontinuidad superior de los  $X_\nu : K_{-\nu} \rightrightarrows K_\nu$ . Sin embargo, ella fue necesaria para deducir que  $G(\Omega)$  era cerrado, lo que implicó  $(\bar{x}, \bar{y}) \in G(\Omega)$ , esto es,  $\bar{y}^\nu \in X_\nu(\bar{x}^{-\nu})$  para cada  $\nu$ .

# Capítulo 3

## Reformulaciones como problemas de desigualdad variacional

En este capítulo presentamos varias reformulaciones del GNEP como problemas de desigualdad variacional. Lo que haremos esencialmente será aplicar las condiciones de optimalidad de la Sección 8, que traducen un problema de minimización del tipo

$$\min f(x), \quad x \in \mathbb{R}^n$$

en un problema de desigualdad variacional, a los problemas de minimización

$$\min_{x^\nu} \theta_\nu(x^\nu, x^{-\nu}), \quad x^\nu \in X_\nu(x^{-\nu}) \quad (10.2)$$

que enfrenta cada jugador<sup>1</sup>. Como vimos en dicha sección, el tipo de reformulación que obtuvimos dependió de las propiedades de  $f$ . Igualmente, las hipótesis que adoptemos determinarán el tipo de reformulación del GNEP que obtendremos aquí.

Antes de comenzar, recordamos y fijamos algunas notaciones y terminología del Capítulo 2. Primero que nada, trabajaremos siempre bajo la hipótesis de Rosen, esto es, existirá siempre un conjunto **convexo** y no vacío  $X \subset \mathbb{R}^n$  tal que cada  $X_\nu(x^{-\nu})$  está dado por

$$X_\nu(x^{-\nu}) = \{x^\nu \in \mathbb{R}^{n_\nu} \mid (x^\nu, x^{-\nu}) \in X\}; \quad (10.3)$$

en particular, los conjuntos  $X_\nu(x^{-\nu})$  también serán convexos. Luego, el GNEP será un NEP si y solo si el conjunto  $X$  tiene la forma

$$X = X_1 \times \cdots \times X_N, \quad (10.4)$$

donde cada  $X_\nu \subset \mathbb{R}^{n_\nu}$  es convexo y no vacío, en cuyo caso

$$X_\nu(x^{-\nu}) = \begin{cases} X_\nu, & \text{si existe } x^\nu \in \mathbb{R}^{n_\nu} \text{ tal que } (x^\nu, x^{-\nu}) \in X, \\ \emptyset, & \text{si } (x^\nu, x^{-\nu}) \notin X \text{ para todo } x^\nu \in \mathbb{R}^{n_\nu}. \end{cases}$$

A menos que se indique lo contrario, vamos a suponer tácitamente que las funciones objetivo de nuestros GNEPs satisfacen la hipótesis de **continuidad**

---

<sup>1</sup>En términos generales, la mayor parte de este capítulo puede verse como una serie de corolarios de los resultados de la Sección 8.

**H1.** cada  $\theta_\nu$  es continua en  $\mathbb{R}^n$ .

Adicionalmente, los GNEPs considerados satisfarán las siguientes hipótesis (no necesariamente todas) de **diferenciabilidad, convexidad y cuasiconvexidad**

**H2.** cada  $\theta_\nu$  es de clase  $C^1$  en  $\mathbb{R}^n$ ;

**H3.** cada  $\theta_\nu$  es convexa respecto a la  $\nu$ -ésima variable;

**H4.** cada  $\theta_\nu$  es cuasiconvexa respecto a la  $\nu$ -ésima variable,

respectivamente. Si  $x = (x^1, \dots, x^N) \in \mathbb{R}^n$ , entonces

$$\Omega(x) = X_1(x^{-1}) \times \dots \times X_N(x^{-N}) \subset \mathbb{R}^n,$$

definiendo esto una multifunción  $\Omega : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  cuyos puntos fijos  $x \in \Omega(x)$  son los **puntos factibles** del GNEP. Sabemos del Lema 10.3 que

$$x \in \Omega(x) \quad \text{si y solo si} \quad x \in X,$$

es decir, los puntos factibles del GNEP **son precisamente los elementos de  $X$** . Toda solución  $\bar{x} \in \mathbb{R}^n$  del GNEP es un vector factible, o sea  $\bar{x} \in X$ . Recíprocamente, un vector factible  $\bar{x} \in X$  es una solución del GNEP si y solo si

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu}) \quad \text{para cualesquiera } \nu \in \{1, \dots, N\} \text{ y } x^\nu \in X_\nu(x^{-\nu}).$$

Vamos a reformular el GNEP primero bajo las hipótesis de diferenciabilidad **H2.** y convexidad **H3.** (Sección 11). Luego debilitaremos la hipótesis de diferenciabilidad, trabajando solo con las hipótesis de continuidad **H1.** y convexidad **H3.** (Sección 12). Después reemplazaremos la hipótesis de convexidad por la de cuasiconvexidad, trabajando solamente con las hipótesis **H1.** y **H4.** (Sección 13). Aunque las reformulaciones serán totales (nos darán condiciones necesarias y suficientes) para el NEP, y parciales (nos darán solo condiciones suficientes) para el GNEP general, en cada caso ellas nos permitirán obtener resultados análogos al teorema de Rosen (Teorema 10.6).

## 11. Caso convexo y diferenciable

Trabajamos primero con GNEPs que satisfacen las hipótesis de diferenciabilidad y convexidad **H2.** y **H3.**. Entonces para cada  $\nu \in \{1, \dots, N\}$  y  $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$  fijos la aplicación

$$\theta_\nu(\cdot, x^{-\nu}) : x^\nu \mapsto \theta_\nu(x^\nu, x^{-\nu}),$$

que el  $\nu$ -ésimo jugador desea minimizar sobre  $X_\nu(x^{-\nu})$ , es convexa y diferenciable. De ahí que la idea para reformular el GNEP sea utilizar la condición de optimalidad de la Subsección 8.1, aplicable a problemas de minimización cuya función objetivo es convexa y diferenciable.

Más precisamente, para cada punto  $x = (x^1, \dots, x^N)$  en  $\mathbb{R}^n$  definamos el vector

$$F(x) := (\nabla_{x^\nu} \theta_\nu(x^\nu, x^{-\nu}))_{\nu=1}^N = (\nabla_{x^1} \theta_1(x^1, x^{-1}), \dots, \nabla_{x^N} \theta_N(x^N, x^{-N})) \in \mathbb{R}^n,$$

donde  $\nabla_{x^\nu} \theta_\nu(x^\nu, x^{-\nu})$  es el gradiente de  $\theta_\nu(\cdot, x^{-\nu})$  en el punto  $x^\nu$ . Entonces existe un resultado clásico que relaciona el GNEP con el problema de desigualdad variacional  $VI(F, X)$  asociado a la función  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ .

**Teorema 11.1.** *Consideremos un GNEP que satisface las hipótesis de diferenciabilidad y convexidad **H2.** y **H3.** Entonces toda solución del problema de desigualdad variacional  $VI(F, X)$  es solución del GNEP.*

PRUEBA. Sea  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N) \in X$  una solución del problema  $VI(F, X)$ . Entonces

$$\sum_{\mu=1}^N \langle \nabla_{x^\mu} \theta_\mu(\bar{x}^\mu, \bar{x}^{-\mu}), z^\mu - \bar{x}^\mu \rangle = \langle F(\bar{x}), z - \bar{x} \rangle \geq 0 \quad \text{para todo } z = (z^1, \dots, z^N) \in X,$$

donde  $\langle \cdot, \cdot \rangle$  representa el producto interno correspondiente en cada caso. Debido a la hipótesis de Rosen, cada componente  $\bar{x}^\nu$  de  $\bar{x}$  está en  $X_\nu(\bar{x}^{-\nu})$ . Además, para cada  $\nu \in \{1, \dots, N\}$  y  $x^\nu \in X_\nu(\bar{x}^{-\nu})$  tenemos que el vector

$$z := (x^\nu, \bar{x}^{-\nu}) = (\bar{x}^1, \dots, \bar{x}^{\nu-1}, x^\nu, \bar{x}^{\nu+1}, \dots, \bar{x}^N)$$

está en  $X$  y tiene la propiedad de que  $z^\mu = \bar{x}^\mu$  si  $\mu \neq \nu$  y  $z^\nu = x^\nu$ , de modo que

$$\langle \nabla_{x^\nu} \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}), x^\nu - \bar{x}^\nu \rangle = \sum_{\mu=1}^N \langle \nabla_{x^\mu} \theta_\mu(\bar{x}^\mu, \bar{x}^{-\mu}), z^\mu - \bar{x}^\mu \rangle \geq 0.$$

La Proposición 8.1 implica por tanto que para cada  $\nu$  fijo el vector  $\bar{x}^\nu$  minimiza la función convexa y diferenciable  $\theta_\nu(\cdot, \bar{x}^{-\nu})$  sobre el conjunto convexo  $X_\nu(\bar{x}^{-\nu})$ , es decir,  $\bar{x}$  es solución del GNEP.  $\square$

Consecuentemente, si encontramos alguna solución del problema  $VI(F, X)$  entonces habremos encontrado un equilibrio de Nash generalizado. La importancia del Teorema 11.1 radica justamente en este hecho, ya que existen muchas más herramientas, teóricas y algorítmicas, para abordar los problemas de desigualdad variacional que para abordar el GNEP. Sin embargo, esta es solo una reformulación parcial del GNEP porque en general hay muchas más soluciones del GNEP que soluciones del problema  $VI(F, X)$ , de manera que, por ejemplo, si determinamos que el problema  $VI(F, X)$  no posee soluciones entonces no podemos afirmar nada sobre el GNEP.

El siguiente ejemplo muestra lo que ya fue aludido en el párrafo anterior: que no toda solución del GNEP es solución del problema  $VI(F, X)$ .

**Ejemplo 11.2.** Consideremos el GNEP del Ejemplo 10.2, que satisface las hipótesis **H2.** y **H3.** Tenemos  $X = \{(x^1, x^2) \mid x^1 + x^2 \leq 1\}$ . Como  $\theta_1(x^1, x^2) = (x^1 - 1)^2$  y  $\theta_2(x^1, x^2) = (x^2 - \frac{1}{2})^2$ , la función  $F$  tiene la forma

$$F(x^1, x^2) = (2(x^1 - 1), 2(x^2 - \frac{1}{2})).$$

Gráficamente, podemos ver que el problema  $VI(F, X)$  posee una única solución en el punto  $\bar{x} := (\frac{3}{4}, \frac{1}{4})$ . Todos los demás puntos del segmento que une  $\alpha := (\frac{1}{2}, \frac{1}{2})$  y  $\beta := (1, 0)$  no son soluciones de  $VI(F, X)$ , pero sí son soluciones del GNEP, como ya fue probado en el Ejemplo 10.2.

No obstante que el GNEP y el problema  $VI(F, X)$  no son equivalentes en general, sí lo son si en vez del GNEP consideramos el NEP, es decir, si suponemos que el conjunto  $X$  del GNEP está dado como en (10.4), o sea  $X = X_1 \times \dots \times X_N$  donde cada  $X_\nu$  es convexo y no vacío.

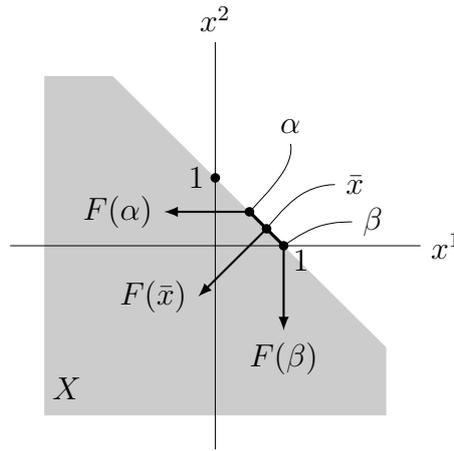


Figura 11.1: Gráfico del Ejemplo 11.2.

**Teorema 11.3.** *Consideremos un GNEP que satisface las hipótesis de diferenciabilidad y convexidad **H2.** y **H3.** Si el conjunto  $X$  está definido como en (10.4), entonces las soluciones del problema de desigualdad variacional  $VI(F, X)$  coinciden con las soluciones del GNEP. En particular, el NEP es equivalente al problema  $VI(F, X)$ .*

PRUEBA. Vamos a mostrar que toda solución del GNEP es solución del problema  $VI(F, X)$ . Sea  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N)$  una solución del GNEP. Entonces  $\bar{x} \in X$  y cada  $\bar{x}^\nu$  minimiza la función convexa y diferenciable  $\theta_\nu(\cdot, \bar{x}^{-\nu})$  sobre el conjunto convexo  $X_\nu(\bar{x}^{-\nu})$ . Sabemos por la Proposición 8.1 que  $\langle \nabla_{x^\nu} \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}), x^\nu - \bar{x}^\nu \rangle \geq 0$  para todo  $x^\nu \in X_\nu(\bar{x}^{-\nu})$ . En particular, dado cualquier vector  $x = (x^1, \dots, x^N)$  en  $X = X_1 \times \dots \times X_N$  tenemos que cada  $x^\nu$  está en  $X_\nu = X_\nu(\bar{x}^{-\nu})$ , de donde

$$\langle F(\bar{x}), x - \bar{x} \rangle = \sum_{\nu=1}^N \langle \nabla_{x^\nu} \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}), x^\nu - \bar{x}^\nu \rangle \geq 0.$$

Esto prueba que  $\bar{x}$  es solución del problema  $VI(F, X)$ .  $\square$

Así, resolver el NEP equivale a resolver el problema  $VI(F, X)$ . Desde un punto de vista histórico, este fue uno de los primeros resultados que conectaron los equilibrios de Nash con el problema de desigualdad variacional; desde un punto de vista práctico, esta reformulación contribuyó mucho al estudio de los NEPs (véase por ejemplo [13]).

Ahora aplicamos nuestra reformulación (Teorema 11.1) para probar un caso particular del teorema de Rosen (Teorema 10.6).

**Teorema 11.4** (Existencia de solución). *Consideremos un GNEP que satisface las hipótesis de diferenciabilidad y convexidad **H2.** y **H3.** Si el conjunto  $X$  es no vacío, convexo y compacto, entonces el GNEP posee solución.*

PRUEBA. De acuerdo con el Teorema 11.1, bastará con mostrar que el problema de desigualdad variacional  $VI(F, X)$  admite alguna solución. Como cada aplicación

$x \mapsto \nabla_{x^\nu} \theta_\nu(x^\nu, x^{-\nu})$  es continua (pues  $\theta_\nu$  es de clase  $C^1$ ), la función  $F$ , que está dada por

$$F(x) = (\nabla_{x^1} \theta_1(x^1, x^{-1}), \dots, \nabla_{x^N} \theta_N(x^N, x^{-N})),$$

es continua. El Teorema 7.4 nos garantiza por tanto que el problema  $VI(F, X)$  posee solución. Esto concluye la prueba del teorema.  $\square$

Hemos visto que la diferenciabilidad de las funciones  $\theta_\nu$ , o sea la hipótesis **H2.**, jugó un papel crucial en los teoremas que acabamos de probar. Esto debido a que la función  $F$  está definida en términos de los gradientes de los  $\theta_\nu$ 's. Desde luego, si omitimos la hipótesis **H2.** ya no podemos utilizar tales gradientes, por lo que dichos teoremas ya no son válidos. Sin embargo, en la siguiente sección veremos que gracias a la hipótesis de convexidad **H3.**, cada teorema de esta sección puede generalizarse de forma natural al caso en que los  $\theta_\nu$ 's no son diferenciables.

**Observación 11.5** (Generalizaciones de los teoremas). En esta observación enumeramos varias extensiones de los teoremas probados hasta ahora, que se desprenden de sus demostraciones.

- (1) Los Teoremas 11.1 y 11.3 se cumplen bajo hipótesis un poco más generales. A saber, como la Proposición 8.1 es válida cuando la función objetivo respectiva es convexa y diferenciable (o pseudoconvexa), las hipótesis **H2.** y **H3.** pueden ser reemplazadas por la condición

$$\text{cada función } \theta_\nu(\cdot, x^{-\nu}) : x^\nu \mapsto \theta_\nu(x^\nu, x^{-\nu}) \text{ es convexa y diferenciable,} \quad (11.1)$$

o, más generalmente, por la **hipótesis de pseudoconvexidad**

$$\text{cada } \theta_\nu \text{ es pseudoconvexa respecto a la } \nu\text{-ésima variable.} \quad (11.2)$$

Es claro que también podemos prescindir de la hipótesis tácita de continuidad **H1.** Más aun, todo funciona si el conjunto  $X$  de la hipótesis de Rosen, en vez de ser convexo, tan solo satisface la hipótesis de que

$$X_\nu(x^{-\nu}) = \{x^\nu \mid (x^\nu, x^{-\nu}) \in X\} \text{ es convexo para cualesquiera } \nu \text{ y } x^{-\nu}. \quad (11.3)$$

Si bien esto último mejora ligeramente el Teorema 11.1, no conseguimos nada nuevo respecto del Teorema 11.3, pues al considerar NEPs (esto es,  $X = X_1 \times \dots \times X_N$ ) tenemos  $X_\nu(x^{-\nu}) = \emptyset$  o  $X_\nu(x^{-\nu}) = X_\nu$ , de modo que la afirmación (11.3) equivale a la convexidad de  $X$ .

- (2) Puede verse que el Teorema 11.4 también se cumple cuando reemplazamos las hipótesis **H2.** y **H3.** por la condición (11.1). En este caso, la continuidad de las aplicaciones  $x \mapsto \nabla_{x^\nu} \theta_\nu(x^\nu, \bar{x}^{-\nu})$  en la prueba de dicho teorema se sigue del Corolario 2.29. Como este corolario es consecuencia de la Proposición 2.27, que no vale cuando la función  $f$  en cuestión es pseudoconvexa, no podemos reemplazar la condición (11.1) por la hipótesis de pseudoconvexidad (11.2). Tampoco podemos debilitar la hipótesis de continuidad **H1.** (véanse las hipótesis del Corolario 2.29), ni la convexidad del conjunto  $X$  de la hipótesis de Rosen, ya que dicha convexidad es necesaria para aplicar el Teorema de Kakutani-Fan (Teorema 4.19) en la prueba del Teorema 7.6.

**Comentario aparte 11.6** (Desigualdad cuasivariacional). Supongamos que el GNEP considerado no verifica la hipótesis de Rosen y que, en vez de tener un conjunto convexo  $X$  tal que (10.3) se satisface, tan solo tenemos la hipótesis de que

$$X_\nu(x^{-\nu}) \text{ es convexo para cualesquiera } \nu \text{ y } x^{-\nu}.$$

Supongamos además que las hipótesis de diferenciabilidad y convexidad **H2.** y **H3.** se cumplen, o, más generalmente, que la hipótesis de pseudoconvexidad (11.2) es satisfecha. Entonces las demostraciones de los Teoremas 11.1 y 11.3 nos permiten ver que un punto factible  $\bar{x} \in \Omega(\bar{x})$  es solución del GNEP si y solo si  $\bar{x}$  es solución del problema de desigualdad cuasivariacional  $QVI(F, \Omega)$ , donde la multifunción  $\Omega$  está dada por

$$\Omega(x) = X_1(x^{-1}) \times \cdots \times X_N(x^{-N}).$$

Obtenemos así que el GNEP es equivalente al problema de desigualdad cuasivariacional  $QVI(F, \Omega)$ . Más aun, esta equivalencia permanece válida si prescindimos de la hipótesis de continuidad **H1.** Obsérvese finalmente que la equivalencia clásica entre el NEP y el problema  $VI(F, X)$ , obtenida en el Teorema 11.3, es un caso particular de esta nueva equivalencia, ya que al considerar el NEP en lugar del GNEP, la hipótesis de Rosen es automáticamente verificada y los problemas  $VI(F, X)$  y  $VI(F, \Omega)$  son los mismos.

## Condiciones KKT

Teniendo en cuenta que toda solución del problema  $VI(F, X)$  es solución del GNEP, pero no recíprocamente, surge una pregunta natural: ¿cuáles son las soluciones del GNEP que se preservan al pasar del GNEP al problema  $VI(F, X)$ ? Aunque no existe una respuesta completa a esta pregunta, podemos utilizar las condiciones KKT (véanse las Secciones 3 y 7) para dar una caracterización de tales soluciones del GNEP en ciertos casos. Con este fin, vamos a suponer que el conjunto convexo  $X$  de la definición del GNEP está dado por

$$X = \{x \in \mathbb{R}^n \mid g_1(x) \leq 0, \dots, g_r(x) \leq 0\}, \quad (11.4)$$

donde  $g_1, \dots, g_r : \mathbb{R}^n \rightarrow \mathbb{R}$  son funciones convexas y diferenciables.

La idea es la siguiente. Supongamos que  $\bar{x} \in X$  es una solución del GNEP. Entonces cada  $\bar{x}^\nu$  minimiza  $\theta_\nu(\cdot, \bar{x}^{-\nu})$  sobre el conjunto convexo  $X_\nu(\bar{x}^{-\nu})$ . Luego, si alguna condición de calificación<sup>2</sup> (CQ) se satisface en  $\bar{x}^\nu$ , entonces existe un vector  $\lambda^\nu = (\lambda_1^\nu, \dots, \lambda_r^\nu)$  de multiplicadores en  $\bar{x}^\nu$  tales que las condiciones KKT

$$\begin{aligned} \nabla_{x^\nu} \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) + \sum_i \lambda_i^\nu \cdot \nabla_{x^\nu} g_i(\bar{x}^\nu, \bar{x}^{-\nu}) &= 0, \\ 0 \leq \lambda^\nu \perp g(\bar{x}) \leq 0, \end{aligned} \quad (11.5)$$

<sup>2</sup>Recordemos que una **condición de calificación (CQ)** para el problema

$$\min f(x), \quad x \in X,$$

donde  $X = \{x \mid g_1(x), \dots, g_r(x) \leq 0\}$ , es una condición suficiente para que en un mínimo local dado  $\bar{x} \in X$  del problema exista un vector de multiplicadores de Lagrange. Véase la página 43.

se cumplen.

Similarmente, si  $\bar{x} \in X$  es solución del problema  $VI(F, X)$  y alguna  $CQ^3$  se satisface en  $\bar{x}$ , entonces existe un vector de multiplicadores  $\lambda = (\lambda_1, \dots, \lambda_r)$  tales que las condiciones KKT

$$\begin{aligned} F(\bar{x}) + \sum_i \lambda_i \cdot \nabla g_i(\bar{x}) &= 0, \\ 0 \leq \lambda \perp g(\bar{x}) &\leq 0, \end{aligned}$$

se satisfacen. Por la definición de  $F$ , este sistema equivale a

$$\begin{aligned} \nabla_{x^\nu} \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) + \sum_i \lambda_i \cdot \nabla_{x^\nu} g_i(\bar{x}^\nu, \bar{x}^{-\nu}) &= 0, \\ 0 \leq \lambda \perp g(\bar{x}) &\leq 0. \end{aligned} \tag{11.6}$$

Con estas observaciones podemos probar el siguiente teorema.

**Teorema 11.7** (Condiciones KKT). *Consideremos un GNEP que satisface las hipótesis de diferenciabilidad y convexidad **H2.** y **H3.** Supongamos que el conjunto  $X$  asociado está dado como en (11.4).*

- (1) *Si  $\bar{x}$  es una solución del problema  $VI(F, X)$  para la que existen multiplicadores  $\lambda \in \mathbb{R}^r$  que verifican (11.6), entonces  $\bar{x}$  es una solución del GNEP que satisface (11.5) con  $\lambda^1 = \dots = \lambda^N = \lambda$ .*
- (2) *Si  $\bar{x}$  es una solución del GNEP en la que existen multiplicadores  $\lambda^1, \dots, \lambda^N$  que verifican (11.5), y tales que  $\lambda^1 = \dots = \lambda^N$ , entonces  $\bar{x}$  es una solución del problema  $VI(F, X)$  en la que  $\lambda = \lambda^1 = \dots = \lambda^N$ .*

PRUEBA. Básicamente, aplicaremos los Teoremas 3.9 y 7.3.

(1) Como los multiplicadores  $\lambda$  satisfacen (11.6), entonces claramente satisfacen (11.5) con  $\lambda^1 = \dots = \lambda^N = \lambda$ . Luego, aplicando el Teorema 3.9 vemos que cada  $\bar{x}^\nu$  minimiza  $\theta_\nu(\cdot, \bar{x}^{-\nu})$  sobre  $X_\nu(\bar{x}^{-\nu})$ , es decir,  $\bar{x}$  es solución del GNEP.

(2) Definiendo  $\lambda := \lambda^1 = \dots = \lambda^N$ , vemos que  $\lambda$  verifica (11.6). Luego, gracias al Teorema 7.3 concluimos que  $\bar{x}$  es solución de  $VI(F, X)$ .  $\square$

**Observación 11.8** (Caso pseudoconvexo). Dado que el Teorema 3.9 es válido cuando la función objetivo es pseudoconvexa (y no solamente convexa y diferenciable), el Teorema 11.7 permanece válido si sustituimos las hipótesis de diferenciabilidad y convexidad **H2.** y **H3.** por la condición (11.1), o, más generalmente, por la hipótesis de pseudoconvexidad (11.2).

**Ejemplo 11.9.** Revisemos el GNEP del Ejemplo 10.5, que verifica las hipótesis **H2.** y **H3.** Como  $\theta_1(x^1, x^2) = x^1 + x^2 - 1$  y  $\theta_2(x^1, x^2) = x^1 x^2$ , la función  $F$  está dada por

$$F(x^1, x^2) = (1, x^1).$$

<sup>3</sup>Una **condición de calificación (CQ)** para el problema  $VI(F, X)$ , donde  $X$  tiene la forma  $\{x \mid g_1(x), \dots, g_r(x) \leq 0\}$ , es una condición suficiente para que en una solución  $\bar{x} \in X$  de  $VI(F, X)$  exista un vector de multiplicadores de Lagrange. Véase la página 70.

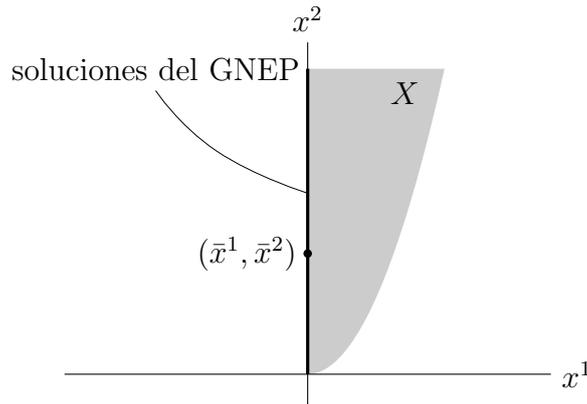


Figura 11.2: Gráfico del Ejemplo 11.9.

Como también  $X = \{(x^1, x^2) \in \mathbb{R}^2 \mid x^1 \geq 0, x^2 \geq (x^1)^2\}$ , las funciones  $g_i$  son

$$g_1(x^1, x^2) = -x^1 \quad \text{y} \quad g_2(x^1, x^2) = (x^1)^2 - x^2.$$

Ya sabemos que el conjunto de soluciones de este GNEP es  $\{0\} \times [0, \infty)$ . Analicemos los multiplicadores de Lagrange en cada punto  $\bar{x} = (\bar{x}^1, \bar{x}^2)$  de este conjunto. Observamos en seguida que

$$F(\bar{x}^1, \bar{x}^2) = (1, 0), \quad \nabla g_1(x^1, x^2) = (-1, 0) \quad \text{y} \quad \nabla g_2(x^1, x^2) = (0, -1).$$

Supongamos primero que  $\bar{x}^2 > 0$ . Entonces el conjunto activo es  $I(\bar{x}) = \{1\}$ , por lo que cada vector de multiplicadores  $\lambda^\nu$  (si existe) tendrá solo una coordenada. Identificamos cada  $\lambda^\nu$  con dicha coordenada<sup>4</sup>. Ahora es fácil ver que  $\lambda^1 = 1$ ,  $\lambda^2 = *$  pues

$$1 + 1 \cdot (-1) = 0 \quad \text{y} \quad 0 + * \cdot 0 = 0,$$

donde  $*$  representa a cualquier número no negativo. Tomando  $\lambda^2 = 1$  tenemos  $\lambda^1 = \lambda^2$ , lo que por la parte (2) del Teorema 11.7 significa que  $\bar{x}$  es solución de  $VI(F, X)$  con vector de multiplicadores  $\lambda = \lambda^1 = \lambda^2 = 1$ ; efectivamente,

$$(1, 0) + 1 \cdot (-1, 0) = (0, 0).$$

Supongamos ahora que  $\bar{x}^2 = 0$ . El conjunto activo es  $I(\bar{x}) = \{1, 2\}$  y cada vector  $\lambda^\nu$  (si existe) debe tener dos coordenadas. Podemos entonces deducir que  $\lambda^1 = (1, *)$  y  $\lambda^2 = (*, 0)$ , pues

$$1 + 1 \cdot (-1) + * \cdot 0 = 0 \quad \text{y} \quad 0 + * \cdot 0 + 0 \cdot (-1) = 0.$$

Escogiendo  $\lambda^1 = (1, 0)$  y  $\lambda^2 = (1, 0)$  tenemos  $\lambda^1 = \lambda^2$ , lo cual, nuevamente, significa que  $\bar{x}$  es solución del problema  $VI(F, X)$  con vector de multiplicadores de Lagrange  $\lambda = \lambda^1 = \lambda^2 = (1, 0)$ ; en efecto,

$$(1, 0) + 1 \cdot (-1, 0) + 0 \cdot (0, -1) = (0, 0).$$

Esto quiere decir en particular que las soluciones de este GNEP coinciden con las soluciones de  $VI(F, X)$ .

<sup>4</sup>Nótese que ya hicimos esto para los vectores de estrategias  $x = (x^1, x^2)$  cuando  $n_1 = n_2 = 1$ . En efecto, al representar cada  $x^\nu$  a una sola coordenada,  $x^\nu$  pasaba a representar a dicha coordenada.

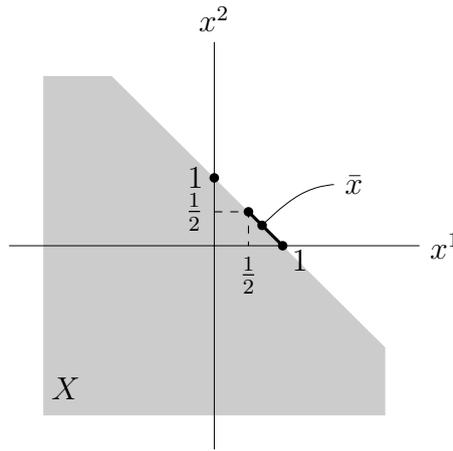


Figura 11.3: Gráfico del Ejemplo 11.10.

**Ejemplo 11.10.** Veamos nuevamente el Ejemplo 10.2. Teníamos allí que  $\theta_1(x^1, x^2) = (x^1 - 1)^2$ ,  $\theta_2(x^1, x^2) = (x^2 - \frac{1}{2})^2$ ,  $X = \{(x^1, x^2) \mid x^1 + x^2 \leq 1\}$  y

$$F(x^1, x^2) = (2(x^1 - 1), 2(x^2 - \frac{1}{2})).$$

Existe una única función  $g_i$  que define  $X$ , a saber,

$$g_1(x^1, x^2) = x^1 + x^2 - 1.$$

A continuación analizamos los multiplicadores de Lagrange en las soluciones del GNEP, o sea en los puntos  $\bar{x} = (\bar{x}^1, \bar{x}^2)$  tales que  $\frac{1}{2} \leq \bar{x}^1 \leq 1$  y  $\bar{x}^1 + \bar{x}^2 = 1$ . Evidentemente,

$$F(\bar{x}^1, \bar{x}^2) = (2(\bar{x}^1 - 1), 2(\bar{x}^2 - \frac{1}{2})) \quad \text{y} \quad \nabla g_1(\bar{x}^1, \bar{x}^2) = (1, 1).$$

Como el conjunto activo  $I(\bar{x}) = \{1\}$  tiene siempre solo un elemento, realizamos la misma identificación del ejemplo anterior. Ahora es fácil ver que  $\lambda^1 = -2(\bar{x}^1 - 1)$  y que  $\lambda^2 = -2(\bar{x}^2 - \frac{1}{2})$  pues

$$2(\bar{x}^1 - 1) + (-2(\bar{x}^1 - 1)) \cdot 1 = 0 \quad \text{y} \quad 2(\bar{x}^2 - \frac{1}{2}) + (-2(\bar{x}^2 - \frac{1}{2})) \cdot 1 = 0.$$

Por otra parte, las CQs de Abadie (véase la página 70) se satisfacen en  $\bar{x}$ , de forma que siempre existen multiplicadores en  $\bar{x}$ , que están únicamente determinados por ser  $\nabla g_1(\bar{x}^1, \bar{x}^2)$  no nulo. Luego, se sigue del Teorema 11.7 que  $\bar{x}$  es solución de  $VI(F, X)$  si y solo si  $\lambda^1 = \lambda^2$ , es decir, si y solo si  $\bar{x} = (\frac{3}{4}, \frac{1}{4})$ . Notemos que ya vimos esto en el Ejemplo 11.2.

## 12. Caso convexo

En esta sección adaptamos las ideas de la sección anterior para reformular el GNEP en la situación más general en que cada función de coste  $\theta_\nu$  no es diferenciable, pero sí convexa respecto a la  $\nu$ -ésima variable. Utilizando los subdiferenciales, que son

el análogo de los gradientes para las funciones convexas que no son diferenciables, construiremos una multifunción  $\Phi$  y usaremos las condiciones de optimalidad de la Subsección 8.2 para relacionar el GNEP con el problema de desigualdad variacional  $VI(\Phi, X)$ . Veremos que este proceso es una generalización natural de lo realizado en la sección anterior.

Consideramos en esta sección GNEPs que satisfacen la hipótesis de convexidad **H3.**, es decir, para cada  $\nu \in \{1, \dots, N\}$  y  $x^{-\nu} \in \mathbb{R}^{n-n_\nu}$  fijos la aplicación

$$\theta_\nu(\cdot, x^{-\nu}) : x^\nu \mapsto \theta_\nu(x^\nu, x^{-\nu})$$

es convexa (y continua). Cabe recordar que implícitamente asumimos que la hipótesis de continuidad **H1.** también es satisfecha.

Escribamos  $\partial_{x^\nu} \theta_\nu(\bar{x}^\nu, x^{-\nu})$  para denotar al subdiferencial de  $\theta_\nu(\cdot, x^{-\nu})$  en el punto  $\bar{x}^\nu \in \mathbb{R}^{n_\nu}$ ; explícitamente,

$$\begin{aligned} \partial_{x^\nu} \theta_\nu(\bar{x}^\nu, x^{-\nu}) &= \partial \theta_\nu(\cdot, x^{-\nu})(\bar{x}^\nu) \\ &= \{v^\nu \in \mathbb{R}^{n_\nu} \mid \theta_\nu(x^\nu, x^{-\nu}) \geq \theta_\nu(\bar{x}^\nu, x^{-\nu}) + \langle v^\nu, x^\nu - \bar{x}^\nu \rangle, \forall x^\nu \in \mathbb{R}^{n_\nu}\}. \end{aligned}$$

Para cada  $x = (x^1, \dots, x^N) \in \mathbb{R}^n$  definimos

$$\Phi(x) = \Phi(x^1, \dots, x^N) = \partial_{x^1} \theta_1(x^1, x^{-1}) \times \dots \times \partial_{x^N} \theta_N(x^N, x^{-N}) \subset \mathbb{R}^n.$$

Tenemos así una multifunción  $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  y, consecuentemente, un problema de desigualdad variacional  $VI(\Phi, X)$ .

Puesto que el subdiferencial es la generalización del gradiente, la multifunción  $\Phi$  es una suerte de generalización de la función  $F$  de la sección anterior. De hecho, el problema de desigualdad variacional  $VI(\Phi, X)$  se relaciona con el GNEP de la misma manera que el problema  $VI(F, X)$  en el caso diferenciable.

**Teorema 12.1.** *Consideremos un GNEP que satisface la hipótesis de convexidad **H3.** Entonces toda solución del problema de desigualdad variacional  $VI(\Phi, X)$  es solución del GNEP.*

PRUEBA. Sea  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N) \in X$  una solución del problema  $VI(\Phi, X)$ . Entonces existe  $v = (v^1, \dots, v^N) \in \Phi(\bar{x})$  tal que

$$\sum_{\mu=1}^N \langle v^\mu, z^\mu - \bar{x}^\mu \rangle = \langle v, z - \bar{x} \rangle \geq 0 \quad \text{para todo } z = (z^1, \dots, z^N) \in X,$$

donde  $\langle \cdot, \cdot \rangle$  representa el producto interno correspondiente en cada caso. Para cada  $\nu \in \{1, \dots, N\}$  y  $x^\nu \in X_\nu(\bar{x}^{-\nu})$  tenemos que el vector

$$z := (x^\nu, \bar{x}^{-\nu}) = (\bar{x}^1, \dots, \bar{x}^{\nu-1}, x^\nu, \bar{x}^{\nu+1}, \dots, \bar{x}^N)$$

está en  $X$  y tiene la propiedad de que  $z^\mu = \bar{x}^\mu$  si  $\mu \neq \nu$  y  $z^\nu = x^\nu$ , de modo que

$$\langle v^\nu, x^\nu - \bar{x}^\nu \rangle = \sum_{\mu=1}^N \langle v^\mu, z^\mu - \bar{x}^\mu \rangle \geq 0.$$

Como  $v^\nu \in \partial_{x^\nu} \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu})$ , lo anterior significa que para cada  $\nu$  fijo el punto  $\bar{x}^\nu$  es solución del problema

$$VI(\partial \theta_\nu(\cdot, \bar{x}^{-\nu}), X_\nu(\bar{x}^{-\nu})),$$

lo cual, por el Teorema 8.4, quiere decir que  $\bar{x}^\nu$  minimiza la función convexa  $\theta_\nu(\cdot, \bar{x}^{-\nu})$  sobre el conjunto convexo  $X_\nu(\bar{x}^{-\nu})$ . Es decir,  $\bar{x}$  es solución del GNEP.  $\square$

De nuevo, si conseguimos conocer alguna solución del problema  $VI(\Phi, X)$  entonces habremos determinado un equilibrio de Nash generalizado. Sin embargo, esta reformulación del GNEP es parcial porque hay en general muchas más soluciones del GNEP que soluciones del problema  $VI(\Phi, X)$ , como se observa en el siguiente ejemplo.

**Ejemplo 12.2.** Modificamos ligeramente el Ejemplo 10.2. Consideremos el GNEP

$$\begin{array}{ll} \min_{x^1} (x^1 - 1)^2 & \min_{x^2} |x^2 - \frac{1}{2}| \\ x^1 + x^2 \leq 1 & x^1 + x^2 \leq 1 \end{array}$$

con conjunto  $X = \{(x^1, x^2) \mid x^1 + x^2 \leq 1\}$ , y cuyo conjunto solución es

$$\{(\bar{x}^1, \bar{x}^2) \mid \frac{1}{2} \leq \bar{x}^1 \leq 1, \bar{x}^1 + \bar{x}^2 = 1\}.$$

Como  $\theta_1(x^1, x^2) = (x^1 - 1)^2$  y  $\theta_2(x^1, x^2) = |x^2 - \frac{1}{2}|$ , el operador  $\Phi$  está dado por

$$\Phi(x^1, x^2) = \begin{cases} \{2(x^1 - 1)\} \times \{-1\}, & \text{si } x^2 < \frac{1}{2}, \\ \{2(x^1 - 1)\} \times [-1, 1], & \text{si } x^2 = \frac{1}{2}, \\ \{2(x^1 - 1)\} \times \{1\}, & \text{si } x^2 > \frac{1}{2}. \end{cases}$$

Geoméricamente, vemos que el problema  $VI(\Phi, X)$  posee una única solución en  $\bar{x} := (\frac{1}{2}, \frac{1}{2})$ , de manera que las demás soluciones del GNEP no son soluciones de  $VI(\Phi, X)$ . Obsérvese que, a diferencia de lo que obtuvimos en el Ejemplo 11.2, donde la solución de  $VI(F, X)$  se daba en  $(\frac{3}{4}, \frac{1}{4})$ , la solución de  $VI(\Phi, X)$  ocurre en  $(\frac{1}{2}, \frac{1}{2})$ .

Al igual que en la sección anterior, no obstante que el GNEP y el problema  $VI(\Phi, X)$  no son equivalentes en general, sí lo son si en vez del GNEP consideramos el NEP.

**Teorema 12.3.** *Consideremos un GNEP que satisface la hipótesis de convexidad **H3**.. Si el conjunto  $X$  está definido como en (10.4), entonces las soluciones del problema de desigualdad variacional  $VI(\Phi, X)$  coinciden con las soluciones del GNEP. En particular, el NEP es equivalente al problema  $VI(\Phi, X)$ .*

PRUEBA. Vamos a mostrar que toda solución del GNEP es una solución del problema  $VI(\Phi, X)$ . Sea  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N)$  una solución del GNEP. Entonces  $\bar{x} \in X$  y cada  $\bar{x}^\nu$  minimiza la función convexa  $\theta_\nu(\cdot, \bar{x}^{-\nu})$  sobre el conjunto convexo  $X_\nu(\bar{x}^{-\nu})$ . Sabemos por el Teorema 8.4 que cada  $\bar{x}^\nu$  es solución del problema

$$VI(\partial \theta_\nu(\cdot, \bar{x}^{-\nu}), X_\nu(\bar{x}^{-\nu})),$$

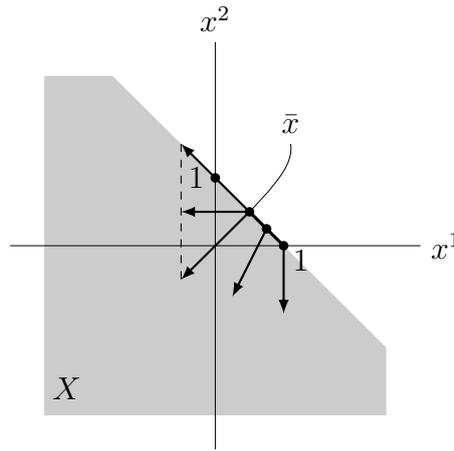


Figura 12.1: Gráfico del Ejemplo 12.2.

lo cual significa que existe  $v^\nu \in \partial_{x^\nu} \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu})$  tal que  $\langle v^\nu, x^\nu - \bar{x}^\nu \rangle \geq 0$  para todo  $x^\nu \in X_\nu(\bar{x}^{-\nu})$ . En particular, dado cualquier  $x = (x^1, \dots, x^N)$  en  $X = X_1 \times \dots \times X_N$  tenemos que cada  $x^\nu$  está en  $X_\nu = X_\nu(\bar{x}^{-\nu})$  y por consiguiente

$$\langle v, x - \bar{x} \rangle = \sum_{\nu=1}^N \langle v^\nu, x^\nu - \bar{x}^\nu \rangle \geq 0,$$

donde  $v := (v^1, \dots, v^N) \in \partial_{x^1} \theta_1(\bar{x}^1, \bar{x}^{-1}) \times \dots \times \partial_{x^N} \theta_N(\bar{x}^N, \bar{x}^{-N}) = \Phi(\bar{x})$ . Esto muestra que  $\bar{x}$  es solución del problema  $VI(\Phi, X)$ .  $\square$

En definitiva, resolver el NEP equivale a resolver el problema  $VI(\Phi, X)$ .

A continuación aplicamos nuestra reformulación (Teorema 12.1) para extender el Teorema 11.4 y obtener otro caso particular del teorema de Rosen (Teorema 10.6).

**Teorema 12.4** (Existencia de solución). *Consideremos un GNEP que satisface la hipótesis de convexidad **H3**. Si el conjunto  $X$  es no vacío, convexo y compacto, entonces el GNEP posee solución.*

PRUEBA. Por el Teorema 12.1, basta con mostrar que el problema de desigualdad variacional  $VI(\Phi, X)$  admite alguna solución. Por definición, el operador  $\Phi$  está dado por

$$\Phi(x) = \partial_{x^1} \theta_1(x^1, x^{-1}) \times \dots \times \partial_{x^N} \theta_N(x^N, x^{-N}).$$

Afirmamos que  $\Phi$  es scs y tiene imágenes convexas, compactas y no vacías. Para ver esto, tan solo debemos probar que cada operador  $\Phi_\nu : \mathbb{R}^n \rightrightarrows \mathbb{R}^{n_\nu}$  definido por

$$\Phi_\nu(x) = \partial_{x^\nu} \theta_\nu(x^\nu, x^{-\nu})$$

es scs y tiene imágenes convexas, compactas y no vacías (Ejemplo 4.6). La Proposición 2.36 nos asegura que  $\Phi_\nu$  tiene imágenes convexas, compactas y no vacías. Y, por otro lado, cada  $\theta_\nu(\cdot, x^{-\nu})$  depende solo de  $x^{-\nu}$ , por lo que la semicontinuidad superior de  $\Phi_\nu$  se sigue de la Proposición 2.37.

Así, el operador  $\Phi$  verifica las hipótesis del Teorema 7.6, el cual nos garantiza que el problema  $VI(\Phi, X)$  posee solución.  $\square$

**Observación 12.5** (Generalizaciones de los teoremas).

- (1) En los Teoremas 12.1 y 12.3 no necesitamos que las funciones  $\theta_\nu$  sean continuas en  $\mathbb{R}^n$ , si bien cada aplicación  $\theta_\nu(\cdot, x^{-\nu})$  es continua por ser convexa. De hecho, solo la hipótesis **H3**. fue usada en sus demostraciones. Esto quiere decir que podemos omitir la hipótesis de continuidad **H1**.. Más aun, todo funciona si el conjunto  $X$  de la hipótesis de Rosen, en vez de ser convexo, tan solo satisface la hipótesis de que

$$X_\nu(x^{-\nu}) = \{x^\nu \mid (x^\nu, x^{-\nu}) \in X\} \text{ es convexo para cualesquiera } \nu \text{ y } x^{-\nu}. \quad (12.1)$$

A pesar de que esto último mejora levemente el Teorema 12.1, no sucede lo mismo con el Teorema 12.3, pues al considerar NEPs (o sea  $X = X_1 \times \cdots \times X_N$ ) tenemos  $X_\nu(x^{-\nu}) = \emptyset$  o  $X_\nu(x^{-\nu}) = X_\nu$ , de manera que (12.1) equivale a la convexidad de  $X$ .

- (2) En cuanto al Teorema 12.4, no podemos omitir la hipótesis **H1**. debido a que ella es necesaria para aplicar la Proposición 2.37. Tampoco podemos debilitar la convexidad del conjunto  $X$  asociado, ya que dicha convexidad es requerida para poder aplicar el teorema de Kakutani-Fan (Teorema 4.19) en la demostración del Teorema 7.6.

**Comentario aparte 12.6** (Desigualdad cuasivariacional). Supongamos que el GNEP considerado no verifica la hipótesis de Rosen y que, en vez de tener un conjunto convexo  $X$  tal que (10.3) se satisface, solo tenemos la hipótesis de que

$$X_\nu(x^{-\nu}) \text{ es convexo para cualesquiera } \nu \text{ y } x^{-\nu}.$$

Supongamos además que la hipótesis de convexidad **H3**. es verificada. Entonces las pruebas de los Teoremas 12.1 y 12.3 nos permiten ver que un punto factible  $\bar{x} \in \Omega(\bar{x})$  es solución del GNEP si y solo si  $\bar{x}$  es solución del problema de desigualdad cuasivariacional  $QVI(\Phi, \Omega)$ , donde, recordemos, la multifunción  $\Omega$  está dada por

$$\Omega(x) = X_1(x^{-1}) \times \cdots \times X_N(x^{-N}).$$

Así, el GNEP es equivalente al problema de desigualdad cuasivariacional  $QVI(\Phi, \Omega)$ , siendo esta equivalencia válida incluso cuando omitimos la hipótesis tácita **H1**.. Nótese finalmente que al considerar el NEP (en lugar del GNEP) la hipótesis de Rosen es automáticamente verificada, y que en este caso los problemas  $VI(\Phi, X)$  y  $VI(\Phi, \Omega)$  son los mismos.

## 13. Caso cuasiconvexo

En esta sección vamos a generalizar los resultados de la sección anterior al caso cuasiconvexo del GNEP. Específicamente, supondremos que cada  $\theta_\nu$  es cuasiconvexa respecto a la  $\nu$ -ésima variable, y emplearemos las herramientas del Análisis cuasiconvexo, a saber, las recientemente introducidas nociones de subniveles ajustados y sus operadores normales (véase la Subsección 5.2).

Partimos de la siguiente observación: en la sección anterior utilizamos los subdiferenciales de las funciones objetivo de cada jugador (que eran convexas) para transformar el GNEP, vía las condiciones de optimalidad de la Subsección 8.3, en un problema de desigualdad variacional. En esta sección haremos lo mismo: usaremos los subniveles ajustados de tales funciones (que ahora son solo cuasiconvexas) y sus operadores normales para convertir el GNEP, por medio de las condiciones de optimalidad de la Subsección 8.3, en un problema de desigualdad variacional. En particular, esto nos permitirá apreciar cuán útiles son dichos operadores normales.

Consideraremos en esta sección GNEPs que satisfacen la hipótesis de cuasiconvexidad **H4.**. Como la hipótesis de continuidad **H1.** se cumple implícitamente, tenemos en particular que para cada  $\nu \in \{1, \dots, N\}$  y  $\theta_\nu(\cdot, x^{-\nu})$  la aplicación

$$\theta_\nu(\cdot, x^{-\nu}) : x^\nu \mapsto \theta_\nu(x^\nu, x^{-\nu})$$

es cuasiconvexa y continua.

Antes de definir el operador multivaluado de la reformulación, introducimos un poco de terminología para simplificar la notación. Dados  $\nu \in \{1, \dots, N\}$  y  $x = (x^1, \dots, x^N) \in \mathbb{R}^n$  escribimos

$$A_\nu(x^{-\nu}) := \arg \min_{\mathbb{R}^{n_\nu}} \theta_\nu(\cdot, x^{-\nu}),$$

y

$$\begin{aligned} S_\nu(x) &:= S_{\theta_\nu(\cdot, x^{-\nu})}^a(x^\nu), \\ N_{\theta_\nu}^a(x^\nu) &:= N_{\theta_\nu(\cdot, x^{-\nu})}^a(x^\nu), \end{aligned}$$

donde  $S_{\theta_\nu(\cdot, x^{-\nu})}^a(x^\nu)$  es el subnivel ajustado de la función cuasiconvexa  $\theta_\nu(\cdot, x^{-\nu})$  en  $x^\nu$  y  $N_{\theta_\nu(\cdot, x^{-\nu})}^a(x^\nu)$  el operador normal correspondiente (véase la Subsección 5.2); explícitamente,

$$N_{\theta_\nu}^a(x^\nu) = \{v^\nu \in \mathbb{R}^{n_\nu} \mid \langle v^\nu, y^\nu - x^\nu \rangle \leq 0, \forall y^\nu \in S_\nu(x)\}.$$

En seguida definimos la multifunción deseada  $N_\theta^a : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ . Dado  $x = (x^1, \dots, x^N) \in \mathbb{R}^n$  sea

$$N_\theta^a(x) = N_\theta^a(x^1, \dots, x^N) := F_1(x) \times \cdots \times F_N(x),$$

donde cada  $F_\nu : \mathbb{R}^n \rightrightarrows \mathbb{R}^{n_\nu}$  está dado por

$$F_\nu(x) = \begin{cases} \overline{B}_\nu(0, 1), & \text{si } x^\nu \in A_\nu(x^{-\nu}), \\ \text{co}(N_{\theta_\nu}^a(x^\nu) \cap S_\nu(0, 1)), & \text{si } x^\nu \notin A_\nu(x^{-\nu}). \end{cases}$$

Aquí,  $\overline{B}_\nu(0, 1)$  es la bola unitaria cerrada de  $\mathbb{R}^{n_\nu}$  y  $S_\nu(0, 1)$  es la esfera unitaria en  $\mathbb{R}^{n_\nu}$ .

De inmediato observamos la similitud entre el operador  $F_\nu$  y la multifunción  $F_f$  definida en (8.2) a partir de una función cuasiconvexa  $f$ . En efecto, utilizando la notación  $F_f$  de la Subsección 8.3 tenemos que

$$F_\nu(x) = F_\nu(x^\nu, x^{-\nu}) = F_{\theta_\nu(\cdot, x^{-\nu})}(x^\nu) \quad \text{para todo } x = (x^1, \dots, x^N) \in \mathbb{R}^n.$$

o, equivalentemente,

$$F_\nu(\cdot, x^{-\nu}) = F_{\theta_\nu(\cdot, x^{-\nu})} \quad \text{para todo } x^{-\nu} \in \mathbb{R}^{n-n_\nu},$$

donde  $F_\nu(\cdot, x^{-\nu}) : \mathbb{R}^{n_\nu} \rightrightarrows \mathbb{R}^{n_\nu}$  es el operador que aplica  $\bar{x}^\nu$  en  $F_\nu(\bar{x}^\nu, x^{-\nu})$ . Lo cual quiere decir que  $F_\nu(\cdot, x^{-\nu})$  es el operador  $F_f$  asociado a la función  $f = \theta_\nu(\cdot, x^{-\nu})$ . Esta observación nos permitirá aplicar los resultados de la Subsección 8.3 al operador  $N_\theta^a$ .

**Observación 13.1.** Cuando pasamos de la Sección 11 a la Sección 12 dijimos que, no pudiendo usar gradientes, íbamos a emplear subdiferenciales, por lo que la multifunción  $\Phi$  era una suerte de generalización de la función  $F$ . Análogamente, aquí no podemos usar subdiferenciales (no tenemos convexidad) pero sí los operadores normales (tenemos cuasiconvexidad), de forma que  $N_\theta^a$  es el sustituto de  $\Phi$  en este nuevo contexto: el problema  $VI(N_\theta^a, X)$  tendrá casi todas las propiedades que el problema  $VI(\Phi, X)$  tenía en la sección anterior.

La siguiente proposición es una clara consecuencia de las Proposiciones 8.5 y 8.6.

**Proposición 13.2.** *Consideremos un GNEP que satisface la hipótesis de cuasiconvexidad  $\mathbf{H4}$ . Entonces los operadores  $F_\nu$  tienen imágenes convexas, compactas y no vacías. Además, dado  $\bar{x} \in \mathbb{R}^n$  se cumple que  $0 \in F_\nu(\bar{x})$  si y solo si  $\bar{x}^\nu \in A_\nu(\bar{x}^{-\nu})$ .*

Examinamos ahora la relación entre el problema de desigualdad variacional  $VI(N_\theta^a, X)$  y el GNEP. Comenzamos demostrando el análogo del Teorema 12.1.

**Teorema 13.3.** *Consideremos un GNEP que satisface la hipótesis de cuasiconvexidad  $\mathbf{H4}$ . Entonces toda solución del problema de desigualdad variacional  $VI(N_\theta^a, X)$  es solución del GNEP.*

PRUEBA. Supongamos que  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N) \in X$  es una solución del problema  $VI(N_\theta^a, X)$ . Entonces existe  $v = (v^1, \dots, v^N) \in N_\theta^a(\bar{x})$  tal que

$$\sum_{\mu=1}^N \langle v^\mu, z^\mu - \bar{x}^\mu \rangle = \langle v, z - \bar{x} \rangle \geq 0, \quad \forall z = (z^1, \dots, z^N) \in X.$$

donde  $\langle \cdot, \cdot \rangle$  representa el producto interno correspondiente en cada caso. Observemos que cada componente  $\bar{x}^\nu$  de  $\bar{x}$  está en  $X_\nu(\bar{x}^{-\nu})$ , ya que  $\bar{x} \in X$ . Ahora bien, para cada  $\nu \in \{1, \dots, N\}$  y  $x^\nu \in X_\nu(\bar{x}^{-\nu})$  tenemos que el vector

$$z := (x^\nu, \bar{x}^{-\nu}) = (\bar{x}^1, \dots, \bar{x}^{\nu-1}, x^\nu, \bar{x}^{\nu+1}, \dots, \bar{x}^N)$$

está en  $X$  y tiene la propiedad de que  $z^\mu = \bar{x}^\mu$  si  $\mu \neq \nu$  y  $z^\nu = x^\nu$ , de forma que

$$\langle v^\nu, x^\nu - \bar{x}^\nu \rangle = \sum_{\mu} \langle v^\mu, z^\mu - \bar{x}^\mu \rangle \geq 0.$$

Como  $v^\nu \in F_\nu(\bar{x}) = F_{\theta_\nu(\cdot, \bar{x}^{-\nu})}(\bar{x}^\nu)$ , lo anterior significa que para cada  $\nu$  fijo el punto  $\bar{x}^\nu$  es solución del problema

$$VI(F_{\theta_\nu(\cdot, \bar{x}^{-\nu})}, X_\nu(\bar{x}^{-\nu})),$$

lo cual, por el Teorema 8.7, implica que  $\bar{x}^\nu$  minimiza la función cuasiconvexa y continua  $\theta_\nu(\cdot, \bar{x}^{-\nu})$  sobre el conjunto convexo  $X_\nu(\bar{x}^{-\nu})$ . Esto muestra que  $\bar{x}$  es solución del GNEP.  $\square$

Esto nos da una condición suficiente para poder afirmar que un punto  $\bar{x}$  es solución del GNEP. Sin embargo, no es cierto que tal condición sea necesaria, es decir, puede haber soluciones del GNEP que no sean soluciones del problema  $VI(N_\theta^a, X)$ . Utilizamos el Ejemplo 8.8 para construir un GNEP que ilustre este punto.

**Ejemplo 13.4.** En  $\mathbb{R}^2$  consideremos el GNEP de  $N = 2$  jugadores

$$\begin{array}{ll} \min_{x^1} f(x^1) & \min_{x^2} x^2 \\ |x^1 - 2| \leq 1, & |x^1 - 2| \leq 1, \\ |x^2| \leq 1, & |x^2| \leq 1, \end{array}$$

donde  $f(t) = \min\{1, |t|\}$ . El conjunto asociado a este GNEP es  $X = [1, 3] \times [-1, 1]$ , por lo que se trata en realidad de un NEP. No es difícil verificar que el conjunto de soluciones de este GNEP es justamente

$$[1, 3] \times \{-1\}.$$

Dado  $\bar{x} = (\bar{x}^1, \bar{x}^2) \in X$ , tenemos

$$A_1(\bar{x}^2) = \{0\} \quad \text{y} \quad A_2(\bar{x}^1) = \emptyset,$$

de manera que

$$S_1(\bar{x}) = S_{\theta_1(\cdot, \bar{x}^2)}^a(\bar{x}^1) = [-\bar{x}^1, \bar{x}^1] \quad \text{y} \quad S_2(\bar{x}) = S_{\theta_2(\bar{x}^1, \cdot)}^a(\bar{x}^2) = (-\infty, \bar{x}^2];$$

luego  $N_{\theta_1}^a(\bar{x}^1) = N_{\theta_2}^a(\bar{x}^2) = [0, \infty)$ , y en consecuencia

$$F_1(\bar{x}) = F_2(\bar{x}) = \{1\},$$

es decir,

$$N_\theta^a(\bar{x}) = \{(1, 1)\}.$$

Ahora es fácil ver gráficamente que  $\bar{x}$  es solución del problema  $VI(N_\theta^a, X)$  si y solo si  $\bar{x} = (1, -1)$ . O sea que todos los demás puntos de  $[1, 3] \times \{-1\}$  son soluciones del GNEP pero no soluciones de  $VI(N_\theta^a, X)$ .

No siendo el GNEP y el problema  $VI(N_\theta^a, X)$  equivalentes en general, nos preguntamos qué condiciones tendríamos que imponer sobre el GNEP para que tal equivalencia se dé (o sea, para que el recíproco del Teorema 13.3 se verifique). Al contrario que en las dos secciones anteriores, cuando consideramos el NEP en lugar del GNEP la equivalencia todavía no ocurre (el GNEP del Ejemplo 13.4 es un NEP). Con el fin de superar esta dificultad, introducimos la siguiente **hipótesis de cuasiconvexidad semiestricta**<sup>5</sup>

**H5.** cada  $\theta_\nu$  es cuasiconvexa semiestricta con respecto a la  $\nu$ -ésima variable,

<sup>5</sup>Recordemos que una función cuasiconvexa  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es cuasiconvexa semiestricta cuando  $f(tx + (1-t)y) < \max\{f(x), f(y)\}$  para cualesquiera  $t \in (0, 1)$  y  $x, y \in \mathbb{R}^n$  con  $f(x) \neq f(y)$ .

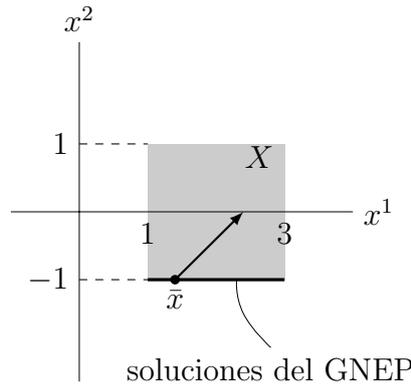


Figura 13.1: Gráfico del Ejemplo 13.4.

la cual, como tácitamente asumimos la hipótesis **H1.**, implica que cada aplicación

$$\theta_\nu(\cdot, x^{-\nu}) : x^\nu \mapsto \theta_\nu(x^\nu, x^{-\nu})$$

es cuasiconvexa semiestricta y continua.

Obviamente, esta hipótesis es más fuerte que la hipótesis de cuasiconvexidad **H4.** Pero es más débil que la hipótesis de convexidad **H3.** pues toda función convexa es cuasiconvexa semiestricta. Esquemáticamente,

$$\mathbf{H3.} \Rightarrow \mathbf{H5.} \Rightarrow \mathbf{H4.}$$

Con la hipótesis **H5.** ya podemos probar el análogo del Teorema 12.3.

**Teorema 13.5.** *Consideremos un GNEP que satisface la hipótesis de cuasiconvexidad semiestricta **H5.** Si el conjunto  $X$  está definido como en (10.4), entonces las soluciones del problema de desigualdad variacional  $VI(N_\theta^a, X)$  coinciden con las soluciones del GNEP. En particular, el NEP es equivalente al problema  $VI(N_\theta^a, X)$ .*

PRUEBA. Sea  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N)$  una solución del GNEP. Entonces  $\bar{x} \in X$  y cada  $\bar{x}^\nu$  minimiza la función cuasiconvexa semiestricta y continua  $\theta_\nu(\cdot, \bar{x}^{-\nu})$  sobre el conjunto convexo  $X_\nu(\bar{x}^{-\nu})$ . Sabemos por el Teorema 8.9 que cada  $\bar{x}^\nu$  es solución del problema

$$VI(F_{\theta_\nu(\cdot, \bar{x}^{-\nu})}, X_\nu(\bar{x}^{-\nu})),$$

lo que significa que existe  $v^\nu \in F_{\theta_\nu(\cdot, \bar{x}^{-\nu})}(\bar{x}^\nu) = F_\nu(\bar{x})$  tal que  $\langle v^\nu, x^\nu - \bar{x}^\nu \rangle \geq 0$  para todo  $x^\nu \in X_\nu(\bar{x}^{-\nu})$ . En particular, dado cualquier  $x = (x^1, \dots, x^N)$  en  $X = X_1 \times \dots \times X_N$  tenemos que cada  $x^\nu$  está en  $X_\nu = X_\nu(\bar{x}^{-\nu})$  y en consecuencia

$$\langle v, x - \bar{x} \rangle = \sum_{\nu=1}^N \langle v^\nu, x^\nu - \bar{x}^\nu \rangle \geq 0,$$

donde  $v := (v^1, \dots, v^N) \in F_1(\bar{x}) \times \dots \times F_N(\bar{x}) = N_\theta^a(\bar{x})$ . Esto prueba que  $\bar{x}$  es solución del problema  $VI(N_\theta^a, X)$ .  $\square$

Hemos encontrado así una reformulación total del NEP para el caso cuasiconvexo semiestricto, en el sentido de que resolver el NEP equivale a resolver el problema  $VI(N_\theta^a, X)$ .

Ahora aplicamos nuestra reformulación (Teorema 13.3) para mostrar una generalización del teorema de Rosen (Teorema 10.6).

**Teorema 13.6** (Existencia de solución). *Consideremos un GNEP que satisfice la hipótesis de cuasiconvexidad semiestricta **H5**. Si el conjunto  $X$  es no vacío, convexo y compacto, entonces el GNEP posee solución.*

PRUEBA. Utilizaremos el Teorema 7.6 para probar que el problema  $VI(N_\theta^a, X)$  posee solución. Entonces tenemos que probar que el operador  $N_\theta^a$ , dado por

$$N_\theta^a(x) = F_1(x) \times \cdots \times F_N(x),$$

es scs y tiene imágenes convexas, compactas y no vacías. Para ver esto basta con mostrar que cada  $F_\nu$  es scs y tiene imágenes convexas, compactas y no vacías (véase el Ejemplo 4.6). Gracias a la Proposición 13.2 las imágenes de  $F_\nu$  ya son convexas, compactas y no vacías, por lo que solo nos falta verificar que  $F_\nu$  es scs.

Como siempre se cumple que  $F_\nu(x^\nu, x^{-\nu}) = F_{\theta_\nu(\cdot, x^{-\nu})}(x^\nu)$ , donde la aplicación  $\theta_\nu(\cdot, x^{-\nu}) : \mathbb{R}^{n_\nu} \rightarrow \mathbb{R}$  es continua y cuasiconvexa semiestricta, concluimos del Teorema 8.10 que  $F_\nu$  es scs. Esto completa la prueba del teorema.  $\square$

**Observación 13.7** (Generalizaciones de los teoremas).

- (1) El Teorema 13.3 se cumple bajo hipótesis más generales. Efectivamente, basta con recordar que el Teorema 8.7, que fue utilizado en su prueba, es válido cuando la función objetivo del problema de minimización respectivo es cuasiconvexa y continua, sin ninguna restricción sobre la región factible respectiva. Luego, la hipótesis explícita **H4**. y la hipótesis implícita **H1**. pueden ser reemplazadas por la condición

$$\text{cada } \theta_\nu \text{ es cuasiconvexa y continua respecto a la } \nu\text{-ésima variable; } \quad (13.1)$$

también, el conjunto  $X$  puede ser arbitrario (no convexo, por ejemplo). Sin embargo, una hipótesis natural, que en cierto modo es necesaria para que la cuasiconvexidad de las funciones  $\theta_\nu(\cdot, x^{-\nu})$  tenga sentido, es que el conjunto  $X$  satisfaga la propiedad de que

$$X_\nu(x^{-\nu}) = \{x^\nu \mid (x^\nu, x^{-\nu}) \in X\} \text{ es convexo para cualesquiera } \nu \text{ y } x^{-\nu}. \quad (13.2)$$

- (2) Hacemos lo mismo con el Teorema 13.5. Examinando el Teorema 8.9, vemos que podemos reemplazar las hipótesis **H4**. y **H1**. del Teorema 13.5 por la condición

$$\begin{aligned} \text{cada } \theta_\nu \text{ es cuasiconvexa semiestricta y continua} \\ \text{respecto a la } \nu\text{-ésima variable.} \end{aligned} \quad (13.3)$$

Vemos además que aunque todo funciona cuando el conjunto  $X$ , en vez de ser convexo, satisfice la propiedad (13.2), en realidad no conseguimos nada nuevo, ya que  $X = X_1 \times \cdots \times X_N$  implica que (13.2) equivale a la convexidad de  $X$ .

- (3) En lo que se refiere al Teorema 13.6, no es posible omitir la hipótesis **H1.** debido a que ella es necesaria para aplicar el Teorema 8.10. Tampoco podemos debilitar la convexidad del conjunto asociado  $X$ , puesto que dicha convexidad es requerida para poder aplicar el teorema de Kakutani-Fan (Teorema 4.19) en la demostración del Teorema 7.6.

**Comentario aparte 13.8** (Desigualdad cuasivariacional). Nuevamente, supongamos que el GNEP considerado no verifica la hipótesis de Rosen y que, en vez de tener un conjunto convexo  $X$  tal que (10.3) se satisface, solo tenemos la hipótesis de que

$$X_\nu(x^{-\nu}) \text{ es convexo para cualesquiera } \nu \text{ y } x^{-\nu}.$$

Supongamos además que la hipótesis de cuasiconvexidad **H4.** se cumple. Entonces la demostración del Teorema 13.3 nos permite ver que toda solución del problema de desigualdad cuasivariacional  $QVI(N_\theta^a, \Omega)$  es solución del GNEP, donde la multifunción  $\Omega$  está dada por

$$\Omega(x) = X_1(x^{-1}) \times \cdots \times X_N(x^{-N}).$$

Recíprocamente, si la hipótesis de cuasiconvexidad semiestricta **H5.** es satisfecha, entonces la demostración del Teorema 13.5 muestra que toda solución del GNEP es solución del problema de desigualdad cuasivariacional  $QVI(N_\theta^a, \Omega)$ . Concluimos así que el GNEP es equivalente al problema de desigualdad cuasivariacional  $VI(N_\theta^a, \Omega)$  bajo la hipótesis **H5.**, mas no necesariamente bajo la hipótesis **H4.** Notemos que la equivalencia entre el NEP y el problema  $VI(N_\theta^a, X)$ , establecida en el Teorema 13.5, es un caso particular de esta nueva equivalencia, ya que al considerar NEPs (en lugar de GNEPs) la hipótesis de Rosen es automáticamente verificada y los problemas  $VI(N_\theta^a, X)$  y  $QVI(N_\theta^a, \Omega)$  son los mismos.

La Observación 13.7 puede también adaptarse a esta reformulación del GNEP como problema de desigualdad cuasivariacional.

# Capítulo 4

## Reformulaciones como problemas de optimización

En este capítulo presentamos varias reformulaciones del GNEP como problemas de optimización (o minimización) del tipo

$$\min f(x), \quad x \in X.$$

Es natural esperar que tales reformulaciones existan, ya que el GNEP consiste en que todos los jugadores resuelvan sus respectivos problemas de minimización

$$\min \theta_\nu(x^\nu, x^{-\nu}), \quad x^\nu \in X_\nu(x^{-\nu})$$

al mismo tiempo. Como es lógico, las propiedades de las funciones  $\theta_\nu$  van a determinar el tipo de reformulación obtenida; veremos en particular que si ellas son de clase  $C^1$ , entonces la función objetivo de la reformulación será de clase  $C^1$ , lo cual es importante pues la mayoría de algoritmos existentes para resolver problemas de minimización se pueden aplicar más eficientemente a problemas *diferenciables*.

Todas las reformulaciones de este capítulo están basadas en el artículo [16].

Antes de comenzar, recordamos y fijamos algunas notaciones y terminología del Capítulo 2. Primero que nada, trabajaremos bajo la hipótesis de Rosen, esto es, siempre existirá un conjunto no vacío  $X \subset \mathbb{R}^n$  tal que cada  $X_\nu(x^{-\nu})$  está dado por

$$X_\nu(x^{-\nu}) = \{x^\nu \in \mathbb{R}^n \mid (x^\nu, x^{-\nu}) \in X\}.$$

También supondremos que el conjunto  $X$  es **convexo** y **cerrado**; en particular, los conjuntos  $X_\nu(x^{-\nu})$  también serán convexos y cerrados. En esta situación, el GNEP considerado será un NEP si y solo si

$$X = X_1 \times \cdots \times X_N,$$

donde cada  $X_\nu \subset \mathbb{R}^{n_\nu}$  es convexo y cerrado, en cuyo caso

$$X_\nu(x^{-\nu}) = \begin{cases} X_\nu, & \text{si existe } x^{\nu u} \in \mathbb{R}^{n_\nu} \text{ tal que } (x^\nu, x^{-\nu}) \in X, \\ \emptyset, & \text{si } (x^\nu, x^{-\nu}) \notin X \text{ para todo } (x^\nu, x^{-\nu}) \in X. \end{cases}$$

Vamos a suponer siempre que las funciones  $\theta_\nu$  verifican la hipótesis de continuidad

**H1.** cada  $\theta_\nu$  es continua en  $\mathbb{R}^n$ ,

y la hipótesis de convexidad

**H3.** cada  $\theta_\nu$  es convexa respecto a la  $\nu$ -ésima variable.

Más tarde también añadiremos la siguiente hipótesis de diferenciabilidad

**H2.** cada  $\theta_\nu$  es de clase  $C^1$  en  $\mathbb{R}^n$ .

Si  $x = (x^1, \dots, x^N) \in \mathbb{R}^n$ , entonces

$$\Omega(x) = X_1(x^{-1}) \times \dots \times X_N(x^{-N}) \subset \mathbb{R}^n;$$

en particular,  $\Omega$  define una multifunción  $\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ , cuyos puntos fijos  $x \in \Omega(x)$  son los **puntos factibles** del GNEP. Del Lema 10.3 sabemos que

$$x \in \Omega(x) \text{ si y solo si } x \in X.$$

En otras palabras, los vectores factibles del GNEP **son precisamente los elementos de  $X$** . Esta será una observación crucial en este capítulo.

## 14. Primeras reformulaciones

Un vector factible  $x \in X$  es solución del GNEP si y solo si las diferencias

$$\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu})$$

son menores o iguales que 0 para cualquier otro vector  $y = (y^1, \dots, y^N) \in \Omega(x)$ . Luego, encontrar puntos de equilibrio del GNEP equivale a conseguir puntos  $x$  tales que las diferencias sean lo más “pequeñas” posible. La función de Nikaidô-Isoda, que jugará un papel central en las reformulaciones que vamos a dar, nos ayuda a “medir” globalmente cuán pequeñas son estas diferencias.

**Definición 14.1.** La función  $\Psi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  dada por

$$\Psi(x, y) = \sum_{\nu=1}^N [\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu})]$$

es llamada la **función de Nikaidô-Isoda** o la función de Ky-Fan del GNEP.

Observamos que

$$\Psi(x, \cdot) \text{ es cóncava para todo } x \in \mathbb{R}^n$$

ya que las aplicaciones  $y^\nu \mapsto -\theta_\nu(y^\nu, x^{-\nu})$  son cóncavas en  $\mathbb{R}^{n_\nu}$  para cada  $\nu$  y  $x^{-\nu}$  fijos. Similarmente,

$$\Psi(\cdot, y) \text{ es convexa para todo } y \in \mathbb{R}^n.$$

Para cada  $x \in X$  definimos

$$V(x) = \sup_{y \in \Omega(x)} \Psi(x, y). \quad (14.1)$$

Esto nos da una función

$$V : X \rightarrow \mathbb{R} \cup \{\infty\},$$

que es no negativa pues  $x \in X$  implica  $x \in \Omega(x)$ , y por tanto

$$V(x) \geq \Psi(x, x) = 0.$$

**Observación 14.2.** Notemos que la función  $V$  ha sido definida en  $X$  y no en  $\mathbb{R}^n$ . Esto se debe a que  $\Omega(x)$  puede ser vacío cuando  $x \notin X$ , en cuyo caso  $V(x) = -\infty$ .

**Proposición 14.3.** *Para que un vector  $\bar{x} \in \mathbb{R}^n$  sea solución del GNEP es necesario y suficiente que  $\bar{x} \in X$  y que  $V(\bar{x}) = 0$ .*

PRUEBA. Toda solución del GNEP es un vector factible. Así, basta con probar que si  $\bar{x} \in X$ , entonces  $\bar{x}$  es solución del GNEP si y solo si  $V(\bar{x}) = 0$ .

Supongamos que  $V(\bar{x}) = 0$ . Dados  $\nu \in \{1, \dots, N\}$  y  $x^\nu \in X_\nu(\bar{x}^{-\nu})$ , debemos mostrar que  $\theta_\nu(\bar{x}^\nu, \bar{x}^\nu) - \theta_\nu(x^\nu, \bar{x}^\nu) \leq 0$ . El vector

$$y := (x^\nu, \bar{x}^{-\nu}) = (\bar{x}^1, \dots, \bar{x}^{\nu-1}, x^\nu, \bar{x}^{\nu+1}, \dots, \bar{x}^N)$$

está en  $X_1(\bar{x}^{-1}) \times \dots \times X_N(\bar{x}^{-N}) = \Omega(\bar{x})$  pues  $\bar{x} \in \Omega(\bar{x})$ , y tiene la propiedad de que  $y^\mu = \bar{x}^\mu$  si  $\mu \neq \nu$  y  $y^\nu = x^\nu$ . Luego,

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \theta_\nu(x^\nu, \bar{x}^{-\nu}) = \sum_{\mu=1}^N [\theta_\mu(\bar{x}^\mu, \bar{x}^{-\mu}) - \theta_\mu(y^\mu, \bar{x}^{-\mu})] = \Psi(\bar{x}, y) \leq V(\bar{x}) = 0.$$

Recíprocamente, si  $\bar{x}$  es solución del GNEP, entonces para todo  $y := (y^1, \dots, y^N)$  en  $\Omega(\bar{x})$  tenemos que  $\Psi(\bar{x}, y) \leq 0$  pues

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \theta_\nu(y^\nu, \bar{x}^{-\nu}) \leq 0$$

para cada  $\nu$ , y por tanto  $V(\bar{x}) \leq 0$ , esto es,  $V(\bar{x}) = 0$ .  $\square$

Así, las soluciones del GNEP son exactamente los mínimos globales  $x$  de la función  $V$  tales que  $V(x) = 0$ . Esto significa que encontrar las soluciones del GNEP equivale a resolver el problema de minimización

$$\min V(x), \quad x \in X. \quad (14.2)$$

Hemos obtenido, por tanto, una reformulación del GNEP como un problema de optimización con restricciones (véase la observación a seguir).

**Observación 14.4.** Nótese que, si bien toda solución del GNEP es solución el Problema (14.2), puede haber soluciones Problema (14.2) que no son soluciones del GNEP, a saber, los mínimos globales  $\bar{x} \in X$  de  $V$  tales que  $V(\bar{x}) > 0$ . No obstante, es claro que si conocemos las soluciones del Problema (14.2) entonces ya sabemos si el GNEP admite soluciones, y, si este fuera el caso, sabemos cuáles son. De ahí que el GNEP sea en realidad equivalente al Problema (14.2).

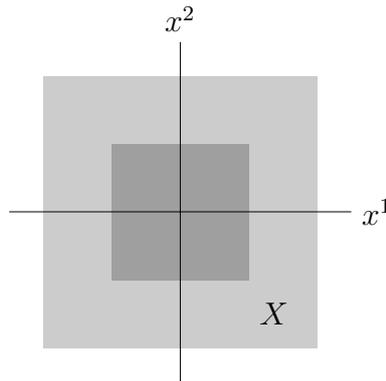


Figura 14.1: Gráfico del Ejemplo 14.5. La región sombreada del centro representa a las soluciones del GNEP.

**Ejemplo 14.5.** Consideremos en  $\mathbb{R}^2$  el GNEP con  $N = 2$  jugadores

$$\begin{array}{ll} \min_{x^1} f(x^1) & \min_{x^2} f(x^2) \\ |x^1| \leq 2, & |x^1| \leq 2, \\ |x^2| \leq 2, & |x^2| \leq 2, \end{array}$$

donde  $f(t) = \max\{|t| - 1, 0\}$ . El conjunto asociado a este GNEP es  $X = [-2, 2] \times [-2, 2]$ , por lo que se trata en realidad de un NEP. Calculamos la función de Nikaidô-Isoda

$$\Psi(x, y) = (f(x^1) - f(y^1)) + (f(x^2) - f(y^2)).$$

Dado  $\bar{x} = (\bar{x}^1, \bar{x}^2) \in X$ , tenemos  $\Omega(\bar{x}) = [-2, 2] \times [-2, 2] = X$ , de donde se sigue que

$$\begin{aligned} V(\bar{x}) &= \sup_{(y^1, y^2) \in \Omega(\bar{x})} (f(\bar{x}^1) - f(y^1) + f(\bar{x}^2) - f(y^2)) \\ &= f(\bar{x}^1) + f(\bar{x}^2) - \inf_{|y^1| \leq 2} f(y^1) - \inf_{|y^2| \leq 2} f(y^2) \\ &= f(\bar{x}^1) + f(\bar{x}^2). \end{aligned}$$

En particular,  $V(\bar{x}) = 0$  si y solo si  $f(\bar{x}^1) = f(\bar{x}^2) = 0$ , es decir, si y solo si  $\bar{x}$  pertenece al conjunto  $[-1, 1] \times [-1, 1]$ . Luego,  $[-1, 1] \times [-1, 1]$  representa a las soluciones de este NEP. Observamos finalmente que el supremo en la definición de  $V$  no es alcanzado en un único punto, sino mas bien en toda la región  $[-1, 1] \times [-1, 1]$ .

**Ejemplo 14.6.** Veamos un GNEP sin soluciones. Consideremos el GNEP de  $N = 2$  participantes

$$\begin{array}{ll} \min_{x^1} x^1 x^2 & \min_{x^2} x^1 - x^2 \\ |x^1| \leq x^2, & |x^1| \leq x^2, \end{array}$$

con conjunto asociado  $X = \{(x^1, x^2) \mid |x^1| \leq x^2\}$ . No es difícil verificar directamente que este GNEP no posee soluciones. Hagamos la verificación vía la función de Nikaidô-Isoda. Tenemos

$$\Psi(x, y) = (x^1 x^2 - y^1 x^2) + ((x^1 - x^2) - (x^1 - y^2)) = x^2(x^1 - y^1) + y^2 - x^2.$$

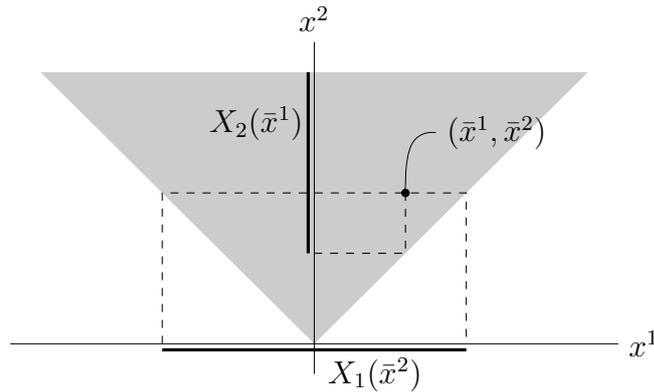


Figura 14.2: Gráfico del Ejemplo 14.6.

Luego, para cada  $\bar{x} = (\bar{x}^1, \bar{x}^2) \in X$  es claro que  $\Omega(\bar{x}) = [-\bar{x}^2, \bar{x}^2] \times [\bar{x}^1, \infty)$ , y por tanto

$$V(\bar{x}) = \sup_{(y^1, y^2) \in \Omega(\bar{x})} (\bar{x}^2(\bar{x}^1 - y^1) + y^2 - \bar{x}^2) = \infty.$$

Así, la función  $V$  es idénticamente infinito. En particular, el GNEP no posee soluciones.

La reformulación que acabamos de conseguir es una consecuencia lógica de la discusión que precedió a la Definición 14.1. De hecho, debido a su simplicidad es natural esperar que ella tenga algunos defectos:

- D1.** No podemos garantizar que la función  $V$  tome solo valores finitos. Es decir, podría ocurrir que  $\sup_{y \in \Omega(x)} \Psi(x, y) = \infty$  para un  $x \in X$  dado (véase el Ejemplo 14.6). Se trata de una gran desventaja si se quiere resolver el Problema (14.2) usando métodos iterativos.
- D2.** Incluso cuando dicho supremo sea finito, este podría no ser alcanzado, o ser alcanzado en más de un punto (véase el Ejemplo 14.5).
- D3.** La función objetivo  $V$  no es necesariamente diferenciable (cuando las funciones  $\theta_\nu$  lo son), debido al ítem anterior (véase también el Ejemplo 14.11). Este es un problema desagradable pues la mayoría de algoritmos que existen para resolver problemas de minimización de este tipo funcionan mejor cuando la función objetivo es diferenciable.

Deseamos encontrar reformulaciones que superen estos defectos. Con este fin, modificamos la función de Nikaidô-Isoda.

**Definición 14.7.** Sea  $\alpha > 0$  un parámetro fijo. La aplicación  $\Psi_\alpha : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  dada por

$$\Psi_\alpha(x, y) = \sum_{\nu=1}^N [\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) - \frac{\alpha}{2} \|x^\nu - y^\nu\|^2]$$

es llamada la **función regularizada de Nikaidô-Isoda**.

En seguida vemos que

$$\Psi_\alpha(x, y) = \Psi(x, y) - \frac{\alpha}{2} \|x - y\|^2,$$

y que en particular

$$\Psi_\alpha(x, y) \leq \Psi(x, y).$$

Luego, se sigue de la Proposición 2.30 que

$$\Psi_\alpha(x, \cdot) \text{ es fuertemente cóncava para todo } x \in \mathbb{R}^n.$$

De manera análoga a como hicimos en (14.1), para cada vector factible  $x \in X$  definimos

$$V_\alpha(x) = \max_{y \in \Omega(x)} \Psi_\alpha(x, y), \quad (14.3)$$

lo cual nos da una aplicación

$$V_\alpha : X \rightarrow \mathbb{R}.$$

Esta función está bien definida. En efecto, dado  $x \in X$ , la aplicación  $\Psi_\alpha(\cdot, y)$  es fuertemente cóncava, y como el conjunto  $\Omega(x)$  es convexo, cerrado y no vacío (pues  $x \in X$ ), por el Corolario 2.31 existe un (único) punto  $y' \in \Omega(x)$  tal que  $\max_{y \in \Omega(x)} \Psi_\alpha(x, y) = \Psi_\alpha(x, y')$ .

Como  $\Psi_\alpha(x, x) = 0$  y  $x \in X$  implica  $x \in \Omega(x)$ , es claro que  $V_\alpha$  es una función no negativa. Luego,

$$0 \leq V_\alpha(x) \leq V(x), \quad \forall x \in X. \quad (14.4)$$

**Observación 14.8.** De nuevo,  $V_\alpha$  ha sido definida solo sobre  $X$  y no en  $\mathbb{R}^n$ , ya que  $\Omega(x)$  podría ser vacío cuando  $x \notin X$ , en cuyo caso  $\sup_{y \in \Omega(x)} \Psi_\alpha(x, y) = -\infty$ .

**Teorema 14.9.** *La función  $V_\alpha$  posee las siguientes propiedades:*

- (1)  $V_\alpha(x) \geq 0$  para todo  $x \in X$ ;
- (2) un vector  $\bar{x} \in \mathbb{R}^n$  es un equilibrio de Nash generalizado si y solo si  $\bar{x} \in X$  y  $V_\alpha(\bar{x}) = 0$ ;
- (3) para cada  $x \in X$  existe un único vector  $y_\alpha(x) = (y_\alpha^1(x), \dots, y_\alpha^N(x)) \in \Omega(x)$  tal que  $V_\alpha(x) = \Psi_\alpha(x, y_\alpha(x))$ .

**PRUEBA.** Los ítems (1) y (3) ya han sido probados, y solo nos falta mostrar el ítem (2). Puesto que toda solución del GNEP es un vector factible podemos suponer que  $\bar{x} \in X$ . Debemos entonces mostrar que  $\bar{x}$  es punto de equilibrio si y solo si  $V_\alpha(\bar{x}) = 0$ .

Si  $\bar{x}$  es solución del GNEP, tenemos  $V(\bar{x}) = 0$  por la Proposición 14.3, y la desigualdad (14.4) muestra que  $V_\alpha(\bar{x}) = 0$ .

Recíprocamente, supongamos que  $V_\alpha(\bar{x}) = 0$ . Dados cualesquiera  $\nu \in \{1, \dots, N\}$  y  $x^\nu \in X_\nu(\bar{x}^{-\nu})$ , queremos ver que  $\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \theta_\nu(x^\nu, \bar{x}^{-\nu}) \leq 0$ . Como  $x^\nu$  y  $\bar{x}^\nu$  están en el conjunto convexo  $X_\nu(\bar{x}^{-\nu})$ , los puntos  $y_t^\nu := t\bar{x}^\nu + (1-t)x^\nu$  del segmento  $[\bar{x}^\nu, x^\nu]$  también están en  $X_\nu(\bar{x}^{-\nu})$ . Procediendo como en la prueba de la Proposición 14.4 obtenemos

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \theta_\nu(y_t^\nu, \bar{x}^{-\nu}) - \frac{\alpha}{2} \|\bar{x}^\nu - y_t^\nu\|^2 \leq 0,$$

y de la concavidad de  $-\theta_\nu(\cdot, \bar{x}^{-\nu})$  se sigue que

$$(1-t)\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - (1-t)\theta_\nu(x^\nu, \bar{x}^{-\nu}) - \frac{\alpha}{2}(1-t)^2\|\bar{x}^\nu - x^\nu\|^2 \leq 0.$$

Esto muestra que  $\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \theta_\nu(x^\nu, \bar{x}^{-\nu}) - \frac{\alpha}{2}(1-t)\|\bar{x}^\nu - x^\nu\|^2 \leq 0$  para todo  $t \in (0, 1)$ , es decir,  $\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \theta_\nu(x^\nu, \bar{x}^{-\nu}) \leq 0$ .  $\square$

Luego, las soluciones del GNEP son los mínimos globales  $\bar{x} \in X$  de  $V_\alpha$  tales que  $V_\alpha(\bar{x}) = 0$ . En consecuencia, el problema de optimización con restricciones

$$\min V_\alpha(x), \quad x \in X. \tag{14.5}$$

es equivalente al GNEP. La Observación 14.4 se aplica aquí con las modificaciones obvias.

Aunque el ítem (3) del Teorema 14.9 indica que esta nueva reformulación supera los defectos **D1.** y **D2.** de la reformulación anterior, no sucede lo mismo con el defecto **D3.** Es decir, cuando las funciones  $\theta_\nu$  son diferenciables, la función objetivo  $V_\alpha$  puede no serlo (véase el Ejemplo 14.11), de modo que muchos de los métodos conocidos para resolver problemas del tipo (14.5) no se pueden usar. En la siguiente sección modificaremos la definición de  $V_\alpha$  para tratar de resolver esta cuestión.

Examinemos más de cerca la función

$$y_\alpha : X \rightarrow \mathbb{R}^n$$

definida en el ítem (3) del Teorema 14.9 que aplica cada  $x \in X$  en el punto  $y_\alpha(x)$  que maximiza la función regularizada  $\Psi_\alpha(x, \cdot)$  sobre  $\Omega(x)$ . Dados  $\bar{x} \in X$  y  $z \in \Omega(\bar{x})$ , la unicidad de  $y_\alpha(\bar{x})$  establecida en dicho ítem nos dice que  $z = y_\alpha(\bar{x})$  si y solo si  $V_\alpha(\bar{x}) = \Psi_\alpha(\bar{x}, z)$ . Además,  $\bar{x} \in \Omega(\bar{x})$  y  $\Psi_\alpha(\bar{x}, \bar{x}) = 0$ . Como por otra parte  $\bar{x}$  es solución del GNEP si y solo si  $V_\alpha(\bar{x}) = 0$ , concluimos que

$$\bar{x} \text{ es solución del GNEP si y solo si } y_\alpha(\bar{x}) = \bar{x}.$$

Así, hemos probado que el GNEP equivale al problema de encontrar los puntos fijos de la función  $y_\alpha$ . Recalcamos este hecho enunciándolo en la siguiente proposición.

**Proposición 14.10.** *Un vector factible  $\bar{x} \in X$  es solución del GNEP si y solo si  $\bar{x}$  es un punto fijo de la aplicación  $y_\alpha$ , es decir, si y solo si  $\bar{x} = y_\alpha(\bar{x})$ .*

**Ejemplo 14.11.** Sea el GNEP

$$\begin{array}{ll} \min_{x^1} x^1 & \min_{x^2} 0 \\ |x^2| - 1 \leq x^1, & |x^2| - 1 \leq x^1, \\ x^1 \leq 0, & x^1 \leq 0, \end{array}$$

con conjunto asociado  $X = \{(x^1, x^2) \mid |x^2| - 1 \leq x^1 \leq 0\}$ , y con funciones de Nikaidô-Isoda

$$\begin{aligned} \Psi(x, y) &= x^1 - y^1, \\ \Psi_\alpha(x, y) &= x^1 - y^1 - \frac{\alpha}{2}(x^1 - y^1)^2 - \frac{\alpha}{2}(x^2 - y^2)^2. \end{aligned}$$

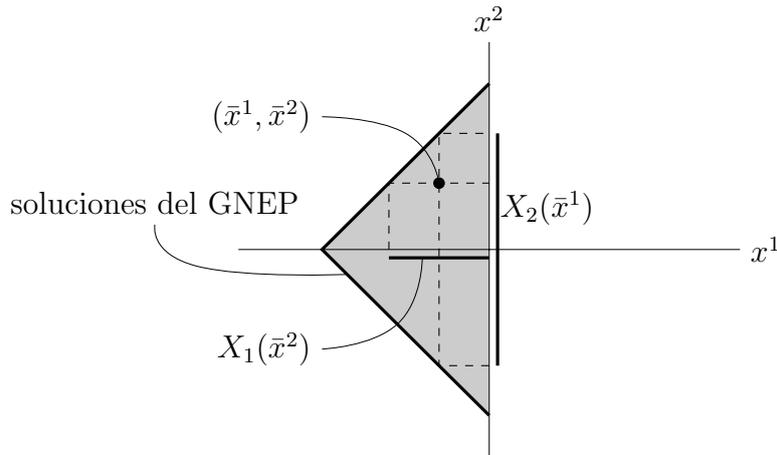


Figura 14.3: Gráfico del Ejemplo 14.11.

Para cada  $\bar{x} = (\bar{x}^1, \bar{x}^2) \in X$ , el conjunto  $\Omega(\bar{x})$  está dado por

$$\Omega(\bar{x}) = X_1(\bar{x}^2) \times X_2(\bar{x}^1) = [|\bar{x}^2| - 1, 0] \times [-\bar{x}^1 - 1, \bar{x}^1 + 1],$$

de manera que

$$V(\bar{x}) = \sup_{(y^1, y^2) \in \Omega(\bar{x})} (\bar{x}^1 - y^1) = \bar{x}^1 - |\bar{x}^2| + 1,$$

y

$$\begin{aligned} V_\alpha(\bar{x}) &= \sup_{(y^1, y^2) \in \Omega(\bar{x})} \left( \bar{x}^1 - y^1 - \frac{\alpha}{2}(\bar{x}^1 - y^1)^2 - \frac{\alpha}{2}(\bar{x}^2 - y^2)^2 \right) \\ &= \max_{|\bar{x}^2| - 1 \leq y^1 \leq 0} \left( \bar{x}^1 - y^1 - \frac{\alpha}{2}(\bar{x}^1 - y^1)^2 \right) - \min_{y^2 \in X_2(\bar{x}^1)} \frac{\alpha}{2}(\bar{x}^2 - y^2)^2, \end{aligned} \quad (14.6)$$

donde el mínimo de la derecha es 0 porque  $\bar{x}^2 \in X_2(\bar{x}^2)$ . Es claro que  $V$  no es diferenciable. Deseamos ver que  $V_\alpha$  tampoco lo es. Para ello, por simplicidad suponemos que  $\alpha = 1$ . Notemos que  $|\bar{x}^2| - 1 \leq y^1 \leq 0$  si y solo si

$$\bar{x}^1 \leq \bar{x}^1 - y^1 \leq \bar{x}^1 - |\bar{x}^2| + 1,$$

donde  $\bar{x}^1 \leq 1$  y  $\bar{x}^1 - |\bar{x}^2| + 1 \leq 1$ . Puesto que la aplicación parabólica  $t \mapsto t - \frac{1}{2}t^2$  definida en  $\mathbb{R}$  posee un máximo global en  $t = 1$ , el máximo de la ecuación (14.6) necesariamente se da cuando  $\bar{x}^1 - y^1 = \bar{x}^1 - |\bar{x}^2| + 1$ , o sea cuando  $y^1 = |\bar{x}^2| - 1$ . Esto significa que

$$y_\alpha(\bar{x}) = (|\bar{x}^2| - 1, \bar{x}^2),$$

y obtenemos en particular que  $y_\alpha$  no es diferenciable. Luego,

$$\begin{aligned} V_\alpha(\bar{x}) &= \bar{x}^1 - (|\bar{x}^2| - 1) - \frac{1}{2} \left( \bar{x}^1 - (|\bar{x}^2| - 1) \right)^2 \\ &= \frac{1}{2} - \frac{1}{2}(\bar{x}^1)^2 - \frac{1}{2}(\bar{x}^2)^2 + \bar{x}^1|\bar{x}^2|, \end{aligned}$$

y concluimos que  $V_\alpha$  tampoco es diferenciable.

## 15. Reformulación diferenciable y con restricciones

En esta sección modificamos el enfoque de la sección anterior para obtener una reformulación del GNEP como un problema de optimización cuya función objetivo será diferenciable (cuando las funciones  $\theta_\nu$  lo son). Partimos de la siguiente observación: como la función  $V_\alpha$  definida en (14.3) satisface la igualdad

$$V_\alpha(x) = \Psi_\alpha(x, y_\alpha(x)),$$

si pudiéramos garantizar que  $y_\alpha$  es diferenciable (cuando los  $\theta_\nu$  lo son) entonces  $V_\alpha$  ya sería diferenciable<sup>1</sup>. Cambiando el conjunto  $\Omega(x)$  por  $X$  en la definición de  $V_\alpha$ , conseguiremos una nueva función “ $y_\alpha$ ” que no será diferenciable, pero que sí será continua, lo cual va a garantizar la diferenciabilidad de la nueva función objetivo. Esto, sin embargo, no viene gratis. El costo que se tiene que pagar es que la reformulación que obtendremos no va a caracterizar a todas las soluciones del GNEP, sino a una subclase importante de ellas: los *equilibrios de Nash normalizados*.

Un vector factible  $\bar{x} \in X$  es un **equilibrio de Nash normalizado** del GNEP si

$$\Psi(\bar{x}, y) \leq 0 \text{ para todo } y \in X,$$

donde  $\Psi$  es la función de Nikaidô-Isoda introducida en la sección anterior. Equivalentemente,  $\bar{x} \in X$  es un equilibrio de Nash normalizado cuando

$$\max_{y \in X} \Psi(\bar{x}, y) = 0.$$

**Proposición 15.1.** *Todo equilibrio de Nash normalizado es solución del GNEP.*

PRUEBA. Supongamos que  $\bar{x}$  es un equilibrio de Nash normalizado, es decir,  $\bar{x} \in X$  y  $\Psi(\bar{x}, y) \leq 0$  para todo  $y \in X$ . Dados  $\nu \in \{1, \dots, N\}$  y  $x^\nu \in X_\nu(\bar{x}^{-\nu})$  el vector  $y := (x^\nu, \bar{x}^{-\nu})$  está en  $X$  y tiene la propiedad de que  $y^\mu = \bar{x}^\mu$  si  $\mu \neq \nu$  y  $y^\nu = x^\nu$ , de modo que

$$\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \theta_\nu(x^\nu, \bar{x}^{-\nu}) = \sum_{\mu=1}^N (\theta_\mu(\bar{x}^\mu, \bar{x}^{-\mu}) - \theta_\mu(y^\mu, \bar{x}^{-\mu})) = \Psi(\bar{x}, y) \leq 0,$$

y por tanto  $\theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) \leq \theta_\nu(x^\nu, \bar{x}^{-\nu})$ . Esto muestra que  $\bar{x}$  es solución del GNEP.  $\square$

El siguiente ejemplo muestra que no todas las soluciones del GNEP son equilibrios de Nash normalizados.

**Ejemplo 15.2.** Revisemos el Ejemplo 14.11. Teníamos allí que  $\theta_1(x^1, x^2) = x^1$ ,  $\theta_2(x^1, x^2) = 0$ ,  $X = \{(x^1, x^2) \mid |x^2| - 1 \leq x^1 \leq 0\}$  y

$$\Psi(x, y) = x^1 - y^1.$$

Las soluciones del GNEP eran los puntos  $(x^1, x^2) \in X$  tales que  $x^1 = |x^2| - 1$ . Ahora bien, un vector factible  $x = (x^1, x^2) \in X$  es un equilibrio de Nash normalizado si y solo si  $x^1 \leq y^1$  para todo  $y = (y^1, y^2) \in X$ , esto es, si y solo si  $x^1 \leq -1$ . Luego, existe un único equilibrio de Nash normalizado, a saber, el punto  $\bar{x} := (-1, 0)$ .

<sup>1</sup>En el caso de  $V_\alpha$ , la función  $y_\alpha$  no es diferenciable (véase el Ejemplo 14.11).

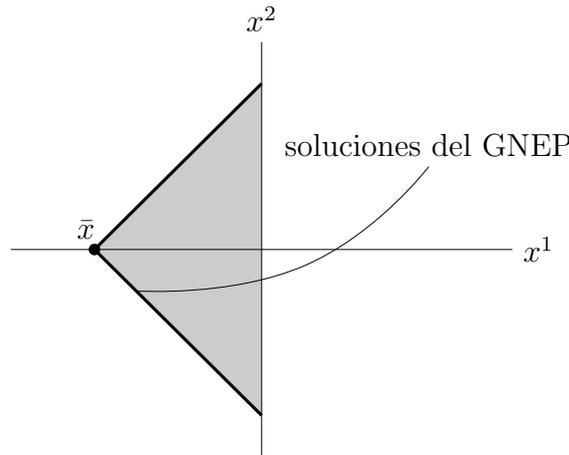


Figura 15.1: Gráfico del Ejemplo 15.2.

Queremos encontrar un problema de minimización, similar a los de la sección previa, que reformule el problema de encontrar los equilibrios de Nash normalizados. Para ello cambiamos  $\Omega(x)$  por  $X$  en (14.3). Para cada vector  $x \in \mathbb{R}^n$  definimos

$$\hat{V}_\alpha(x) = \max_{y \in X} \Psi_\alpha(x, y). \quad (15.1)$$

Obtenemos así una aplicación

$$\hat{V}_\alpha : \mathbb{R}^n \rightarrow \mathbb{R},$$

que está bien definida. De hecho, dado  $x \in \mathbb{R}^n$ , como la función  $\Psi_\alpha(x, \cdot)$  es fuertemente cóncava y el conjunto  $X$  es convexo, cerrado y no vacío, el Corolario 2.31 nos asegura que existe un (único) vector  $y' \in X$  tal que  $\max_{y \in X} \Psi_\alpha(x, y) = \Psi_\alpha(x, y')$ .

**Observación 15.3.** A diferencia de  $V$  y  $V_\alpha$ , la función  $\hat{V}_\alpha$  está definida en todo  $\mathbb{R}^n$ .

La función  $\hat{V}_\alpha$  tiene propiedades similares a las de  $V_\alpha$ , como se ve en el siguiente teorema, que es el resultado análogo al Teorema 14.9 para  $V_\alpha$ .

**Teorema 15.4.** *La función  $\hat{V}_\alpha$  posee las siguientes propiedades:*

- (1)  $\hat{V}_\alpha(x) \geq 0$  para todo  $x \in X$ ;
- (2) un vector  $\bar{x} \in \mathbb{R}^n$  es un equilibrio de Nash normalizado si y solo si  $\bar{x} \in X$  y  $\hat{V}_\alpha(\bar{x}) = 0$ .
- (3) para cada  $x \in \mathbb{R}^n$  existe un único vector  $\hat{y}_\alpha(x) = (\hat{y}_\alpha^1(x), \dots, \hat{y}_\alpha^N(x)) \in X$  tal que  $\Psi_\alpha(x, \hat{y}_\alpha(x)) = \hat{V}_\alpha(x)$ ; además, la aplicación  $\hat{y}_\alpha : x \mapsto \hat{y}_\alpha(x)$  es continua; en particular,  $\hat{V}_\alpha$  es continua.

PRUEBA. (1) Basta con notar que  $\Psi_\alpha(x, x) = 0$ .

(2) Utilizando la parte (1) y recordando que  $\Psi_\alpha(\bar{x}, y) \leq \Psi(\bar{x}, y)$ , deducimos que si  $\bar{x}$  es un equilibrio de Nash normalizado, entonces  $\bar{x} \in X$  y

$$0 \leq \hat{V}_\alpha(\bar{x}) = \max_{y \in X} \Psi_\alpha(\bar{x}, y) \leq \max_{y \in X} \Psi(\bar{x}, y) = 0,$$

es decir,  $\hat{V}_\alpha(\bar{x}) = 0$ .

Recíprocamente, supongamos que  $\bar{x} \in X$  es tal que  $\hat{V}_\alpha(\bar{x}) = 0$ , y tomemos cualquier  $y \in X$ . Entonces los puntos  $t\bar{x} + (1-t)y$  del segmento  $[\bar{x}, y]$  están en  $X$ , y en consecuencia  $\Psi_\alpha(\bar{x}, t\bar{x} + (1-t)y) \leq \hat{V}_\alpha(\bar{x}) = 0$ , esto es,

$$\Psi(\bar{x}, t\bar{x} + (1-t)y) \leq \frac{\alpha}{2} \|\bar{x} - t\bar{x} - (1-t)y\|^2 = \frac{\alpha}{2} (1-t)^2 \|\bar{x} - y\|^2.$$

Luego la concavidad de  $\Psi(\bar{x}, \cdot)$  implica que

$$(1-t)\Psi(\bar{x}, y) = t\Psi(\bar{x}, \bar{x}) + (1-t)\Psi(\bar{x}, y) \leq \frac{\alpha}{2} (1-t)^2 \|\bar{x} - y\|^2,$$

y por tanto  $\Psi(\bar{x}, y) \leq \frac{\alpha}{2} (1-t) \|\bar{x} - y\|^2$  para todo  $t \in (0, 1)$ , es decir,  $\Psi(\bar{x}, y) \leq 0$ . Esto prueba que  $\bar{x}$  es un equilibrio de Nash normalizado.

(3) La existencia y unicidad de  $\hat{y}_\alpha(x)$  ya fueron establecidas justo después de definir  $\hat{V}_\alpha$ . Para establecer la continuidad de  $\hat{y}_\alpha$  usaremos el Teorema 6.6. Hacemos  $f := \Psi_\alpha$ ,  $\Omega(x) := X$  para cada  $x \in \mathbb{R}^n$ , y, según la nomenclatura de dicho teorema, tendremos  $M = \hat{y}_\alpha$ . Luego, es claro que  $f$  es continua, que  $f(x, \cdot)$  es cóncava para cada  $x$ , que  $\Omega$  es cerrado y sci, y que  $M$  es univaluado. Así, en virtud de este teorema concluimos que  $\hat{y}_\alpha$  es una función continua. Finalmente, la continuidad de  $\hat{V}_\alpha$  es una clara consecuencia de la continuidad de  $\hat{y}_\alpha$ .  $\square$

El teorema anterior nos dice que encontrar los equilibrios de Nash normalizados equivale a resolver el siguiente problema de optimización con restricciones

$$\min \hat{V}_\alpha(x), \quad x \in X. \quad (15.2)$$

La Observación 14.4, con obvias modificaciones, también es válida aquí.

Nuestro siguiente objetivo será mostrar que la función continua  $\hat{V}_\alpha$  es diferenciable cuando las funciones  $\theta_\nu$  lo son, y que consecuentemente (15.2) representa un problema de optimización con restricciones y con función objetivo *diferenciable*. Por el momento solo sabemos que  $\hat{V}_\alpha$  es continua, ya que  $\hat{y}_\alpha$  es continua y  $\hat{V}_\alpha(x) = \Psi_\alpha(x, \hat{y}_\alpha(x))$ .

Pero antes de esto, vale la pena enfatizar que la función

$$\hat{y}_\alpha : \mathbb{R}^n \rightarrow X \subset \mathbb{R}^n,$$

que asocia a cada  $x \in \mathbb{R}^n$  con el punto  $\hat{y}_\alpha(x)$  que maximiza la aplicación  $\Psi_\alpha(x, \cdot)$  en  $X$ , está definida en todo  $\mathbb{R}^n$ , y *no solamente* en  $X$ . (La función  $y_\alpha$  estaba definida en  $X$ .) Asimismo, sus puntos fijos son exactamente los equilibrios de Nash normalizados.

**Proposición 15.5.** *Un vector  $\bar{x} \in \mathbb{R}^n$  es un equilibrio de Nash normalizado si y solo si  $\bar{x}$  es un punto fijo de la aplicación  $\hat{y}_\alpha$ , es decir, si y solo si  $\bar{x} = \hat{y}_\alpha(\bar{x})$ .*

PRUEBA. Esta proposición es el resultado análogo a la Proposición 14.10 para  $\hat{y}_\alpha$ . La unicidad de  $\hat{y}_\alpha(\bar{x})$  nos asegura que  $\bar{x} = \hat{y}_\alpha(\bar{x})$  si y solo si  $\bar{x} \in X$  y  $\hat{V}_\alpha(\bar{x}) = \Psi_\alpha(\bar{x}, \bar{x})$ , esto es, si y solo si  $\bar{x} \in X$  y  $\hat{V}_\alpha(\bar{x}) = 0$ , lo cual equivale al hecho de que  $\bar{x}$  sea un equilibrio de Nash normalizado.  $\square$

Volvamos al Problema (15.2). A partir de ahora supondremos, en lo que resta de esta sección, que las funciones  $\theta_\nu$  son de clase  $C^1$ . En particular,  $\Psi_\alpha$  es de clase  $C^1$ . Queremos mostrar que, en esta situación,  $\hat{V}_\alpha$  también es de clase  $C^1$ . Más precisamente, probaremos el siguiente teorema.

**Teorema 15.6.** *La función  $\hat{V}_\alpha$  es diferenciable en  $\mathbb{R}^n$ , y su gradiente está dado por*

$$\begin{aligned} \nabla \hat{V}_\alpha(x) = & \sum_{\nu=1}^N [\nabla \theta_\nu(x^\nu, x^{-\nu}) - \nabla \theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})] \\ & + \begin{pmatrix} \nabla_{x^1} \theta_1(\hat{y}_\alpha^1(x), x^{-1}) \\ \vdots \\ \nabla_{x^N} \theta_N(\hat{y}_\alpha^N(x), x^{-N}) \end{pmatrix} - \alpha(x - \hat{y}_\alpha(x)). \end{aligned} \quad (15.3)$$

En particular,  $\hat{V}_\alpha$  es de clase  $C^1$ .

Usaremos el teorema de Danskin (Teorema 6.7) para demostrar el Teorema 15.6. Con tal fin definimos

$$U := \mathbb{R}^n, \quad K := X \quad \text{y} \quad f := \Psi_\alpha,$$

de modo que, según la nomenclatura del teorema de Danskin, tendremos  $v = \hat{V}_\alpha$  y  $M = \hat{y}_\alpha$ . Entonces  $f$  es continua en  $U \times K$ ;  $\nabla_x f(x, y)$  existe y es continua en  $\Omega \times K$ ; y para cada  $x \in \mathbb{R}^n$ , por la continuidad de  $M$  (parte (3) del Teorema 15.4), existe una vecindad  $V$  de  $x$  tal que  $M$  (como multifunción) tiene imágenes no vacías en  $V$  y  $M(V)$  es acotado. Así, las condiciones de dicho teorema son satisfechas, por lo que deducimos que  $\hat{V}_\alpha$  es diferenciable en  $U = \mathbb{R}^n$  y que su gradiente  $\nabla \hat{V}_\alpha(x)$  es igual a  $\nabla_x \Psi_\alpha(x, y)|_{y=\hat{y}_\alpha(x)}$ , o sea,

$$\nabla \hat{V}_\alpha(x) = \nabla_x \Psi_\alpha(x, \hat{y}_\alpha(x)).$$

Luego el Teorema 15.6 se sigue del siguiente lema.

**Lema 15.7.** *El gradiente respecto a  $x$  de  $\Psi_\alpha$  está dado por*

$$\nabla_x \Psi_\alpha(x, y) = \sum_{\nu=1}^N [\nabla \theta_\nu(x^\nu, x^{-\nu}) - \nabla \theta_\nu(y^\nu, x^{-\nu})] + \begin{pmatrix} \nabla_{x^1} \theta_1(y^1, x^{-1}) \\ \vdots \\ \nabla_{x^N} \theta_N(y^N, x^{-N}) \end{pmatrix} - \alpha(x - y).$$

PRUEBA. Escribiendo

$$\Psi_\alpha(x, y) = \eta(x, y) - \omega(x, y) - \frac{\alpha}{2} \|x - y\|^2,$$

donde  $\eta(x, y) = \sum_{\nu=1}^N \theta_\nu(x^\nu, x^{-\nu})$  y  $\omega(x, y) = \sum_{\nu=1}^N \theta_\nu(y^\nu, x^{-\nu})$ , vemos que el gradiente respecto a  $x$  del último sumando es  $-\alpha(x - y)$ , y que además

$$\nabla_x \eta(x, y) = \sum_{\nu=1}^N \nabla \theta_\nu(x^\nu, x^{-\nu}).$$

Debemos entonces mostrar que

$$\nabla_x \omega(x, y) = \sum_{\nu=1}^N \nabla \theta_\nu(y^\nu, x^{-\nu}) - \begin{pmatrix} \nabla_{x^1} \theta_1(y^1, x^{-1}) \\ \vdots \\ \nabla_{x^N} \theta_N(y^N, x^{-N}) \end{pmatrix}.$$

Puesto que cada  $x^\mu$  representa las  $n_\mu$  variables del  $\mu$ -ésimo jugador, basta con mostrar que si  $x_i$  es una variable controlada por el  $\mu$ -ésimo jugador entonces

$$\frac{\partial \omega}{\partial x_i}(x, y) = \sum_{\nu=1}^N \frac{\partial \theta_\nu}{\partial x_i}(y^\nu, x^{-\nu}) - \frac{\partial \theta_\mu}{\partial x_i}(y^\mu, x^{-\mu}). \quad (15.4)$$

Sea  $\mu \in \{1, \dots, N\}$  y supongamos que  $x_i$  es una variable controlada por el jugador  $\mu$ . Sean las funciones  $\gamma_\nu : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$  definidas por  $\gamma_\nu(x, y) = (y^\nu, x^{-\nu})$ . Luego  $\omega = \sum_{\nu=1}^N \theta_\nu \circ \gamma_\nu$ , y puesto que la aplicación  $\theta_\mu \circ \gamma_\mu : (x, y) \mapsto \theta_\mu(y^\mu, x^{-\mu})$  no depende de las variables  $x^\mu$  controladas por el  $\mu$ -ésimo jugador, tenemos

$$\frac{\partial \omega}{\partial x_i}(x, y) = \sum_{\nu=1}^N \frac{\partial(\theta_\nu \circ \gamma_\nu)}{\partial x_i}(x, y) = \sum_{\nu=1, \nu \neq \mu}^N \frac{\partial(\theta_\nu \circ \gamma_\nu)}{\partial x_i}(x, y).$$

Escribiendo  $\gamma_\nu = ((\gamma_\nu)_1, \dots, (\gamma_\nu)_n) : (x, y) \mapsto (y^\nu, x^{-\nu})$ , vemos que  $\frac{\partial(\gamma_\nu)_i}{\partial x_i}(x, y) = \delta_{ij}$  si  $\nu \neq \mu$  (donde  $\delta_{ij}$  es el delta de Kronecker), de modo que aplicando la Regla de la Cadena (Teorema 1.21) obtenemos

$$\begin{aligned} \sum_{\nu=1, \nu \neq \mu}^N \frac{\partial(\theta_\nu \circ \gamma_\nu)}{\partial x_i}(x, y) &= \sum_{\nu=1, \nu \neq \mu}^N \sum_{j=1}^n \frac{\partial \theta_\nu}{\partial x_j}(\gamma_\nu(x, y)) \cdot \delta_{ij} = \sum_{\nu=1, \nu \neq \mu}^N \frac{\partial \theta_\nu}{\partial x_i}(y^\nu, x^{-\nu}) \\ &= \sum_{\nu=1}^N \frac{\partial \theta_\nu}{\partial x_i}(y^\nu, x^{-\nu}) - \frac{\partial \theta_\mu}{\partial x_i}(y^\mu, x^{-\mu}), \end{aligned}$$

lo cual prueba (15.4). □

## Puntos estacionarios

Ahora que sabemos que el problema

$$\min \hat{V}_\alpha(x), \quad x \in X \quad (15.2)$$

supera los defectos **D1.**, **D2.** y **D3.** que enumeramos en la sección anterior, recordemos que la mayoría de algoritmos usados en optimización para resolver este tipo de problemas sirven para hallar sus puntos estacionarios (véase la página 42). Así, surge una pregunta natural: ¿cuándo un punto estacionario del Problema (15.2) es una solución óptima?

Debido a que estamos interesados en los equilibrios de Nash normalizados, que constituyen una subclase de las soluciones óptimas del Problema (15.2), aquí intentaremos responder la siguiente pregunta: ¿cuándo un punto estacionario del Problema (15.2) es un equilibrio de Nash normalizado?

**Lema 15.8.** *Dado cualquier  $z \in \mathbb{R}^n$  tenemos*

$$\sum_{\nu=1}^N [\nabla_{x^\nu} \theta_\nu(\hat{y}_\alpha^\nu(z), z^{-\nu}) - \alpha(z^\nu - \hat{y}_\alpha^\nu(z))]^T (x^\nu - \hat{y}_\alpha^\nu(z)) \geq 0, \quad \forall x \in X.$$

*En particular, si  $z \in X$  la siguiente desigualdad se verifica*

$$\sum_{\nu=1}^N [\nabla_{x^\nu} \theta_\nu(\hat{y}_\alpha^\nu(z), z^{-\nu}) - \alpha(z^\nu - \hat{y}_\alpha^\nu(z))]^T (z^\nu - \hat{y}_\alpha^\nu(z)) \geq 0.$$

PRUEBA. Recordemos que  $\hat{y}_\alpha(z)$  es la única solución del problema

$$\min \phi(y), \quad y \in X,$$

donde  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$  es la función definida por  $\phi(y) = -\Psi_\alpha(z, y)$ . Esto implica que  $\hat{y}_\alpha(z)$  debe satisfacer la condición de optimalidad dada en la Proposición 3.1:

$$\langle \nabla \phi(\hat{y}_\alpha(z)), x - \hat{y}_\alpha(z) \rangle \geq 0, \quad \forall x \in X. \quad (15.5)$$

De la definición de  $\phi$  vemos que

$$\phi(x) = -\sum_{\nu=1}^N \theta_\nu(z^\nu, z^{-\nu}) + \sum_{\nu=1}^N \theta_\nu(x^\nu, z^{-\nu}) + \frac{\alpha}{2} \|x - z\|^2.$$

El primer sumando  $-\sum_{\nu=1}^N \theta_\nu(z^\nu, z^{-\nu})$  de esta igualdad es constante, y, además, el gradiente del tercer sumando es igual a  $\alpha(x - z)$ . Afirmamos que

$$\nabla \varphi(x) = \begin{pmatrix} \nabla_{x^1} \theta_1(x^1, z^{-1}) \\ \vdots \\ \nabla_{x^N} \theta_N(x^N, z^{-N}) \end{pmatrix},$$

donde  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$  está dada por  $\varphi(x) = \sum_{\nu=1}^N \theta_\nu(x^\nu, z^{-\nu})$ . Supongamos por el momento que la afirmación es cierta. Tenemos entonces que

$$\nabla \phi(x) = \begin{pmatrix} \nabla_{x^1} \theta_1(x^1, z^{-1}) \\ \vdots \\ \nabla_{x^N} \theta_N(x^N, z^{-N}) \end{pmatrix} + \alpha(x - z) = \begin{pmatrix} \nabla_{x^1} \theta_1(x^1, z^{-1}) - \alpha(z^1 - x^1) \\ \vdots \\ \nabla_{x^N} \theta_N(x^N, z^{-N}) - \alpha(z^N - x^N) \end{pmatrix}.$$

Usando esta expresión, la ecuación (15.5) se convierte en

$$\begin{pmatrix} \nabla_{x^1} \theta_1(\hat{y}_\alpha^1(z), z^{-1}) - \alpha(z^1 - \hat{y}_\alpha^1(z)) \\ \vdots \\ \nabla_{x^N} \theta_N(\hat{y}_\alpha^N(z), z^{-N}) - \alpha(z^N - \hat{y}_\alpha^N(z)) \end{pmatrix}^T (x - \hat{y}_\alpha(z)) \geq 0, \quad \forall x \in X,$$

de donde, desarrollando el producto de matrices obtenemos la primera parte del lema. La segunda parte es obvia.

Solo nos falta probar la afirmación. Debemos mostrar que si  $x_i$  es una variable controlada por el  $\mu$ -ésimo jugador entonces

$$\frac{\partial \varphi}{\partial x_i}(x) = \frac{\partial \theta_\mu}{\partial x_i}(x^\mu, z^{-\mu}), \quad \forall x \in \mathbb{R}^n. \quad (15.6)$$

Para cada  $\nu \in \{1, \dots, N\}$  sea la función  $\gamma_\nu : \mathbb{R}^n \rightarrow \mathbb{R}^n$  dada por  $\gamma_\nu(x) = (x^\nu, z^{-\nu})$ . Entonces  $\varphi = \sum_{\nu=1}^N \theta_\nu \circ \gamma_\nu$ , y como la aplicación  $\theta_\nu \circ \gamma_\nu : x \mapsto \theta_\nu(x^\nu, z^{-\nu})$  no depende de  $x_i$  cuando  $\nu \neq \mu$ , se sigue que

$$\frac{\partial \varphi}{\partial x_i}(x) = \sum_{\nu=1}^N \frac{\partial (\theta_\nu \circ \gamma_\nu)}{\partial x_i}(x) = \frac{\partial (\theta_\mu \circ \gamma_\mu)}{\partial x_i}(x).$$

Escribiendo  $\gamma_\mu = ((\gamma_\mu)_1, \dots, (\gamma_\mu)_n) : x \mapsto (x^\mu, z^{-\mu})$ , tenemos que  $\frac{\partial (\gamma_\mu)_j}{\partial x_i}(x) = \delta_{ij}$ , donde  $\delta_{ij}$  es el delta de Kronecker. La Regla de la Cadena implica por tanto que

$$\frac{\partial (\theta_\mu \circ \gamma_\mu)}{\partial x_i}(x) = \sum_{j=1}^n \frac{\partial \theta_\mu}{\partial x_j}(\gamma_\mu(x)) \cdot \delta_{ij} = \frac{\partial \theta_\mu}{\partial x_i}(x^\mu, z^{-\mu}),$$

y combinando esto con la ecuación anterior obtenemos (15.6).  $\square$

Sea ahora  $\bar{x} \in X$  un punto estacionario del Problema (15.2), esto es,

$$\nabla \hat{V}_\alpha(\bar{x})^T(x - \bar{x}) \geq 0, \quad \forall x \in X.$$

De la representación del gradiente  $\nabla \hat{V}_\alpha(\bar{x})$  en el Teorema 15.6, obtenemos que para  $x = \hat{y}_\alpha(\bar{x}) \in X$  se cumple que

$$\begin{aligned} 0 &\leq \sum_{\nu=1}^N [\nabla \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \nabla \theta_\nu(\hat{y}_\alpha^\nu(\bar{x}), \bar{x}^{-\nu})]^T (\hat{y}_\alpha(\bar{x}) - \bar{x}) \\ &\quad + \sum_{\nu=1}^N \nabla_{x^\nu} \theta_\nu(\hat{y}_\alpha^\nu(\bar{x}), \bar{x}^{-\nu})^T (\hat{y}_\alpha(\bar{x})^\nu - \bar{x}^\nu) - \alpha (\bar{x} - \hat{y}_\alpha(\bar{x}))^T (\hat{y}_\alpha(\bar{x}) - \bar{x}) \\ &= \sum_{\nu=1}^N [\nabla \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \nabla \theta_\nu(\hat{y}_\alpha^\nu(\bar{x}), \bar{x}^{-\nu})]^T (\hat{y}_\alpha(\bar{x}) - \bar{x}) \\ &\quad + \sum_{\nu=1}^N [\nabla_{x^\nu} \theta_\nu(\hat{y}_\alpha^\nu(\bar{x}), \bar{x}^{-\nu}) - \alpha (\bar{x}^\nu - \hat{y}_\alpha^\nu(\bar{x}))]^T (\hat{y}_\alpha(\bar{x})^\nu - \bar{x}^\nu). \end{aligned}$$

Ahora bien, el Lema 15.8 nos asegura que el segundo término de esta suma es menor o igual que 0, y que por tanto el primer sumando es mayor o igual que 0.

En conclusión, para todo punto estacionario  $\bar{x} \in X$  del Problema (15.2) tenemos

$$\sum_{\nu=1}^N [\nabla \theta_\nu(\bar{x}^\nu, \bar{x}^{-\nu}) - \nabla \theta_\nu(\hat{y}_\alpha^\nu(\bar{x}), \bar{x}^{-\nu})]^T (\hat{y}_\alpha(\bar{x}) - \bar{x}) \geq 0.$$

Esto motiva la siguiente suposición, que representa una respuesta parcial a la pregunta planteada al inicio de esta subsección.

**Suposición 15.9.** Dado  $x \in \mathbb{R}^n$  fijo con  $x \neq \hat{y}_\alpha(x)$ , se cumple que

$$\sum_{\nu=1}^N [\nabla\theta_\nu(x^\nu, x^{-\nu}) - \nabla\theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})]^T (x - \hat{y}_\alpha(x)) > 0.$$

En otras palabras, la Suposición 15.9 se satisface en  $x \in \mathbb{R}^n$  cuando

$$x \neq \hat{y}_\alpha(x) \quad \text{implica} \quad \sum_{\nu=1}^N [\nabla\theta_\nu(x^\nu, x^{-\nu}) - \nabla\theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})]^T (x - \hat{y}_\alpha(x)) > 0.$$

**Teorema 15.10.** Sea  $\bar{x} \in X$  un punto estacionario del Problema (15.2). Si la Suposición 15.9 se cumple en  $x = \bar{x}$ , entonces  $\bar{x}$  es un equilibrio de Nash normalizado para el GNEP.

PRUEBA. La discusión previa muestra que si  $\bar{x}$  es un punto estacionario en el que la Suposición 15.9 se satisface, entonces  $\bar{x} = \hat{y}_\alpha(\bar{x})$ , o equivalentemente,  $\bar{x}$  es un equilibrio de Nash normalizado.  $\square$

La Suposición 15.9 puede interpretarse como un cierto tipo de monotonía estricta o positividad definida. Para apreciar este hecho supongamos que las funciones  $\theta_\nu$  son cuadráticas y están definidas de la siguiente manera

$$\theta_\nu(x) = (x^\nu)^T A_{\nu\nu} x^\nu + \sum_{\mu=1, \mu \neq \nu}^N (x^\nu)^T A_{\nu\mu} x^\mu,$$

donde cada matriz  $A_{\nu\mu} \in \mathbb{R}^{n_\nu \times n_\mu}$  es el  $(\nu, \mu)$  subbloque de la matriz

$$A := (A_{\nu\mu})_{\nu, \mu=1}^N \in \mathbb{R}^{n \times n}.$$

Suponemos también que las matrices  $A_{\nu\mu}$  son simétricas.

**Proposición 15.11.** Si la matriz  $A$  es positiva definida, entonces la Suposición 15.9 se satisface en todos los puntos  $x \in \mathbb{R}^n$ .

PRUEBA. Usando la identidad  $(x^\nu)^T A_{\nu\mu} x^\mu = (x^\mu)^T A_{\nu\mu}^T x^\nu = (x^\mu)^T A_{\nu\mu} x^\nu$ , derivamos directamente las funciones  $\theta_\nu$  para obtener

$$\nabla_{x^\mu} \theta_\nu(x^\nu, x^{-\nu}) = A_{\nu\mu} x^\nu \quad \text{si } \mu \neq \nu$$

y

$$\nabla_{x^\nu} \theta_\nu(x^\nu, x^{-\nu}) = 2A_{\nu\nu} x^\nu + \sum_{\mu=1, \mu \neq \nu}^N A_{\nu\mu} x^\mu = A_{\nu\nu} x^\nu + \sum_{\mu=1}^N A_{\nu\mu} x^\mu.$$

Luego

$$\sum_{\nu=1}^N [\nabla\theta_\nu(x^\nu, x^{-\nu}) - \nabla\theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})]^T (x - \hat{y}_\alpha(x))$$

$$\begin{aligned}
 &= \sum_{\nu=1}^N \sum_{\mu=1}^N [\nabla_{x^\mu} \theta_\nu(x^\nu, x^{-\nu}) - \nabla_{x^\mu} \theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})]^T (x^\mu - \hat{y}_\alpha^\mu(x)) \\
 &= \sum_{\mu \neq \nu} [\dots] + \sum_{\mu=\nu} [\dots] \\
 &= 2 \sum_{\nu, \mu=1}^N (x^\nu - \hat{y}_\alpha^\nu(x))^T A_{\nu\mu} (x^\nu - \hat{y}_\alpha^\mu(x)) \\
 &= 2(x - \hat{y}_\alpha(x))^T A(x - \hat{y}_\alpha(x)),
 \end{aligned}$$

y consecuentemente  $x \neq \hat{y}_\alpha(x)$  implica que este número es positivo.  $\square$

**Observación 15.12.** Notemos que a fin de obtener la reformulación del GNEP con  $\hat{V}_\alpha$ , lo que utilizamos esencialmente fue el hecho de que la aplicación  $y \mapsto \Psi_\alpha(x, y)$  era fuertemente cóncava para cada  $x \in \mathbb{R}^n$  fijo (véanse los Teoremas 15.4 y 15.6 y la Proposición 15.5). Por ello, si  $y \mapsto \Psi(x, y)$  fuera fuertemente cóncava para cada  $x$ , obtendríamos una reformulación similar sin utilizar la función de Nikaidô-Isoda regularizada, a saber, mediante la función

$$\bar{V}(x) := \max_{y \in X} \Psi(x, y),$$

que tendría las mismas propiedades que  $\hat{V}_\alpha$ . Ahora bien, una condición suficiente para que esto ocurra (para que  $\Psi(x, \cdot)$  sea fuertemente cóncava para cada  $x$ ) es que las funciones de beneficio  $\theta_\nu$  sean fuertemente convexas en  $x^\nu$ , como se verifica fácilmente. Luego, para el GNEP en el “caso fuertemente convexo” no necesitamos regularizar la función de Nikaidô-Isoda para obtener los resultados deseados.

**Observación 15.13.** De las definiciones dadas en las ecuaciones (14.3) y (15.1), vemos que la única diferencia entre  $V_\alpha$  y  $\hat{V}_\alpha$  es que, al definir las, se emplean los conjuntos  $\Omega(x)$  y  $X$  respectivamente. Por ello, si nos restringimos al NEP, esto es, si  $X = X_1 \times \dots \times X_N$  donde cada  $X_\nu \subset \mathbb{R}^{n_\nu}$  es convexo, cerrado y no vacío, es fácil ver que  $\Omega(x) = X$  para cada  $x \in X$ . Luego,  $V_\alpha = \hat{V}_\alpha$  y en particular  $V_\alpha$  es de clase  $C^1$ . Así, para el NEP las reformulaciones (14.5) y (15.2) coinciden.

## 16. Reformulación diferenciable y sin restricciones

En esta sección adaptamos las ideas de la sección anterior para obtener una reformulación del GNEP como un problema de optimización *sin* restricciones. La idea principal será usar dos parámetros  $\alpha$  y  $\beta$  distintos y trabajar con la diferencia entre sus funciones  $\hat{V}$  correspondientes, para luego aplicar los resultados de la sección anterior. La reformulación y sus propiedades irán surgiendo de manera natural de lo que ya sabemos de la sección previa. Desde luego, la reformulación de esta sección no va a caracterizar a todas las soluciones del GNEP, sino solamente a los equilibrios de Nash normalizados.

Comenzamos estableciendo el contexto bajo el que trabajaremos. Sean los parámetros  $\alpha, \beta \in \mathbb{R}$  tales que  $0 < \alpha < \beta$ , y consideremos las funciones regularizadas de Nikaidô-Isoda correspondientes

$$\begin{aligned}\Psi_\alpha(x, y) &= \sum_{\nu=1}^N [\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) - \frac{\alpha}{2} \|x^\nu - y^\nu\|^2] = \Psi(x, y) - \frac{\alpha}{2} \|x - y\|^2, \\ \Psi_\beta(x, y) &= \sum_{\nu=1}^N [\theta_\nu(x^\nu, x^{-\nu}) - \theta_\nu(y^\nu, x^{-\nu}) - \frac{\beta}{2} \|x^\nu - y^\nu\|^2] = \Psi(x, y) - \frac{\beta}{2} \|x - y\|^2;\end{aligned}$$

sean también

$$\begin{aligned}\hat{V}_\alpha(x) &= \max_{y \in X} \Psi_\alpha(x, y) = \Psi_\alpha(x, \hat{y}_\alpha(x)), \\ \hat{V}_\beta(x) &= \max_{y \in X} \Psi_\beta(x, y) = \Psi_\beta(x, \hat{y}_\beta(x)),\end{aligned}\tag{16.1}$$

como en la ecuación (15.1) y según el Teorema 15.4. Como ya vimos, estas dos funciones están bien definidas en todo  $\mathbb{R}^n$ . Luego, podemos definir a partir de ellas una nueva función

$$\hat{V}_{\alpha\beta}(x) := \hat{V}_\alpha(x) - \hat{V}_\beta(x), \quad x \in \mathbb{R}^n,\tag{16.2}$$

que es continua por el Teorema 15.4. Como veremos más adelante, esta será la función objetivo de nuestro problema de optimización sin restricciones.

A continuación establecemos un resultado preliminar.

**Lema 16.1.** *Para cada  $x \in \mathbb{R}^n$  se verifica*

$$\frac{\beta - \alpha}{2} \|x - \hat{y}_\beta(x)\|^2 \leq \hat{V}_{\alpha\beta}(x) \leq \frac{\beta - \alpha}{2} \|x - \hat{y}_\alpha(x)\|^2.$$

PRUEBA. De la ecuación (16.1), tenemos

$$\Psi_\beta(x, \hat{y}_\alpha(x)) \leq \hat{V}_\beta(x) = \Psi_\beta(x, \hat{y}_\beta(x)) \quad \text{y} \quad \Psi_\alpha(x, \hat{y}_\beta(x)) \leq \hat{V}_\alpha(x) = \Psi_\alpha(x, \hat{y}_\alpha(x)),$$

de donde se sigue

$$\begin{aligned}\frac{\beta - \alpha}{2} \|x - \hat{y}_\beta(x)\|^2 &= \Psi_\alpha(x, \hat{y}_\beta(x)) - \Psi_\beta(x, \hat{y}_\beta(x)) \leq \hat{V}_\alpha(x) - \hat{V}_\beta(x) \\ &\leq \Psi_\alpha(x, \hat{y}_\alpha(x)) - \Psi_\beta(x, \hat{y}_\alpha(x)) = \frac{\beta - \alpha}{2} \|x - \hat{y}_\alpha(x)\|^2. \quad \square\end{aligned}$$

Este lema nos permite obtener un resultado análogo al Teorema 15.4 para la función  $\hat{V}_{\alpha\beta}$ , que caracteriza los equilibrios de Nash normalizados.

**Teorema 16.2.** *La función  $\hat{V}_{\alpha\beta}$  satisface las siguientes propiedades:*

- (1)  $\hat{V}_{\alpha\beta}(x) \geq 0$  para todo  $x \in \mathbb{R}^n$ ;
- (2) un vector  $\bar{x} \in \mathbb{R}^n$  es un equilibrio de Nash normalizado si y solo si  $\bar{x}$  es un mínimo global de  $\hat{V}_{\alpha\beta}$  tal que  $\hat{V}_{\alpha\beta}(\bar{x}) = 0$ .

PRUEBA. Es evidente que la parte (1) se sigue del Lema 16.1. Para la parte (2), utilizando la Proposición 15.5 para  $\alpha$  y  $\beta$ , tenemos que  $\bar{x} \in \mathbb{R}^n$  es un equilibrio de Nash normalizado si y solo si  $\bar{x} = \hat{y}_\beta(\bar{x})$  o  $\bar{x} = \hat{y}_\alpha(\bar{x})$ . Luego, basta con notar que, por el Lema 16.1,  $\bar{x} = \hat{y}_\alpha(\bar{x})$  implica  $\hat{V}_{\alpha\beta}(\bar{x}) = 0$ , y que  $\hat{V}_{\alpha\beta}(\bar{x}) = 0$  implica  $\bar{x} = \hat{y}_\beta(\bar{x})$ .  $\square$

Así, encontrar los equilibrios de Nash normalizados equivale a resolver el problema de minimización *sin* restricciones

$$\min \hat{V}_{\alpha\beta}(x), \quad x \in \mathbb{R}^n. \quad (16.3)$$

Nuevamente, la Observación 14.4 es válida aquí con las respectivas modificaciones.

Como nuestro objetivo es dar una reformulación con función objetivo diferenciable, de aquí en adelante vamos a suponer que las funciones  $\theta_\nu$  son de clase  $C^1$ . Como  $\hat{V}_{\alpha\beta} = \hat{V}_\alpha - \hat{V}_\beta$ , de acuerdo con el Teorema 15.6, la función  $\hat{V}_{\alpha\beta}$  también será de clase  $C^1$ , y su gradiente será la diferencia de los gradientes de  $\hat{V}_\alpha$  y  $\hat{V}_\beta$ .

**Teorema 16.3.** *La función  $\hat{V}_{\alpha\beta}$  es de clase  $C^1$  y su gradiente está dado por*

$$\begin{aligned} \nabla \hat{V}_{\alpha\beta}(x) = & \sum_{\nu=1}^N [\nabla \theta_\nu(\hat{y}_\beta^\nu(x), x^{-\nu}) - \nabla \theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})] \\ & + \begin{pmatrix} \nabla_{x^1} \theta_1(\hat{y}_\alpha^1(x), x^{-1}) - \nabla_{x^1} \theta_1(\hat{y}_\beta^1(x), x^{-1}) \\ \vdots \\ \nabla_{x^N} \theta_N(\hat{y}_\alpha^N(x), x^{-N}) - \nabla_{x^N} \theta_N(\hat{y}_\beta^N(x), x^{-N}) \end{pmatrix} \\ & - \alpha(x - \hat{y}_\alpha(x)) + \beta(x - \hat{y}_\beta(x)). \end{aligned}$$

El Teorema 16.3 implica en particular que el Problema (16.3) es un problema de optimización *sin* restricciones y con función objetivo *diferenciable*.

## Puntos estacionarios

Como se hizo notar en la sección anterior, es importante saber cuándo un punto estacionario de un problema de optimización es una solución óptima (véase la página 42). En nuestro caso, es importante saber cuándo un punto estacionario del Problema (16.3) es un equilibrio de Nash normalizado. Dado que (16.3) es un problema *sin* restricciones, sus puntos estacionarios son aquellos vectores  $\bar{x} \in \mathbb{R}^n$  tales que  $\nabla \hat{V}_{\alpha\beta}(\bar{x}) = 0$ .

**Lema 16.4.** *Para cada  $z \in \mathbb{R}^n$  se cumple que*

$$\begin{aligned} \sum_{\nu=1}^N [\nabla_{x^\nu} \theta_\nu(\hat{y}_\alpha^\nu(z), z^{-\nu}) - \nabla_{x^\nu} \theta_\nu(\hat{y}_\beta^\nu(z), z^{-\nu}) \\ - \alpha(z^\nu - \hat{y}_\alpha^\nu(z)) + \beta(z^\nu - \hat{y}_\beta^\nu(z))]^T (\hat{y}_\beta^\nu(z) - \hat{y}_\alpha^\nu(z)) \geq 0. \end{aligned}$$

PRUEBA. Sabemos del Lema 15.8 que

$$\sum_{\nu=1}^N [\nabla_{x^\nu} \theta_\nu(\hat{y}_\alpha^\nu(z), z^{-\nu}) - \alpha(z^\nu - \hat{y}_\alpha^\nu(z))]^T (x^\nu - \hat{y}_\alpha^\nu(z)) \geq 0, \quad \forall x \in X,$$

y que en particular para  $x = \hat{y}_\beta^\nu(z)$

$$\sum_{\nu=1}^N [\nabla_{x^\nu} \theta_\nu(\hat{y}_\alpha^\nu(z), z^{-\nu}) - \alpha(z^\nu - \hat{y}_\alpha^\nu(z))]^T (\hat{y}_\beta^\nu(z) - \hat{y}_\alpha^\nu(z)) \geq 0.$$

Análogamente,

$$\sum_{\nu=1}^N [\nabla_{x^\nu} \theta_\nu(\hat{y}_\beta^\nu(z), z^{-\nu}) - \beta(z^\nu - \hat{y}_\beta^\nu(z))]^T (\hat{y}_\alpha^\nu(z) - \hat{y}_\beta^\nu(z)) \geq 0.$$

Sumando estas dos desigualdades obtenemos lo que buscábamos.  $\square$

Ahora daremos una condición similar a la Suposición 15.9, que representará una respuesta parcial a la cuestión de bajo qué condiciones un punto estacionario es un equilibrio de Nash normalizado.

**Suposición 16.5.** Dado  $x \in \mathbb{R}^n$  fijo con  $\hat{y}_\alpha(x) \neq \hat{y}_\beta(x)$ , la siguiente desigualdad se satisface

$$\sum_{\nu=1}^N [\nabla \theta_\nu(\hat{y}_\beta^\nu(x), x^{-\nu}) - \nabla \theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})]^T (\hat{y}_\beta(x) - \hat{y}_\alpha(x)) > 0.$$

Equivalentemente, la Suposición 16.5 se satisface en  $x \in \mathbb{R}^n$  cuando

$$\hat{y}_\alpha(x) \neq \hat{y}_\beta(x) \text{ implica } \sum_{\nu=1}^N [\nabla \theta_\nu(\hat{y}_\beta^\nu(x), x^{-\nu}) - \nabla \theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})]^T (\hat{y}_\beta(x) - \hat{y}_\alpha(x)) > 0.$$

Con esta suposición podemos establecer el resultado análogo al Teorema 15.10 para el problema (16.3).

**Teorema 16.6.** Sea  $\bar{x} \in \mathbb{R}^n$  un punto estacionario del Problema (16.3), esto es,

$$\nabla \hat{V}_{\alpha\beta}(\bar{x}) = 0.$$

Si la Suposición 16.5 se satisface en  $x = \bar{x}$ , entonces  $\bar{x}$  es un equilibrio de Nash normalizado.

PRUEBA. Sabemos que  $\nabla \hat{V}_{\alpha\beta}(\bar{x}) = 0$ . Del Teorema 16.3 se sigue que

$$\begin{aligned} 0 &= \sum_{\nu=1}^N [\nabla \theta_\nu(\hat{y}_\beta^\nu(\bar{x}), \bar{x}^{-\nu}) - \nabla \theta_\nu(\hat{y}_\alpha^\nu(\bar{x}), \bar{x}^{-\nu})] \\ &\quad + \begin{pmatrix} \nabla_{x^1} \theta_1(\hat{y}_\alpha^1(\bar{x}), \bar{x}^{-1}) - \nabla_{x^1} \theta_1(\hat{y}_\beta^1(\bar{x}), \bar{x}^{-1}) \\ \vdots \\ \nabla_{x^N} \theta_N(\hat{y}_\alpha^N(\bar{x}), \bar{x}^{-N}) - \nabla_{x^N} \theta_N(\hat{y}_\beta^N(\bar{x}), \bar{x}^{-N}) \end{pmatrix} \\ &\quad - \alpha(\bar{x} - \hat{y}_\alpha(\bar{x})) + \beta(\bar{x} - \hat{y}_\beta(\bar{x})). \end{aligned} \tag{16.4}$$

Multiplicando por  $(\hat{y}_\beta(\bar{x}) - \hat{y}_\alpha(\bar{x}))$  obtenemos

$$\begin{aligned} 0 = & \sum_{\nu=1}^N [\nabla\theta_\nu(\hat{y}_\beta^\nu(\bar{x}), \bar{x}^{-\nu}) - \nabla\theta_\nu(\hat{y}_\alpha^\nu(\bar{x}), \bar{x}^{-\nu})]^T (\hat{y}_\beta(\bar{x}) - \hat{y}_\alpha(\bar{x})) \\ & + \sum_{\nu=1}^N [\nabla_{x^\nu}\theta_\nu(\hat{y}_\alpha^\nu(\bar{x}), \bar{x}^{-\nu}) - \nabla_{x^\nu}\theta_\nu(\hat{y}_\beta^\nu(\bar{x}), \bar{x}^{-\nu}) \\ & \quad - \alpha(\bar{x}^\nu - \hat{y}_\alpha^\nu(\bar{x})) + \beta(\bar{x}^\nu - \hat{y}_\beta^\nu(\bar{x}))]^T (\hat{y}_\beta^\nu(\bar{x}) - \hat{y}_\alpha^\nu(\bar{x})), \end{aligned}$$

donde el segundo sumando es mayor o igual que 0 por el Lema 16.4, es decir,

$$\sum_{\nu=1}^N [\nabla\theta_\nu(\hat{y}_\beta^\nu(\bar{x}), \bar{x}^{-\nu}) - \nabla\theta_\nu(\hat{y}_\alpha^\nu(\bar{x}), \bar{x}^{-\nu})]^T (\hat{y}_\beta(\bar{x}) - \hat{y}_\alpha(\bar{x})) \leq 0,$$

Como la Suposición 16.5 se satisface en  $x = \bar{x}$ , lo anterior significa que  $\hat{y}_\beta(\bar{x}) = \hat{y}_\alpha(\bar{x})$ . Así,  $\bar{x}$  es un equilibrio de Nash normalizado por la Proposición 15.5, ya que de (16.4) se sigue que  $(\beta - \alpha)(\bar{x} - \hat{y}_\alpha(\bar{x})) = 0$  y por consiguiente  $\bar{x} = \hat{y}_\alpha(\bar{x})$ .  $\square$

Terminamos analizando la Suposición 16.5 como ya hicimos antes con la Suposición 15.9. Así que supongamos nuevamente que las funciones  $\theta_\nu$  son cuadráticas y están definidas del siguiente modo

$$\theta_\nu(x) = (x^\nu)^T A_{\nu\nu} x^\nu + \sum_{\mu=1, \mu \neq \nu}^N (x^\nu)^T A_{\nu\mu} x^\mu,$$

donde cada matriz  $A_{\nu\mu} \in \mathbb{R}^{n_\nu \times n_\mu}$  es el  $(\nu, \mu)$  subbloque de la matriz

$$A := (A_{\nu\mu})_{\nu, \mu=1}^N \in \mathbb{R}^{n \times n}.$$

Suponemos también que las matrices  $A_{\nu\mu}$  son simétricas.

**Proposición 16.7.** *Si la matriz  $A$  es definida positiva, entonces la Suposición 16.5 se satisface en todos los puntos  $x \in \mathbb{R}^n$ .*

PRUEBA. La prueba es prácticamente igual a la demostración de la Proposición 15.11, con  $\hat{y}_\beta(x)$  en vez de  $x$ . Recordemos que

$$\nabla_{x^\mu}\theta_\nu(x^\nu, x^{-\nu}) = \begin{cases} A_{\nu\mu}x^\nu, & \text{si } \mu \neq \nu, \\ A_{\nu\nu}x^\nu + \sum_{\mu=1}^N A_{\nu\mu}x^\mu, & \text{si } \mu = \nu. \end{cases}$$

Entonces

$$\begin{aligned} & \sum_{\nu=1}^N [\nabla\theta_\nu(\hat{y}_\beta^\nu(x), x^{-\nu}) - \nabla\theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})]^T (\hat{y}_\beta(x) - \hat{y}_\alpha(x)) \\ & = \sum_{\nu=1}^N \sum_{\mu=1}^N [\nabla_{x^\mu}\theta_\nu(\hat{y}_\beta^\nu(x), x^{-\nu}) - \nabla_{x^\mu}\theta_\nu(\hat{y}_\alpha^\nu(x), x^{-\nu})]^T (\hat{y}_\beta^\mu(x) - \hat{y}_\alpha^\mu(x)) \\ & = 2 \sum_{\nu, \mu=1}^N (\hat{y}_\beta^\nu(x) - \hat{y}_\alpha^\nu(x))^T A_{\nu\mu} (\hat{y}_\beta^\mu(x) - \hat{y}_\alpha^\mu(x)) \\ & = 2(\hat{y}_\beta^\nu(x) - \hat{y}_\alpha(x))^T A (\hat{y}_\beta^\nu(x) - \hat{y}_\alpha(x)), \end{aligned}$$

y por consiguiente  $\hat{y}_\beta(x) \neq \hat{y}_\alpha(x)$  implica que este número es positivo.  $\square$

# Capítulo 5

## Conclusiones y recomendaciones

El Problema de Equilibrio de Nash Generalizado (GNEP) es la generalización natural del Problema de Equilibrio de Nash (NEP). El GNEP se adecúa mejor que el NEP a las situaciones de la vida real, pero es mucho más difícil de estudiar. Resolver el GNEP es complicado, ya que consiste en resolver al mismo tiempo varios problemas de minimización, cuyos conjuntos factibles pueden variar conforme intentamos resolverlos. Es por ello que las reformulaciones del GNEP son importantes.

Existen varios resultados que garantizan que un GNEP dado posee solución. Obtuvimos varios usando las reformulaciones del Capítulo 3. Pero el resultado más importante existente en la literatura posiblemente sea el Teorema 10.8.

En el caso convexo y diferenciable del GNEP (o sea  $X$  es convexo y cada  $\theta_\nu$  es de clase  $C^1$  en  $\mathbb{R}^n$  y convexa respecto a la  $\nu$ -ésima variable), toda solución del problema de desigualdad variacional  $VI(F, X)$ , donde la función  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  es aquella cuya definición se dio en la Sección 11, es solución del GNEP mas no toda solución del GNEP es solución del problema  $VI(F, X)$ . La equivalencia entre los dos problemas se da en el contexto de los NEPs.

Si el conjunto  $X$  tiene la forma  $\{x \in \mathbb{R}^n \mid g_i(x) \leq 0, \forall i = 1, \dots, r\}$  donde cada  $g_i$  es convexa y diferenciable, existe una caracterización de las soluciones del GNEP que son soluciones de  $VI(F, X)$ , obtenida utilizando las condiciones KKT<sup>1</sup>: si  $\bar{x}$  es una solución de  $VI(F, X)$  con multiplicadores  $\lambda$ , entonces  $\bar{x}$  es solución del GNEP con multiplicadores  $\lambda^1 = \dots = \lambda^N = \lambda$ ; y si  $\bar{x} \in X$  es una solución del GNEP con multiplicadores  $\lambda^1 = \dots = \lambda^N$ , entonces  $\bar{x}$  es solución de  $VI(F, X)$  con multiplicadores  $\lambda = \lambda^1 = \dots = \lambda^N$ .

En el caso convexo del GNEP (o sea  $X$  es convexo y cada  $\theta_\nu$  es continua en  $\mathbb{R}^n$  y convexa respecto a la  $\nu$ -ésima variable), toda solución del problema de desigualdad variacional  $VI(\Phi, X)$ , donde el operador  $\Phi : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  es aquel cuya definición se dio en la Sección 12, es solución del GNEP mas no toda solución del GNEP es solución

---

<sup>1</sup>Recordemos que si  $\bar{x} \in X$  es solución de  $VI(F, X)$  y alguna condición de calificación se satisface en  $\bar{x}$  entonces existe un vector de multiplicadores de Lagrange  $\lambda$  en  $\bar{x}$  (véase la Sección 7); similarmente, si  $\bar{x} = (\bar{x}^1, \dots, \bar{x}^N)$  es una solución del GNEP, o sea si cada  $\bar{x}^\nu$  minimiza  $\theta_\nu(\cdot, \bar{x}^{-\nu})$  sobre  $X_\nu(\bar{x}^{-\nu})$ , y se cumplen condiciones de calificación para cada problema de minimización, entonces existen vectores de multiplicadores de Lagrange  $\lambda^1, \dots, \lambda^N$  en  $\bar{x}$ , uno por cada problema (véase la Sección 3.)

del problema  $VI(\Phi, X)$ . La equivalencia entre los dos problemas se da en el contexto de los NEPs.

En el caso cuasiconvexo del GNEP (o sea  $X$  es convexo y cada  $\theta_\nu$  es continua en  $\mathbb{R}^n$  y cuasiconvexa respecto a la  $\nu$ -ésima variable), toda solución del problema de desigualdad variacional  $VI(N_\theta^a, X)$ , donde el operador multivaluado  $N_\theta^a : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$  es aquel cuya definición se dio en la Sección 13, es solución del GNEP. A diferencia de lo obtenido en el caso convexo, aquí  $X$  no tiene necesariamente que ser convexo. No obstante, el recíproco solo se da en el contexto de los NEPs, cuando cada función de coste  $\theta_\nu$  es cuasiconvexa semiestricta y cuando  $X$  es convexo.

Para concluir que el GNEP tiene solución es suficiente garantizar que los problemas de desigualdad variacional asociados tienen solución. Así, si el conjunto  $X$  de la definición del GNEP es convexo, compacto y no vacío entonces el problema tiene solución cuando cada función  $\theta_\nu$  es cuasiconvexa semiestricta respecto a la  $\nu$ -ésima variable.

Los respectivos problemas de desigualdad cuasivariacional  $QVI(F, \Omega)$ ,  $QVI(\Phi, \Omega)$  y  $QVI(N_\theta^a, \Omega)$ , que extienden las reformulaciones anteriores, son equivalentes al GNEP en cada caso.

Para las conclusiones restantes suponemos que  $X$  es convexo y cerrado, y que cada función  $\theta_\nu$  es continua en  $\mathbb{R}^n$  y convexa respecto de la  $\nu$ -ésima variable.

Un vector  $\bar{x} \in \mathbb{R}^n$  es solución del GNEP si y solo si cualquiera de las siguientes condiciones se satisface

- (a)  $\bar{x}$  es punto fijo del operador  $\mathcal{S}$ ;
- (b)  $\bar{x}$  minimiza  $V$  sobre  $X$  y  $V(\bar{x}) = 0$ , donde  $V$  es la función definida en (14.1);
- (c)  $\bar{x}$  minimiza  $V_\alpha$  sobre  $X$  y  $V_\alpha(\bar{x}) = 0$ , donde  $V_\alpha$  es la función definida en (14.3);
- (d)  $\bar{x}$  es un punto fijo de la aplicación  $y_\alpha$ , donde  $y_\alpha$  es la función definida en el Teorema 14.9.

Las funciones  $V$  y  $V_\alpha$  no son necesariamente diferenciables cuando las funciones  $\theta_\nu$  lo son. Así, los problemas (14.2) y (14.5) son reformulaciones del GNEP como problemas de minimización *con* restricciones y con función objetivo *no diferenciable*.

Todo equilibrio de Nash normalizado es solución del GNEP, mas no toda solución del GNEP es un equilibrio de Nash normalizado. Un vector  $\bar{x} \in \mathbb{R}^n$  es un equilibrio de Nash normalizado si y solo si cualquiera de las siguientes condiciones se satisface

- (a)  $\bar{x}$  minimiza  $\hat{V}_\alpha$  sobre  $X$  y  $\hat{V}_\alpha(\bar{x}) = 0$ , donde  $\hat{V}_\alpha$  es la función definida en (15.1);
- (b)  $\bar{x}$  es un punto fijo de la aplicación  $\hat{y}_\alpha$ , donde  $\hat{y}_\alpha$  es la función definida en el Teorema 15.4;
- (c)  $\bar{x}$  minimiza  $\hat{V}_{\alpha\beta}$  sobre  $\mathbb{R}^n$ , donde  $\hat{V}_{\alpha\beta}$  es la función definida en (16.2).

Las funciones  $\hat{V}_\alpha$  y  $\hat{V}_{\alpha\beta}$  son de clase  $C^1$  cuando las funciones  $\theta_\nu$  lo son. Luego, el problema (15.2) es una reformulación del problema de hallar los equilibrios de Nash normalizados como un problema de minimización *con* restricciones y con función

objetivo *diferenciable*, mientras que el problema (16.3) es otra reformulación de dicho problema como un problema de minimización *sin* restricciones y con función objetivo *diferenciable*.

Bajo la hipótesis de la Suposición 15.9 todo punto estacionario del problema (15.2) es un equilibrio de Nash normalizado. Análogamente, bajo la hipótesis de la Suposición 16.5 todo punto estacionario del problema (16.3) es un equilibrio de Nash normalizado.

Si cada función de beneficio  $\theta_\nu$  fuera fuertemente convexa respecto a  $x^\nu$  no sería necesario regularizar la función de Nikaidô-Isoda. De hecho, trabajando con  $V$  y  $\Psi$  en vez de con  $\hat{V}_\alpha$  y  $\Psi_\alpha$  obtenemos una reformulación idéntica a (15.2) como problema de minimización con función objetivo diferenciable.

Las reformulaciones (14.5) y (15.2) coinciden en el contexto de los NEPs.

Puesto que en esta tesis hemos estudiado el GNEP y hemos presentado varias reformulaciones del problema, recomendamos que este estudio sea continuado desde un punto de vista práctico (nuestro estudio ha sido solamente teórico). Es decir, se podrían considerar problemas de la vida real específicos, modelarlos con el GNEP, reformularlos como problemas de optimización o de desigualdad variacional, y luego intentar estudiarlos a partir de dichas reformulaciones.

Otro camino a seguir consistiría en analizar las reformulaciones desde un punto de vista algorítmico, esto es, intentar ver cuán eficientemente funcionan con el GNEP los algoritmos clásicos empleados para resolver problemas de optimización y problemas de desigualdad variacional. En este caso posiblemente sea necesario trabajar bajo ciertas hipótesis, como por ejemplo la diferenciabilidad de las funciones  $\theta_\nu$ .

# Bibliografía

- [1] Aubin, J.P.; Frankowska, H., *Set-valued Analysis*, Birkhauser, Boston, 1990.
- [2] Arrow, K.; Debreu, G. *Existence of an Equilibrium for a Competitive Economy*, *Econometrica*, 22, pp. 265-290, 1954.
- [3] Aussel, D.; Dutta, J., *Generalized Nash Equilibrium Problem. Variational Inequality and Quasiconvexity*, *Operations Research Letters* 36, pp. 461-464, 2008.
- [4] Aussel, D.; Hadjisavvas, N., *Adjusted Sublevel Sets, Normal Operator and Quasiconvex Programming*, *SIAM Journal of Optimization*, Vol. 16, 2005, pp. 433-457.
- [5] Bertsekas, D. P.; Nedic, A.; Ozdaglar, A., *Convex Analysis and Optimization*, Athena Scientific, Belmont, Massachusetts, 2003.
- [6] Bertsekas, D., *Nonlinear Programming*, Athena Scientific, Belmont, Massachusetts, 1999.
- [7] Borde, J.; Crouzeix, J.-P., *Continuity Properties of the Normal Cone to the Sublevel Sets of a Quasiconvex Function*, *J. Optim. Theory Appl.* 66, 1990, pp. 415-429.
- [8] Borwein, J.; Lewis, A., *Convex Analysis and Nonlinear Optimization*, CMS Books in Mathematics, Springer, 2006.
- [9] Crouzeix, J. P.; Ocaña, E.; Sosa, W., *Análisis Convexo*, Instituto de Matemáticas y Ciencias Afines, IMCA, Lima, 2003.
- [10] Debreu, G., *Theory of Values*. Yale University Press, New Haven, USA, 1959.
- [11] Facchinei F.; Fischer, A.; Piccialli, V. On generalized Nash games and variational inequalities. *Operations Research Letters*, 35, pp. 159–174, 2007.
- [12] Facchinei F.; Kanzow, C. Generalized Nash Equilibrium Problems. *Annals of Operations Research*, 175, pp. 177–211, 2010.
- [13] Facchinei, F.; Pang, J.-S., *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Volumes I and II. Springer, New York, NY, 2003.
- [14] Granas, A.; Dugundji, J., *Fixed Point Theory*, Springer, 2003.

- 
- [15] García Ramos, Y.; *Un estudio sobre funciones pseudoconvexas*, Tesis de maestría, Lima, 2003.
- [16] von Heusinger, A.; Kanzow, C., *Optimization reformulations of the generalized Nash equilibrium problem using Nikaidô-Isoda-Type functions*, *Comput. Optim. Appl.* 43, 2009, no. 3, pp. 353–377.
- [17] Hilario Poma, C., *El problema de equilibrio de Nash generalizado y desigualdad variacional*, Seminario de Tesis 2, Universidad Nacional de Ingeniería, Lima, 2013.
- [18] Hiriart-Urruty, J.-B.; Lemaréchal, C. *Convex Analysis and Minimization Algorithms*, Springer-Verlag, 2003. Two volumes.
- [19] Hoffman, K.; Kunze, R., *Linear Algebra*. Prentice-Hall, Englewoods Cliffs, N. J., 1971.
- [20] Hogan, W., *Point-to-set maps in mathematical programming*, *SIAM Review* 15, 1973, pp. 591–603.
- [21] Ichiishi, T., *Game theory for economic analysis*, Academic Press, 1983.
- [22] Karamardian, S.; Schaible, S. *Seven Kinds of Monotone Maps*, *Journal of Optimization Theory and Applications*, Vol 66, No. 1, 1990, pp. 37–46.
- [23] Lima, E., *Álgebra Lineal*. Textos del IMCA, Instituto de Matemática y Ciencias Afines, 1998.
- [24] Lima, E., *Curso de Análise, Vol. 2*, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1989.
- [25] Lima, E., *Espaços Métricos*, Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1979.
- [26] Nash J. F., *Equilibrium points in n-person games*, *Proc. Nat. Acad. Sci. U. S. A.* 36, 1950, pp. 48-49.
- [27] Nash J. F., *Non-cooperative Games*, *Ann. of Math.* 54, 1951, pp. 286-295.
- [28] Rockafellar R.T.. *Convex Analysis*, Princeton University Press, Princeton, 1998.
- [29] Rosen, J.B. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica*. *Econometrica*, 33, pp. 520–534, 1965.
- [30] Rudin, W., *Principles of Mathematical Analysis*, McGraw-Hill Inc., 1976.
- [31] Rudin, W., *Real and Complex Analysis*. McGraw-Hill, Inc, 1987.
- [32] Zalinescu, C., *Convex Analysis in General Vector Spaces*, World Scientific Publishing Co., 2002.