

UNIVERSIDAD NACIONAL DE INGENIERÍA  
Facultad de Ciencias

Escuela Profesional de Matemática



INFORME DE SUFICIENCIA  
PARA OPTAR EL TITULO PROFESIONAL DE  
LICENCIADO EN MATEMÁTICA

Titulado:

*MÉTODO DE PLANOS DE CORTE Y CENTRO ANALÍTICO*

Presentado por:

LOO JAVIER MAMANI QUEA

Asesor:

WILLIAM ECHEGARAY CASTILLO

LIMA-PERÚ

2011

# Dedicatoria

A Zaira el motor de mi vida,  
y Mabel el timón de mi destino

## Agradecimientos

Quiero expresar mi gratitud al profesor Wilfredo Sosa Sandoval por sus orientaciones, enseñanzas y su inestimable ayuda en la elaboración de este trabajo.

Agradecer también al profesor William Echeagaray Castillo por mostrarme que un profesor también puede ser un amigo, agradecer el intercambio de ideas y consejos que han ayudado en este trabajo.

En último lugar, pero no por ello menos importante, agradecer el apoyo moral a mi esposa Mabel que durante todo este tiempo me ha animado a realizar este trabajo y con la que espero poder recuperar los momentos que perdí.

## Resumen

El presente trabajo tiene por objetivo presentar una opción de como resolver problemas de programación lineal (PL) cuando la cantidad de variables es muy grande, este método es conocido con el nombre de ACCPM (Analytic Center Cutting Plane Method). El método usa las técnicas de optimización denominada puntos interiores, la cual inicia en el interior de la región factible a diferencia del método SIMPLEX que evoluciona por sus puntos extremos, este método disminuye considerablemente el tiempo de solución de ciertos problemas de optimización lineal de gran tamaño, ya que para llegar a la solución no utiliza los extremos.

Presentaremos una descripción del método denominado planos de corte el cual se basa en el cálculo del centro analítico de la región factible, Se presentan las generalidades de los métodos de plano de corte, de punto interior, de elipses y el cálculo del centro analítico.

# Índice

1. INTRODUCCIÓN	1
2. PRELIMINARES	2
3. PROGRAMACIÓN LINEAL	19
4. MÉTODO DE PUNTOS INTERIORES	22
5. CENTRO ANALÍTICO	23
6. MÉTODO DE PLANOS DE CORTE	24
7. CENTRO ANALÍTICO EN $\mathbb{R}^n$	29
7.1. Condiciones de optimalidad . . . . .	31
7.2. Algoritmo ACCPM . . . . .	33
8. RESULTADOS NUMÉRICOS	34
9. CONCLUSIONES Y RECOMENDACIONES	36
10. APÉNDICE	37
BIBLIOGRAFÍA	40

# 1. INTRODUCCIÓN

Resolver un problema de Programación Matemática es buscar el máximo (o el mínimo) de una función algebraica de variables ligadas por ecuaciones o inecuaciones algebraicas de cualquier grado llamadas restricciones. En el caso particular, donde la función a maximizar (minimizar) y todas las restricciones son de primer grado, el problema recibe el nombre de Programación Lineal (PL), cuyo objetivo es buscar un óptimo (máximo o mínimo) de una función lineal (función objetivo) de  $n$  variables  $x_j$  relacionadas entre si por ecuaciones o inecuaciones lineales llamadas restricciones.

La programación lineal ha sido a lo largo del siglo pasado una de las técnicas mas innovadoras y que mas ha contribuido al desarrollo económico, social y bélico, entre otros. Actualmente la programación lineal se está aplicando en muchas actividades, las cuales han ido más allá de las aplicaciones militares e industriales, para incluir hospitales, instituciones financieras, bibliotecas, planeación urbana, sistemas de transporte y sistemas de comercialización. El avance tecnológico, específicamente en el área de cómputo, ha puesto al alcance de todos la programación lineal.

La programación lineal se centra en solucionar las siguientes interrogantes: ¿Cuál es la forma más eficiente de asignar ciertos recursos escasos para conseguir el más alto beneficio? ¿Cuál es la mejor manera de asignar rutas a una flota de transporte de carga para que lo haga con el menor costo posible? ¿Cuántas ventanillas deben colocarse en un banco en las horas normales y en las horas y días "punta" para que los clientes no se desesperen y se vayan al otro banco? ¿Cuántas cajas registradoras deben habilitar un supermercado para que el largo de las colas no entorpezca la circulación de los clientes que aún están comprando?

La PL encuentra su aplicación práctica en casi todas las facetas de los negocios, desde la publicidad hasta la planificación de la producción. Problemas de transporte, distribución y planificación global de la producción son sus objetos más comunes. Otras aplicaciones de la programación lineal son las Finanzas, la Administración de Producción y Operaciones, los Recursos Humanos y la Distribución, de ahí vemos que la PL es una herramienta importante en la toma de decisiones y en la planificación de ciertas estrategias a tomar en el futuro.

Durante décadas, el método simplex domino la solución de este tipo de problemas, dada su fácil implementación y el tiempo de resolución. Hoy en día tenemos muchos programas de computación que resuelven problemas de PL mediante simplex, pero la dificultad es que mientras mas variables ingresemos, el tiempo que se demora es mayor y los errores de aproximación y redondeo aumentan considerablemente, lo cual hace que el método simplex ya no sea el indicado.

Gracias a la inquietud de grados matemáticos, tales como Karmakar, de mejorar el tiempo de solución de problemas a mediana y gran escala y de disminuir la acumulación de errores de redondeo, surgen diversos métodos para resolver problemas de PL, entre ellos los de punto interior.

Tradicionalmente los problemas de PL se resuelven usando el método SIMPLEX.

el cual trabaja con los vertices de la región factible hasta encontrar el óptimo del problema, mientras mas vertices mas tiempo de resolución, más costos en computación e implementación y mayores errores de redondeo. Dada esta necesidad de contar con un método adecuado para resolver este tipo de problemas, queremos plantear en el presente trabajo una alternativa para resolver este tipo de problemas de programación lineal, es decir dada una función objetivo como logramos maximizarlo (minimizarlo) mejorando el tiempo de resolución respecto al simplex, para ello utilizaremos un algoritmo de puntos interiores basado en consideraciones geométricas, conocido con el nombre de ACCPM (Analytic Center Cutting Plane Method), el método que plantearemos combina el *algoritmo elipsoidal* con el de *planos de corte*. El algoritmo elipsoidal se usará para crear dentro de la región factible una elipsoide de volumen máximo cuyo centro será llamado centro analítico, por donde se trazará un hiperplano el cual dividirá la región factible en dos, se elegirá una nueva región factible "menor" que la anterior por donde se volverá a trazar un hiperplano para volver a dividir la nueva región factible en dos y así sucesivamente, esto se hará de forma iterativa, cuando las iteraciones tiendan al infinito el volumen de los elipsoides tenderá a cero, ese hecho hace que este método converja a una solución del problema.

## 2. PRELIMINARES

En esta sección proporcionaremos un pequeño resumen acerca de conceptos básicos en  $\mathbb{R}^n$ , de álgebra y programación lineal que resultarán fundamentales para el buen entendimiento de este trabajo ya que será el lugar donde nos desenvolveremos. Indicaremos la notación y algunos hechos conocidos sobre el campo del álgebra lineal y teoría de optimización. Estos temas pueden ser encontrados en muchos de los libros de Algebra Lineal y de Programación Lineal.

**Definición 2.1** Sean  $\mathcal{A}$  y  $\mathcal{B}$  dos conjuntos. El producto cartesiano  $\mathcal{A} \times \mathcal{B}$  es el conjunto de parejas ordenadas que tienen como primera componente un elemento de  $\mathcal{A}$  y como segunda componente un elemento de  $\mathcal{B}$  (véase [7]). Esto es:

$$\mathcal{A} \times \mathcal{B} := \left\{ \begin{pmatrix} a \\ b \end{pmatrix} / a \in \mathcal{A} \text{ y } b \in \mathcal{B} \right\}$$

donde  $\begin{pmatrix} a \\ b \end{pmatrix}$  es llamado par ordenado del producto cartesiano  $\mathcal{A} \times \mathcal{B}$ .

**Ejemplo 2.1** Si consideramos  $\mathcal{A} = \mathcal{B} = \mathbb{R}$  obtenemos

$$\mathbb{R} \times \mathbb{R} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} / a \in \mathbb{R} \text{ y } b \in \mathbb{R} \right\}$$

**Definición 2.2** Dado  $n$  conjuntos  $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$  se llama producto cartesiano

$\mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_n$  al conjunto de todas las  $n$ -uplas ordenadas  $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  tales que cada

componente  $x_i$  pertenece al conjunto  $\mathcal{A}_i$ .

$$\mathcal{A}_1 \times \mathcal{A}_2 \times \dots \times \mathcal{A}_n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} / x_i \in \mathcal{A}_i \right\}$$

Es común el Producto Cartesiano de un conjunto por si mismo dos o mas veces, en ese caso la notación se abrevia:

- $\mathcal{A} \times \mathcal{A} = \mathcal{A}^2$
- $\mathcal{A} \times \mathcal{A} \times \mathcal{A} = \mathcal{A}^3$
- $\mathcal{A} \times \mathcal{A} \times \dots \times \mathcal{A} = \mathcal{A}^n$  (n veces)

**Ejemplo 2.2** .

1.  $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$
2.  $\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3$

**Definición 2.3** Se define  $\mathbb{R}^n$  como el producto cartesiano de n veces el conjunto  $\mathbb{R}$  y cada elemento de  $\mathbb{R}^n$  se representara como una columna de n numeros reales, donde a cada número real se le llamará componente y será identificado por un subíndice, de acuerdo a la posición que ocupa (véase [7]).

**Ejemplo 2.3**  $a = \begin{pmatrix} 2 \\ -8 \\ 0 \\ 4 \end{pmatrix}$  es un vector en  $\mathbb{R}^4$  y  $a_3 = 0$  es su tercera componente.

Al vector, cuyas componentes son todos ceros, sera llamado vector nulo y sera representado por 0.

**Definición 2.4** Definimos a una matriz de orden  $m \times n$  como una fila ordenada de n vectores en  $\mathbb{R}^m$ , donde a cada vector se le llamará columna de la matriz y será identificado por un subíndice de acuerdo a la posición que ocupa (véase [7]). Si denotamos por A a la matriz de orden  $m \times n$ , entonces  $A_i$  denotará la columna i de A y  $a^i$  denotará la fila i de A. Según esta notación tenemos que

$$A = [A_1, A_2, \dots, A_n] = \begin{pmatrix} a^1 \\ \vdots \\ a^m \end{pmatrix}$$

**Ejemplo 2.4** Para  $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \\ 5 & 6 \end{pmatrix}$  que es una matriz de orden  $3 \times 2$  tenemos que

$A_2 = \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}$  que es la segunda columna de A y  $a^1 = \begin{pmatrix} 1 & 2 \end{pmatrix}$  es primera fila de A.



Ahora se hace necesario recordar lo que es un Espacio Vectorial real, para saber todo lo que podemos hacer en  $\mathbb{R}^n$ , y dotar a  $\mathbb{R}^n$  de una estructura vectorial, definamos a los espacios vectoriales reales.

**Definición 2.5** *Un espacio vectorial real  $V$  es un conjunto cuyos elementos son llamados vectores (véase [7]), en el cual están definidas dos operaciones:*

1. *La adición:*

$$+ : V \times V \rightarrow V$$

*tal que  $+(u, v) := u + v$*

2. *La multiplicación por un escalar:*

$$\cdot : \mathbb{R} \times V \rightarrow V$$

*tal que  $\cdot(\alpha, v) := \alpha v$  las cuales satisfacen las siguientes condiciones:*

1.1) *Conmutatividad:  $u + v = v + u$ .*

1.2) *Asociatividad:  $(u + v) + w = u + (v + w)$ .*

1.3) *Vector nulo:  $\exists o \in V$  tal que  $u + o = o + u = u \quad \forall u \in V$*

1.4) *Inverso Aditivo:  $\forall u \in V, \exists -u \in V$  tal que  $u + (-u) = 0$ .*

2.1) *Distributividad:  $(\alpha + \lambda)u = \alpha u + \lambda u$   
 $\alpha(u + v) = \alpha u + \alpha v$*

2.2) *Multiplicación por 1:  $1v = v, \forall v \in V$*

Considerando las siguientes operaciones

$$1. + : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \text{ tal que } + \left( \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right) = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

$$2. \cdot : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n \text{ tal que } \cdot \left( \lambda, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right) = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}$$

se verifica que  $\mathbb{R}^n$  es un espacio vectorial real.

**Ejemplo 2.5** *El conjunto de vectores*

$$V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} / x_i \in \mathbb{R} \right\}$$

*define el espacio  $\mathbb{R}^3$*

**Definición 2.6** *Un subconjunto  $L \subseteq \mathbb{R}^n$  es llamado un subespacio vectorial real de  $\mathbb{R}^n$  si se cumple lo siguiente:*

1.  $0 \in L$
2.  $\forall x \in L \text{ y } \forall y \in L, x + y \in L.$
3.  $\forall \lambda \in \mathbb{R} \text{ y } \forall x \in L, \lambda x \in L.$

**Comentario 2.1** Si  $W$  es un sub espacio vectorial de  $\mathbb{R}^n$  entonces  $W$  también es un espacio vectorial sobre  $\mathbb{R}$ .

**Ejemplo 2.6** Para el espacio  $V = \mathbb{R}^2$  tenemos como sub espacio  $W \subset V$ , donde  $W = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} / x_1 = 2x_2 \right\}$  es una recta que pasa por el origen

**Ejemplo 2.7** En  $\mathbb{R}^3$  aquel plano que no pase por el origen no es un subespacio ya que no contiene al cero.

Como hemos visto anteriormente cuando sumamos dos vectores de un espacio vectorial o multiplicamos un vector de un subespacio por un escalar, el resultado es un elemento de este subespacio. Cuando hacemos esto repetidas veces lo que obtenemos es una combinación lineal de vectores.

**Definición 2.7** Sea  $U \subset \mathbb{R}^n$  un espacio vectorial real,  $u_1, u_2, \dots, u_n \in U$  y  $a_1, a_2, \dots, a_n$  números reales. Entonces el vector

$$u = a_1 u_1 + a_2 u_2 + \dots + a_n u_n$$

es un elemento de  $U$  al que llamaremos combinación lineal de  $u_1, u_2, \dots, u_n$

**Comentario 2.2** Observe que en la definición de combinación lineal, la cantidad de elementos involucrados en la suma es finito. Es decir que cuando se habla sobre combinaciones lineales, la cantidad de vectores involucrados en la combinación lineal es finita.

**Definición 2.8** Se dice que  $U \subset \mathbb{R}^n$  genera  $\mathbb{R}^n$  si cada elemento  $v \in \mathbb{R}^n$  es una combinación lineal de elementos de  $U$ .

**Ejemplo 2.8** Sean  $v_1, v_2 \in V \subset \mathbb{R}^2$ , con  $v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ , y  $v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , y  $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$  tenemos que

$$\begin{pmatrix} x \\ y \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

o sea  $V$  genera a  $\mathbb{R}^2$  ya que

$$v = xv_1 + yv_2$$

**Definición 2.9** Se dice que  $U \subset \mathbb{R}^n$  es linealmente independiente si cualquier representación del vector nulo como una combinación lineal de elementos de  $U$  implica que los coeficientes de la combinación lineal son ceros (véase [7]). Es decir que para  $u_1, u_2, \dots, u_n \in U$  y  $a_1, a_2, \dots, a_n \in \mathbb{R}$  y

$$a_1 u_1 + a_2 u_2 + \dots + a_n u_n = 0$$

implica que  $a_1 = a_2 = \dots = a_n = 0$ .

En caso que exista algún  $a_i \neq 0$  diremos que  $u_1, u_2, \dots, u_n$  es linealmente dependiente.

**Ejemplo 2.9** Los vectores  $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$  son linealmente independientes en  $\mathbb{R}^3$

Esto es cierto ya que al encontrar las posibles soluciones en

$$a_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + a_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + a_3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

obtenemos  $a_1 = a_2 = a_3 = 0$

Ahora estamos interesados en encontrar, dentro de un espacio vectorial  $V$ , un conjunto finito de vectores, tales que cualquier otro vector de  $V$  sea una combinación lineal de ellos. En otras palabras, queremos determinar un conjunto de vectores que genere  $V$  y tal que todos los elementos sean realmente necesarios para generar  $V$ .

**Definición 2.10** Se dice que  $U \subset \mathbb{R}^n$  es una base de  $\mathbb{R}^n$  si

1.  $U$  genera  $\mathbb{R}^n$ .
2. Todo subconjunto finito de  $U$  es linealmente independiente.

**Ejemplo 2.10**  $V = \mathbb{R}^2$ ,  $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  y  $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  donde  $\{e_1, e_2\}$  es base de  $V$ , conocida como base canónica de  $\mathbb{R}^2$

Es fácil verificar que en  $\mathbb{R}^n$ , el conjunto  $\{e_i\}_{i=1}^n$ , donde  $e_i$  es un vector cuya  $i$ -ésima componente es 1 y el resto son ceros, es una base para  $\mathbb{R}^n$ . A partir de este resultado, muy conocido, se verifica que toda base de  $\mathbb{R}^n$  tiene  $n$  elementos.

**Definición 2.11** Se define la dimensión de  $\mathbb{R}^n$  como la cantidad de vectores de cualquiera de sus bases (véase [7]).

**Ejemplo 2.11** Para  $V = \mathbb{R}^2$  tenemos que  $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$  y  $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$  son bases de  $\mathbb{R}^2$ . Entonces la dimensión de  $\mathbb{R}^2$  es 2

Como  $\{e_i\}_{i=1}^n$  es una base de  $\mathbb{R}^n$ , entonces la dimensión de  $\mathbb{R}^n$  es  $n$ .

**Definición 2.12** Dados dos vectores  $a, b \in \mathbb{R}^n$  diferentes (véase [3]), la línea o recta que pasa a través de  $a$  y  $b$  está definida por:

$$L = \{\lambda a + \mu b / \lambda + \mu = 1\} = \{\lambda a + (1 - \lambda)b / \lambda \in \mathbb{R}\}$$

Es claro que este conjunto contiene a los puntos  $a$  y  $b$ .

**Ejemplo 2.12** En  $\mathbb{R}^2$ , la recta que pasa por los puntos  $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$  y  $\begin{pmatrix} 6 \\ 3 \end{pmatrix}$  esta representada por

$$L = \left\{ \lambda \begin{pmatrix} 4 \\ 2 \end{pmatrix} + (1 - \lambda) \begin{pmatrix} 6 \\ 3 \end{pmatrix} / \lambda \in \mathbb{R} \right\}$$

y que también puede representarse por

$$L = \{x \in \mathbb{R}^2 / x_1 = 2x_2\}$$

**Comentario 2.3** Si  $0 \leq \lambda \leq 1$  entonces el conjunto que obtenemos es solamente el segmento de línea que une a los puntos  $a$  y  $b$ .

**Definición 2.13 (Segmento de recta)** Sabemos que las ecuaciones paramétricas de una recta en el espacio contienen un parámetro libre,  $t \in \mathbb{R}$ . Cada vez que  $t$  toma un valor diferente en los reales, se genera un nuevo punto a lo largo de la recta infinita. Sin embargo, si en lugar de tener la condición  $t \in \mathbb{R}$ , el parámetro  $t$  se limitara a tomar valores dentro de un intervalo  $t_1 \leq t \leq t_2$  en los reales, entonces éste ya no generaría todos los puntos de la recta infinita, sino tan sólo un segmento de la recta. Dada la recta  $L$  en  $\mathbb{R}^n$  que contiene al punto  $x_0$  y es paralela al vector  $v$ , las ecuación

$$x = x_0 + vt$$

con  $t_1 \leq t \leq t_2$  genera un segmento de recta de  $L$ .

El segmento de recta que une los puntos  $p$  y  $q$  se denotará como  $\overline{pq}$ .

**Ejemplo 2.13** Encontremos la ecuación del segmento de recta que une los puntos

$p = \begin{pmatrix} -3 \\ 2 \\ -3 \end{pmatrix}$  y  $q = \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix}$ . Para ello, podemos tomar el vector de dirección  $v =$

$\begin{pmatrix} 4 \\ -3 \\ 7 \end{pmatrix}$ , de modo que el segmento de recta que une esos puntos queda perfectamente

descrito por la ecuación

$$\overline{pq} = \left\{ x \in \mathbb{R}^3 / x = \begin{pmatrix} -3 \\ 2 \\ -3 \end{pmatrix} + t \begin{pmatrix} 4 \\ -3 \\ 7 \end{pmatrix}, \quad 0 \leq t \leq 1 \right\}$$

En efecto, cuando  $t = 0$  se obtiene el punto  $p$ , cuando  $t = 1$  se obtiene el punto  $q$  y para  $0 < t < 1$  se generan todos los puntos intermedios entre  $p$  y  $q$ .

**Definición 2.14** Un conjunto  $A \subset \mathbb{R}^n$  es llamado un conjunto afín si contiene a la recta que pasa a través de cualesquiera dos de sus puntos (véase [3]), esto es

$$\forall x, y \in A. \quad \forall \lambda \in \mathbb{R}: \quad \lambda x + (1 - \lambda)y \in A$$

**Ejemplo 2.14**  $\mathbb{R}^n$  es un conjunto afín.

**Ejemplo 2.15**  $M = \{a\}$  para  $a \in \mathbb{R}^n$  es un conjunto afín de  $\mathbb{R}^n$

Un conjunto  $C \subset \mathbb{R}^n$  es afín si y solo si es de la forma

$$C = \{x + l/x \in C, l \in L\}$$

donde  $L$  es un subespacio vectorial de  $\mathbb{R}^n$  asociado a  $C$ . Es decir, un conjunto afín es un subespacio desplazado del origen. La dimensión de un conjunto afín  $x + L$  es la de su correspondiente subespacio  $L$ .

Se puede ver que cualquier conjunto afín es convexo aunque el recíproco no es cierto en general.

**Definición 2.15** Se define la dimensión de un conjunto afín como la dimensión de su subespacio paralelo (véase [5]).

Los conceptos geométricos de longitud, distancia y perpendicularidad, que son bien conocidos para  $\mathbb{R}^2$  y  $\mathbb{R}^3$ , se definirán ahora para  $\mathbb{R}^n$ . Estos conceptos proporcionan herramientas geométricas potentes para resolver problemas aplicativos entre ellos el nuestro. Una de las herramientas importantes en el desarrollo de este trabajo son los "hiperplanos" que los definiremos como sigue.

**Definición 2.16** Se dice que  $H \subset \mathbb{R}^n$  es un hiperplano si  $H$  es un conjunto afín y la dimensión de  $H$  es  $n - 1$ .

**Ejemplo 2.16** En  $\mathbb{R}^2$  toda recta es un hiperplano.

La siguiente definición sera usada posteriormente para caracterizar a los hiperplanos.

**Definición 2.17** Se dice que  $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  es un producto interno si se cumplen las siguientes propiedades (véase [6]) :

1.  $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$  para cada  $u, v, w \in \mathbb{R}^n$ .
2.  $\langle \lambda u, v \rangle = \lambda \langle u, v \rangle$  para cada  $\lambda \in \mathbb{R}, u, v \in \mathbb{R}^n$ .
3.  $\langle u, v \rangle = \langle v, u \rangle$  para cada  $u, v \in \mathbb{R}^n$ .
4.  $\langle u, u \rangle > 0$  para cada  $u \in \mathbb{R}^n \setminus \{0\}$

**Ejemplo 2.17** La aplicación  $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  definida por

$$\langle x, y \rangle = x_1 y_1 + \cdots + x_n y_n = \sum_{i=1}^n x_i y_i \quad (1)$$

es un producto interno en  $\mathbb{R}^n$ , llamado producto interno usual de  $\mathbb{R}^n$  o producto euclideo.

La noción geométrica de perpendicularidad, puede ser modelada usando este concepto de producto interno de la forma siguiente.

**Definición 2.18** Se dice que dos vectores  $u$  y  $v$  en  $\mathbb{R}^n$  son ortogonales si  $\langle u, v \rangle = 0$  (véase [6]).

**Comentario 2.4** Recordemos que el producto de una matriz  $A \in \mathbb{R}^{m \times n}$  y un vector  $x \in \mathbb{R}^n$  esta definido por:

$$Ax = \sum_{i=1}^n x_i A_i = \begin{pmatrix} \langle a^1, x \rangle \\ \vdots \\ \langle a^m, x \rangle \end{pmatrix}$$

donde  $A_i$  y  $a^i$  son las columnas y filas de  $A$  respectivamente.

Por otro lado, el concepto de producto interno esta relacionado con el siguiente concepto.

**Definición 2.19** Se dice que  $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$  es una norma si se cumple las siguientes condiciones:

1. Positividad:  $\|x\| \geq 0 \forall x \in \mathbb{R}^n$ ,  $\|x\| = 0$  si y solo si  $x = 0$ .
2. Homogeneidad:  $\|\lambda x\| = |\lambda| \|x\| \forall \lambda \in \mathbb{R}, \forall x \in \mathbb{R}^n$ .
3. Desigualdad triangular:  $\|x + y\| \leq \|x\| + \|y\| \forall x \in \mathbb{R}^n, \forall y \in \mathbb{R}^n$ .

Un vector unitario es aquel cuya norma es 1.

**Ejemplo 2.18** Si definimos  $\|x\| = \sqrt{\langle x, x \rangle}$ , entonces  $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$  es una norma. Es decir que todo producto interno induce una norma. Si el producto interno es el euclideo, entonces la norma es llamada euclidea.

**Comentario 2.5** En general una norma en  $\mathbb{R}^n$  es una función  $N : \mathbb{R}^n \rightarrow \mathbb{R}$  que verifica las propiedades de norma planteadas anteriormente.

La norma euclidiana es la norma usual la cual la denotamos por  $\|x\| = \|x\|_2$  pero no es la única, otros ejemplos de norma son la norma del máximo o norma del infinito  $\|x\|_\infty$  definido como

$$\|x\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_n|\}$$

y la norma de suma o norma uno  $\|x\|_1$  definido como

$$\|x\|_1 = |x_1| + |x_2| + \dots + |x_n|$$

En realidad, en  $\mathbb{R}^n$  se pueden definir infinitas normas (vease [6]), por ejemplo para cada  $p \in \mathbb{R}$  con  $p \geq 1$ , la aplicacion  $\|\cdot\|_p : \mathbb{R}^n \rightarrow \mathbb{R}$  definida por  $\|x\| = (\sum_{i=1}^n |x_i|^p)^{1/p}$  es una norma en  $\mathbb{R}^n$ .

Como vemos existen diferentes tipos de norma, pero todas ellas son equivalentes en  $\mathbb{R}^n$  ya que podemos obtener  $\forall x \in \mathbb{R}^n$  la relación

$$\|x\|_1 \leq \|x\|_2 \leq \|x\|_\infty \leq n \|x\|_1$$

Ahora estamos listos para caracterizar a las aplicaciones lineales  $T : \mathbb{R}^n \rightarrow \mathbb{R}$ , la intuición geométrica es la siguiente: Como  $\mathbb{R}$  caracteriza a una recta, comúnmente llamada recta real, entonces  $T$  transforma  $\mathbb{R}^n$  en una recta que sin pérdida de generalidad podemos asumir que esta recta pertenece a  $\mathbb{R}^n$  y pasa por el origen. Sea  $v_1$  un vector no nulo de dicha recta, tal que  $\|v_1\|^2 = \langle v_1, v_1 \rangle = 1$ , luego construyo  $\{v_i\}_{i=1}^n$  una base ortonormal para  $\mathbb{R}^n$ , tal que  $\langle v_i, v_j \rangle = 0 \forall i \neq j$  y  $\|v_i\| = 1 \forall i$ . Como  $\{v_i\}$  es una base, entonces para cada  $v \in \mathbb{R}^n$  entonces existen  $\{t_i\}_{i=1}^n \subset \mathbb{R}$  tal que  $v = \sum_{i=1}^n t_i v_i$  y puedo modelar  $T$  como  $T(v) = t_i$ , donde  $t_i = \langle v, v_i \rangle = \sum_{j=1}^n t_j \langle v_j, v_i \rangle$ , entonces tenemos que existe un  $v_i \in \mathbb{R}^n$  tal que  $T(v) = \langle v_i, v \rangle$ . Esta intuición geométrica es conocida en Algebra Lineal como el Teorema de representación de Riesz, el cual pasamos a enunciar.

**Proposición 2.1 (Teorema de Representación de Riesz)**

$T : \mathbb{R}^n \rightarrow \mathbb{R}$  es una aplicación lineal si y solo si existe  $a \in \mathbb{R}^n$  tal que

$$T(x) = \langle a, x \rangle$$

(ver demostración en ([12]) página 83).

**Comentario 2.6** Sea  $V$  un espacio con producto interno sobre  $\mathbb{R}$ . A las transformaciones lineales  $T : V \rightarrow \mathbb{R}$  las llaman también funcionales lineales.

**Ejemplo 2.19** La funcional lineal  $T : \mathbb{R}^4 \rightarrow \mathbb{R}$  definido mediante

$T(x) = 4x_1 - 7x_2 + 3x_3 + x_4$  se puede escribir de la forma

$$T(x) = \langle x, y \rangle \quad \forall x \in \mathbb{R}^4$$

con  $y = \begin{pmatrix} 4 \\ -7 \\ 3 \\ 1 \end{pmatrix}$

Ahora podemos caracterizar a los hiperplanos, usando la caracterización de las funcionales lineales de la forma siguiente.

**Proposición 2.2**  $H \subset \mathbb{R}^n$  es un hiperplano si y solo si  $\exists u \in \mathbb{R}^n \setminus \{0\}$  y  $\lambda \in \mathbb{R}$  tal que  $H = H(u, \lambda) = \{x \in \mathbb{R}^n / \langle u, x \rangle = \lambda\}$

(Ver la demostración en ([11]) página 13)

**Comentario 2.7** En un espacio de dimensión uno (como una recta), un hiperplano es un punto; divide una recta en dos semirectas. En un espacio bidimensional (como el plano  $xy$ ), un hiperplano es una recta; divide el plano en dos semiplanos. En un espacio tridimensional, un hiperplano es un plano; divide el espacio en dos semiespacios. Este concepto también puede ser aplicado a espacios de cuatro dimensiones y más, donde estos objetos divisores se llaman simplemente hiperplanos.

**Ejemplo 2.20** Un hiperplano de  $\mathbb{R}^2$  es  $H = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 / 2x + 3y = 6 \right\}$  que es la recta que corta al eje  $OX$  en el punto  $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$  y al eje  $OY$  en el punto  $\begin{pmatrix} 0 \\ 2 \end{pmatrix}$

**Ejemplo 2.21** Un hiperplano de  $\mathbb{R}^3$  es  $H = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 / 2x + 3y + 2z = 6 \right\}$  que es el plano que corta al eje  $OX$  en el punto  $\begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}$  al eje  $OY$  en el punto  $\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$  y al eje  $OZ$  en el punto  $\begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$

**Definición 2.20** Un poliedro es la intersección de un número finito de hiperplanos (véase ([3])).

$$P = \{x \in \mathbb{R}^n / Ax \leq b \quad A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m\}$$

es un poliedro.

Si  $P$  es acotado se le llamará politopo.

**Ejemplo 2.22** La solución del sistema lineal de inecuaciones

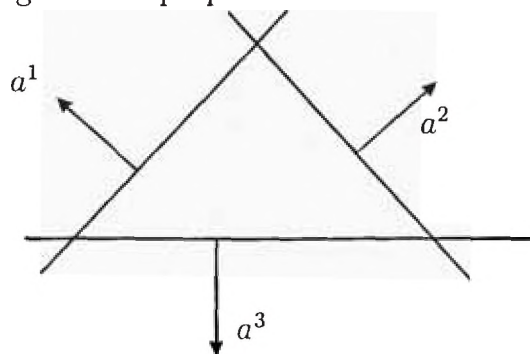
$$\begin{aligned} x_1 + x_2 &\leq 3 \\ -3x_1 + 5x_2 &\leq 15 \\ -x_2 &\leq 1 \end{aligned}$$

que es equivalente a

$$\begin{pmatrix} 1 & 1 \\ -3 & 5 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 3 \\ 15 \\ 1 \end{pmatrix}$$

es un poliedro. (véase figura 1)

Figura 1: Hiperplanos con direcciones  $a^i$



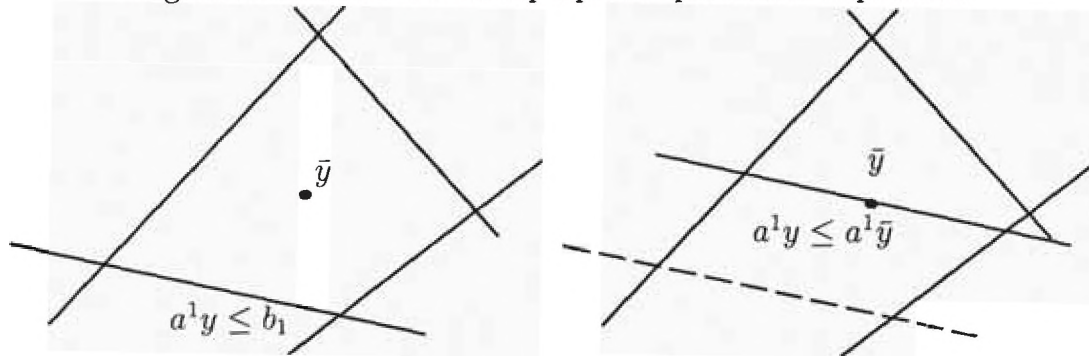


**Ejemplo 2.23** Veamos como trasladar hacia el punto  $\bar{y} \in P$  un hiperplano que define el poliedro

$$P = \{y \in \mathbb{R}^n / Ay \leq b \quad A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m\}$$

Si una inecuación definida por  $P$  como  $a^1 y \leq b_1$  la cambiamos por  $a^1 y \leq a^1 \bar{y}$ , la inecuación es trasladada de forma paralela hacia el punto  $\bar{y}$ , la cual ahora divide el poliedro inicial en dos partes. (véase figura(2))

Figura 2: Traslado de un hiperplano que define el poliedro



**Ejemplo 2.24** Dados los puntos  $p = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ ,  $q = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$  y el vector dirección

$v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ , encontremos el punto de intersección (si lo hay), de la recta que pasa por el punto  $p$  en la dirección del vector  $v$  y del hiperplano que pasa por  $q$  y es normal al vector  $v$ . (ver figura 3)

La recta que pasa por  $P$  en la dirección  $v$  es el conjunto de puntos

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \lambda \in \mathbb{R}$$

El hiperplano que pasa por  $q$  y es normal al vector  $v$ , es el conjunto de puntos

$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3$  para los cuales se satisface la ecuación:

$$\langle x, v \rangle = x_1 + 2x_2 + 3x_3 = 6$$

Sea  $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$  el punto de intersección entre la recta y el hiperplano, entonces

$$y = p + \lambda^* v \quad \Rightarrow \quad y = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \lambda^* \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

para algún  $\lambda^* \in \mathbb{R}$  y también

$$\langle y, v \rangle = 6 \implies y_1 + 2y_2 + 3y_3 = 6$$

De donde tenemos que:

$$p + \lambda^* v = q$$

utilizando propiedades de producto interno encontramos que:

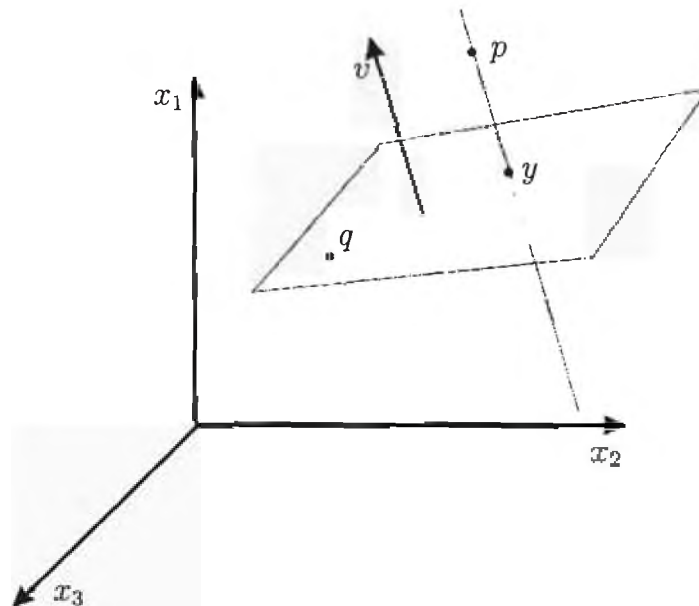
$$\lambda^* = \frac{\langle q - p, v \rangle}{\|v\|^2} = \frac{1}{14}$$

En consecuencia las coordenadas del punto buscado están dadas por:

$$y = p + \lambda^* v = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \frac{1}{14} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$y = \left( \frac{15}{14}, \frac{-12}{14}, \frac{31}{14} \right)^t$$

Figura 3: Corte de una recta con un hiperplano en  $\mathbb{R}^3$



**Definición 2.21** Sea  $H = \{x \in \mathbb{R}^n / \langle c, x \rangle = \alpha \quad c \in \mathbb{R}^n, \alpha \in \mathbb{R}\}$  un hiperplano de  $\mathbb{R}^n$

El **semiespacio cerrado positivo** asociado a  $H$  es  $H^+ = \{x \in \mathbb{R}^n / \langle c, x \rangle \geq \alpha\}$

El **semiespacio cerrado negativo** asociado a  $H$  es  $H^- = \{x \in \mathbb{R}^n / \langle c, x \rangle \leq \alpha\}$

**Ejemplo 2.25** Los semiespacios cerrados  $H^+$  y  $H^-$  asociados a  $H = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 / 2x + 3y = 6 \right\}$  de  $\mathbb{R}^2$  son: (ver figura 4 y 5)

$$H^+ = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 / 2x + 3y \geq 6 \right\}$$

$$H^- = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 / 2x + 3y \leq 6 \right\}$$

Figura 4: Semiespacio positivo

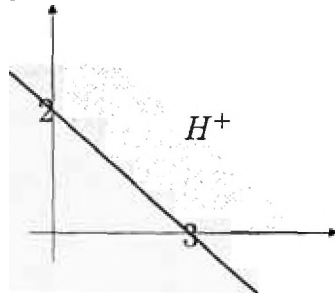
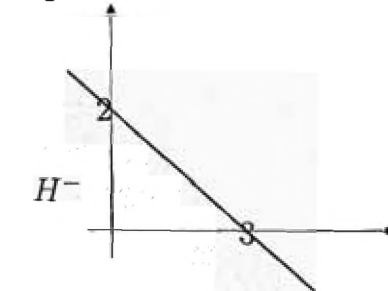


Figura 5: Semiespacio negativo



**Proposición 2.3 (Distancia de un punto a un hiperplano)**

Sea  $H$  un hiperplano de ecuación  $a_1x_1 + \dots + a_nx_n = c$  o equivalentemente  $\langle a, x \rangle = c$

y sea el punto  $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ , la distancia del punto  $y$  al hiperplano  $H$  está dada por

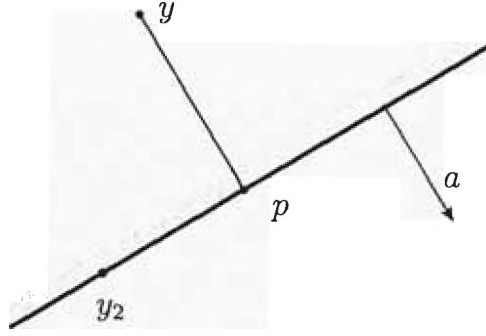
$$d(y, H) = \frac{|c - \langle y, a \rangle|}{\|a\|} = \frac{|c - a_1y_1 - \dots - a_ny_n|}{\sqrt{a_1^2 + \dots + a_n^2}}$$

**Demostración**

Definamos el hiperplano  $H$  como

$$H(a, c) = \{x \in \mathbb{R}^n / \langle a, x \rangle = c, \quad a, x \in \mathbb{R}^n \quad c \in \mathbb{R}\}$$

Figura 6: Distancia de un punto a un hiperplano



tomemos  $y_2 \in H(a, c)$  es decir  $\langle a, y_2 \rangle = c$  tomemos el punto  $p$  perteneciente al hiperplano  $H$  tal que (ver figura 6)

$p = y + ta$  y además  $(y_2 - p) \perp a$  es decir, tenemos:

$$\begin{aligned} 0 &= \langle y_2 - p, a \rangle \\ &= \langle y_2 - y - ta, a \rangle \\ &= \langle y_2, a \rangle - \langle y, a \rangle - t \langle a, a \rangle \\ &= c - \langle y, a \rangle - t \langle a, a \rangle \end{aligned}$$

de donde obtenemos:

$$|t| = \frac{|c - \langle y, a \rangle|}{\langle a, a \rangle}$$

Luego la distancia de  $y$  a  $p$  será:

$$\|y - p\| = \|y - y - ta\| = |t| \|a\| = \frac{|c - \langle y, a \rangle|}{\|a\|^2} \|a\| = \frac{|c - \langle y, a \rangle|}{\|a\|}$$

**Definición 2.22** Para el poliedro  $P = \{x \in \mathbb{R}^n / Bx \leq d\}$ , definimos el poliedro

normalizado como  $P_N = \{x \in \mathbb{R}^n / B_N x \leq d_N\}$  donde  $B_N = \begin{pmatrix} \frac{b^1}{\|b^1\|} \\ \vdots \\ \frac{b^n}{\|b^n\|} \end{pmatrix}$  y  $d_N =$

$$\begin{pmatrix} \frac{d^1}{\|b^1\|} \\ \vdots \\ \frac{d^n}{\|b^n\|} \end{pmatrix}$$

**Ejemplo 2.26** Para el poliedro  $P = \left\{ x \in \mathbb{R}^2 / \begin{pmatrix} 2 & 3 \\ 4 & 5 \\ 6 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 5 \\ 7 \\ 4 \end{pmatrix} \right\}$  tenemos

$$P_N = \left\{ x \in \mathbb{R}^2 / \begin{pmatrix} \frac{2}{\sqrt{13}} & \frac{3}{\sqrt{13}} \\ \frac{4}{\sqrt{41}} & \frac{5}{\sqrt{41}} \\ \frac{6}{\sqrt{40}} & \frac{2}{\sqrt{40}} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} \frac{5}{\sqrt{13}} \\ \frac{7}{\sqrt{41}} \\ \frac{4}{\sqrt{40}} \end{pmatrix} \right\}$$

**Definición 2.23** Sea  $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$  una norma en  $\mathbb{R}^n$ . Para cada  $x \in \mathbb{R}^n$  y cada  $\epsilon > 0$ , se define la bola con centro en  $x$  y radio  $\epsilon$  al conjunto

$$B(x, \epsilon) = \{y \in \mathbb{R}^n / \|x - y\| < \epsilon\}$$

para tener una intuición geométrica de las bolas se requiere del siguiente concepto.

**Definición 2.24** Sea  $A \subset \mathbb{R}^n$ . Se define el interior de  $A$  (véase ([6])) como

$$\overset{\circ}{A} = \{x \in A / \exists \epsilon > 0, B(x, \epsilon) \subset A\}$$

Evidentemente, las bolas dependen de la norma, por ejemplo se puede verificar en  $\mathbb{R}^2$ , que las bolas con centro en  $x$  y radio  $\epsilon$  representan:

1. El interior de un cuadrado de lado  $2\epsilon$  y centro en  $x$  cuando la norma es  $\|\cdot\|_\infty$ ,
2. El interior de un rombo de lados iguales con longitud  $\sqrt{2}\epsilon$  y centro en  $x$ , cuando la norma es con  $p = 1$ .
3. El interior de un círculo de radio  $\epsilon$  y centro en  $x$ , cuando la norma es con  $p = 2$ .

Otra de las estructuras matemáticas importantes es la "topológica", pues sin ella es imposible hablar de abiertos, cerrados, continuidad, etc.

**Definición 2.25** Se dice que una familia  $\Gamma$  de subconjuntos de  $\mathbb{R}^n$  es una topología si se cumplen las siguientes condiciones:

1.  $\emptyset \in \Gamma$  y  $\mathbb{R}^n \in \Gamma$ .
2. Si  $\{A_i\}_{i \in I} \subset \Gamma$ , entonces  $\bigcup_{i \in I} A_i \in \Gamma$ .
3. Si para cada  $p \in \mathbb{N}$  y  $\{A_i\}_{i=1}^p \subset \Gamma$ , then  $\bigcap_{i \in I} A_i \in \Gamma$ .

**Ejemplo 2.27** Si  $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$  es una norma en  $\mathbb{R}^n$ , entonces

$$\Gamma = \{A \subset \mathbb{R}^n / \overset{\circ}{A} = A\}$$

es una topología para  $\mathbb{R}^n$ .

Este ejemplo nos dice que cada norma induce una topología, pero uno de los grandes resultados del análisis, nos dice que todas las normas en  $\mathbb{R}^n$  inducen la misma topología, la cual es llamada topología euclidiana. Por este hecho, se dice que todas las normas en  $\mathbb{R}^n$  son equivalentes.

**Definición 2.26** Considerando las topologías euclidianas, una función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  es continua en  $x \in \mathbb{R}^n$  si para cada  $\epsilon > 0$  existe  $\delta \in (0, \epsilon]$ , tal que  $\forall y \in B(x, \delta)$  se tiene que  $f(y) \in B(f(x), \epsilon)$ . Se dice que  $f$  es continua si es continua en cada  $x \in \mathbb{R}^n$  (véase ([6]))

**Definición 2.27** Un conjunto  $K \subset \mathbb{R}^n$  es convexo si  $\forall x, y \in K$  y  $\lambda \in [0: 1]$  implica que  $\lambda x + (1 - \lambda)y \in K$

**Ejemplo 2.28** Todo hiperplano  $H = \{x \in \mathbb{R}^n / \langle x, c \rangle = \alpha\}$  es un conjunto convexo, ya que si  $\forall u, v \in H$  y  $\lambda \in [0; 1]$  sucede que  $\lambda u + (1 - \lambda)v \in H$ , ya que

$$\langle c, (\lambda u + (1 - \lambda)v) \rangle = \alpha$$

. En efecto:

$$\langle c, (\lambda u + (1 - \lambda)v) \rangle = \lambda \langle c, u \rangle + (1 - \lambda) \langle c, v \rangle = \lambda \alpha + (1 - \lambda) \alpha = \alpha$$

$$u \in H \Rightarrow \langle c, u \rangle = \alpha; \quad v \in H \Rightarrow \langle c, v \rangle = \alpha$$

**Ejemplo 2.29** El conjunto  $E = \{(x, y) \in \mathbb{R}^2 / \frac{x^2}{4} + \frac{y^2}{9} \leq 1\}$  es un conjunto convexo.

**Proposición 2.4** La intersección de conjuntos convexos es convexo.

**Demostración.**

Sean los conjuntos convexos  $K_1, K_2$  y los puntos  $x, y \in K_1 \cap K_2$  entonces  $x, y \in K_1$  como  $K_1$  es convexo entonces el segmento definido por  $\lambda x + (1 - \lambda)y \in K_1$ .

Análogamente para  $x, y \in K_2$  como  $K_2$  es convexo tenemos  $\lambda x + (1 - \lambda)y \in K_2$ . Luego como  $\lambda x + (1 - \lambda)y$  esta contenido simultáneamente en  $K_1$  y  $K_2$  tenemos que  $K_1 \cap K_2$  es convexo.

**Definición 2.28** Sea  $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ , siendo  $S$  convexo. Se dice que la función  $f$  es **estrictamente convexa** en  $S$  si  $\forall u, v \in S$  y  $\forall \lambda \in (0, 1)$  se cumple que

$$f(\lambda u + (1 - \lambda)v) < \lambda f(u) + (1 - \lambda)f(v)$$

En términos geométricos significa que la cuerda que une cualquier par de puntos de la gráfica de  $f$  está por encima de dicha gráfica (véase ([5])).

**Definición 2.29** Sea  $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ , siendo  $S$  convexo. Se dice que la función  $f$  es **convexa** en  $S$  si  $\forall u, v \in S$  y  $\forall \lambda \in [0, 1]$  se cumple que

$$f(\lambda u + (1 - \lambda)v) \leq \lambda f(u) + (1 - \lambda)f(v)$$

En términos geométricos significa que la cuerda que une cualquier par de puntos de la gráfica de  $f$  no está por debajo de dicha gráfica.

**Ejemplo 2.30**  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $f(x) = |x|$  es una función convexa ya que

$$|\lambda u + (1 - \lambda)v| \leq \lambda |u| + (1 - \lambda)|v|$$

**Definición 2.30** Sea  $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ , siendo  $S$  convexo. Se dice que la función  $f$  es **estrictamente cóncava** en  $S$  si  $\forall u, v \in S$  y  $\forall \lambda \in (0, 1)$  se cumple que

$$f(\lambda u + (1 - \lambda)v) > \lambda f(u) + (1 - \lambda)f(v)$$

En términos geométricos significa que la cuerda que une cualquier par de puntos de la gráfica de  $f$  está por debajo de dicha gráfica.

**Definición 2.31** Sea  $f : S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ , siendo  $S$  convexo. Se dice que la función  $f$  es **cóncava** en  $S$  si  $\forall u, v \in S$  y  $\forall \lambda \in [0, 1]$  se cumple que

$$f(\lambda u + (1 - \lambda)v) \geq \lambda f(u) + (1 - \lambda)f(v)$$

En términos geométricos significa que la cuerda que une cualquier par de puntos de la gráfica de  $f$  no está por encima de la gráfica.

**Proposición 2.5 (Teorema de Weierstrass)** Si  $A \subset \mathbb{R}^n$  es compacto (cerrado y acotado) y  $f : A \rightarrow \mathbb{R}$  es continua en  $A$ , presenta un máximo y mínimo global en  $A$ , pues por ser continua  $f$  en el compacto  $A$ , el conjunto  $f(A) \subset \mathbb{R}$  también es compacto.

( véase la demostración en ([8]) página 7)

Si  $m$  y  $M$  son los respectivos extremos inferior y superior de  $f(A)$ , ambos pertenecen a  $f(A)$  ya que  $f(A)$  es cerrado, por lo tanto existen  $x_0, x_1 \in A$  tales que  $f(x_0) = m$  y  $f(x_1) = M$ , siendo  $f(x_0) \leq f(x) \leq f(x_1), \forall x \in A$ .

**Proposición 2.6** Sea  $A \subseteq \mathbb{R}^n$  convexo no vacío y  $f : A \rightarrow \mathbb{R}$  estrictamente cóncava (convexa) en  $A$ . Si  $f$  tiene máximo (mínimo) global en  $A$ , es único.

**Demostración** (Por reducción al absurdo)

Si la función estrictamente cóncava  $f$  presenta máximo global en el punto  $x_1 \in A$  y existe  $x_2 \in A$  tal que  $f(x_1) = f(x_2)$  entonces, por ser convexo  $A$ , si  $\lambda \in [0; 1]$ , el punto  $\lambda x_1 + (1 - \lambda)x_2$  es de  $A$ , y el valor de  $f$  en él es mayor que  $f(x_1)$ , lo que contradice la hipótesis de que  $f$  presente máximo global en  $x_1$ :

Por ser  $f$  estrictamente cóncava

$$f(\lambda x_1 + (1 - \lambda)x_2) > \lambda f(x_1) + (1 - \lambda)f(x_2) = f(x_1)$$

pues  $f(x_1) = f(x_2)$ .

**Proposición 2.7** Si  $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$  es de clase  $C^1$  en el abierto  $A \subseteq \mathbb{R}^n$ , una condición necesaria para que  $x_0 \in A$  sea óptimo local es que  $\nabla f(x_0) = 0$ .

En efecto: si  $f$  tiene un óptimo local en el punto  $x_0 \in A$ , la derivada de  $f$  en  $x_0$  según cualquier vector de  $\mathbb{R}^n$  debe ser nula, pues si existe un vector  $v \in \mathbb{R}^n$  tal que  $D_v f(x_0) \neq 0$ , un desplazamiento infinitesimal en la dirección de  $v$  hace aumentar (disminuir) el valor de  $f$  si  $D_v f(x_0) > 0$  ( $D_v f(x_0) < 0$ ), y un desplazamiento infinitesimal en la dirección de  $-v$  hace disminuir (aumentar) el valor de  $f$  si  $D_v f(x_0) > 0$  ( $D_v f(x_0) < 0$ ), lo que es absurdo, pues así  $f$  no presenta un óptimo local en  $x_0 \in A$ . Por ser  $f$  de clase  $C^1$  en  $A$ , es diferenciable en  $x_0 \in A$ , y eso garantiza que  $D_v f(x_0) = \nabla f(x_0) \cdot v = 0, \forall v \in \mathbb{R}^n$  sólo si ocurre que  $\nabla f(x_0) = 0$

**Comentario 2.8** Si la función  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$  es convexa, entonces esa condición también es suficiente.

**Proposición 2.8** Sea  $f$  una función estrictamente convexa. Si  $f$  tiene al menos un mínimo, entonces ese mínimo es único.

### 3 PROGRAMACIÓN LINEAL

La programación lineal es el mecanismo mas natural para formular una gran variedad de problemas con mínimo esfuerzo. Un problema de programación lineal (PL) es un programa matemático en el cual la función objetivo es lineal en las incógnitas y las restricciones constan de igualdades y desigualdades lineales. La forma exacta de estas restricciones puede variar de unos problemas a otros, el formato general es:

$$(L) = \begin{cases} \text{minimizar } \sum_{i=1}^n c_i x_i \\ \text{sujeto a} \\ \sum_{j=1}^n a_{ij} x_j \leq b_i & i = 1, \dots, p \\ \sum_{j=1}^n a_{ij} x_j = b_i & i = p + 1, \dots, m \\ x_j \geq 0 & j = 1, \dots, q \\ x_j \in R & j = q + 1, \dots, n \end{cases}$$

Lo anterior mediante una serie de artificios se puede llevar a una forma equivalente, el cual es llamado el "Problema Estándar de Programación Lineal", el cual es definido como:

$$(PL) = \begin{cases} \text{minimizar } c^t x \\ \text{sujeto a} \\ Ax = b \\ x \geq 0 \end{cases}$$

**Ejemplo 3.1** Toda restricción de igualdad puede expresarse mediante dos restricciones de desigualdad.

$$2x_1 + 3x_2 = 7 \iff \begin{cases} 2x_1 + 3x_2 \leq 7 \\ 2x_1 + 3x_2 \geq 7 \end{cases}$$

**Ejemplo 3.2** Introduciendo una variable adicional, llamada de holgura, toda restricción de desigualdad puede expresarse mediante una restricción de igualdad y otra de no negatividad de la variable de holgura introducida.

$$2x_1 + 3x_2 \leq 7 \iff \begin{cases} 2x_1 + 3x_2 + x_3 = 7 \\ x_3 \geq 0 \end{cases}$$

$$2x_1 + 3x_2 \geq 7 \iff \begin{cases} 2x_1 + 3x_2 - x_3 = 7 \\ x_3 \geq 0 \end{cases}$$

**Ejemplo 3.3** El sentido de la desigualdad cambia multiplicando ambos miembros por -1.

$$2x_1 + 3x_2 \leq 7 \iff -2x_1 - 3x_2 - x_3 \geq -7$$

En un problema de programación lineal se puede observar que.

1. La función objetivo es un hiperplano y, por tanto, un conjunto convexo.
2. Cada restricción es un semiespacio cerrado y, por tanto, un conjunto convexo.



3. Hay un conjunto finito de restricciones y la intersección es un conjunto convexo

En 1961, Philip Wolfe estableció que el siguiente problema

$$(P) \begin{cases} \text{maximizar } f(y) \\ \text{sujeto a} \\ g(y) = \begin{pmatrix} g_1(y) \\ \vdots \\ g_n(y) \end{pmatrix} \leq 0 \end{cases}$$

tiene un problema dual asociado, el cual es definido por

$$(D) \begin{cases} \text{minimizar } f(y) - \sum_{i=1}^n x_i g_i(y) \\ \text{sujeto a} \\ \nabla f(y) = \sum_{i=1}^n x_i \nabla g_i(y) \\ x \geq 0 \end{cases}$$

Donde las funciones  $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$  son convexas y diferenciables para cada  $i \in \{1, \dots, n\}$  y  $f$  es una función cóncava y diferenciable  $\{y \in \mathbb{R}^m / g(y) \leq 0\}$ . Considerando  $f(y) = b^t y$  y  $g_i(y) = A_i^t y - c_i$  para cada  $i \in \{1, \dots, n\}$ , entonces tenemos que las funciones  $f$  y  $g_i$  satisfacen las hipótesis de Wolfe. Además el problema (P) se reduce al siguiente problema al cual llamaremos *Problema Canónico de la Programación Lineal*.

$$(DL) \begin{cases} \text{maximizar } b^t y \\ \text{sujeto a} \\ A^t y \leq c \end{cases}$$

Así también el problema dual de (P), o sea (D), se reduce al formato del problema standard de programación lineal, esto es

$$(DL) \begin{cases} \text{minimizar } c^t x \\ \text{sujeto a} \\ Ax = b \\ x \geq 0 \end{cases}$$

Esto implica que el problema standard de programación lineal (PL) es el dual del problema canónico de programación lineal (DL).

**Proposición 3.1** Sean  $y \in FD = \{z \in \mathbb{R}^m / A^t z \leq c\}$  y  $x \in F = \{x \in \mathbb{R}^n / Ax = b, x \geq 0\}$  cualesquiera, entonces se tiene que  $c^t x \geq b^t y$ .

**Demostración:** Sean  $y \in FD = \{z \in \mathbb{R}^m / A^t z \leq c\}$  y  $x \in F = \{x \in \mathbb{R}^n / Ax = b, x \geq 0\}$  cualesquiera.

$$\begin{aligned} c^t x &= c^t x + (b - Ax)^t y && \text{pues } x \in F \\ &= b^t y + (c - A^t y)^t x && \text{reordenando} \\ &\geq b^t y && \text{pues } y \in FD \end{aligned}$$

**Proposición 3.2** (Condiciones de Karush-Kuhn-Tucker)  $\bar{x}$  es una solución de (PL) si, y solo si, existen  $y \in \mathbb{R}^m$  y  $\bar{s} \in \mathbb{R}^n$  tales que

$$\begin{array}{lll} c - \bar{s} - A^t \bar{y} = 0 & \bar{s} \geq 0 & \text{factibilidad primal} \\ A \bar{x} = b & \bar{x} \geq 0 & \text{factibilidad dual} \\ \bar{s}_i \bar{x}_i = 0 & \forall i \in \{1, \dots, n\} & \text{complementaridad} \end{array}$$

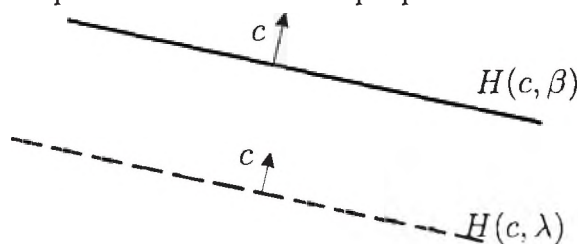
La prueba de esto se realizará en la sección "criterio de optimalidad"

**Comentario 3.1** Observe que:

- $\bar{y} \in FD$ , pues  $c - A^t \bar{y} = \bar{s} \geq 0$
- $b^t \bar{y} = c^t \bar{x}$ , pues  $c^t \bar{x} - y^t A \bar{x} = \bar{s}^t \bar{x} = 0$
- $\bar{y}$  es solución de (DL). pues  $\text{Sup}_{y \in FD} b^t y \leq c^t \bar{x} = b^t \bar{y}$  con  $\bar{y} \in FD$ .

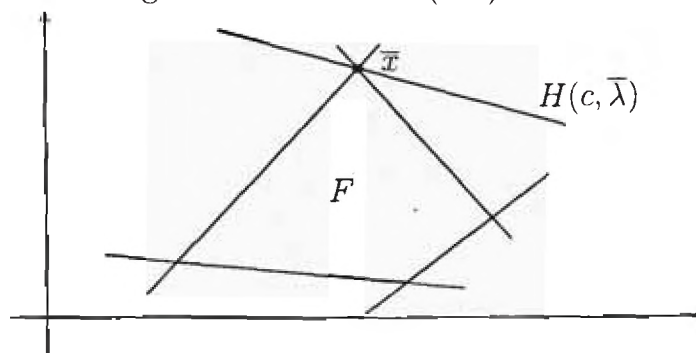
Debemos recordar que para  $\lambda \in \mathbb{R}$  tenemos el hiperplano  $H(c, \lambda) = \{x \in \mathbb{R}^n / \langle c, x \rangle = \lambda\}$  si cambiamos  $\lambda$  por  $\beta$  con  $\beta > \lambda$  obtenemos un hiperplano paralelo al anterior (se desplaza, ver figura 7).

Figura 7: Desplazamiento de un hiperplano en  $\mathbb{R}^2$  ( $\beta > \lambda$ )



Además para el poliedro  $F = \{x \in \mathbb{R}^n / Ax \leq b, x \geq 0\}$ , (ver figura 8) si  $\bar{x}$  es solución de (PL) tal que  $\bar{\lambda} = \langle c, \bar{x} \rangle$  se interpreta como el hiperplano  $H(c, \bar{\lambda})$  tal que  $F \subset H^+(c, \bar{\lambda})$

Figura 8: Solución de (PL) en  $\mathbb{R}^2$



**Definición 3.1**

- $x^*$  es solución óptima si

$$c^t x^* = \min\{c^t x / Ax = b, \quad x \geq 0\}$$

- Un problema de maximización equivale a uno de minimización

$$\max c^t x \Leftrightarrow \min -c^t x$$

Todo programa de maximización (minimización) es convexo, pues la función objetivo, por ser lineal, es cóncava (convexa), y el conjunto de soluciones factibles es convexo, ya que todas las restricciones son lineales. Así por cierto teorema podemos garantizar la solución óptima, si existe, es global.

## 4. MÉTODO DE PUNTOS INTERIORES

Los métodos clásicos para resolver problemas no lineales del tipo  $(P)$ , pueden ser clasificados (según la literatura existente) como: *Métodos de Puntos Interiores* y *Métodos de Puntos Exteriores*. Estas denominaciones fueron dadas porque los primeros tratan de buscar una solución desde el interior del conjunto factible, mientras que los segundos lo hacen desde el exterior del conjunto factible.

Por otro lado, hasta el año 1979, la programación lineal siempre fue considerada independiente de la programación no lineal, pues la programación lineal era considerada como un problema combinatorial y el SIMPLEX usa explícitamente esa estructura combinatorial, debido a que trata de obtener una solución moviéndose a través de sus puntos extremos.

La estrategia del algoritmo de punto interior es implementar, en el espacio original en el que se encuentra el politopo, una serie de transformaciones en el mismo. El politopo original es transformado en cada iteración en otro politopo, en cada paso es transformado buscando la mejor solución y se regresa a su forma original al final del paso (vease [?]).

Consideremos el problema dual de programación lineal escrito en la forma estándar:

$$\begin{aligned} & \text{Máximizarse} \quad b^t y \\ & \text{sujeto a} \quad A^t y \leq c \end{aligned}$$

donde  $y \in \mathbb{R}^m$ ,  $b \in \mathbb{R}^m$  y  $A$  es una matriz  $m \times n$ .

Una vez dado el sistema y un punto inicial, el cual debe ser factible e interior, es decir, debe satisfacer las restricciones  $A^t y < c$ , se busca una manera de movernos en la dirección del gradiente de la función objetivo hacia el próximo punto, el cual debe ser de igual forma factible e interior. Este proceso debe ser repetido hasta que no pueda haber mayor incremento en el valor de la función objetivo.

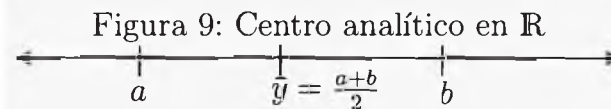
En el presente trabajo, vamos a enfatizar la importancia de dos componentes del

algoritmo de programación lineal de Punto Interior: centrado y direcciones de ascenso. El centrado provee el potencial de una buena reducción del costo en un simple paso, mientras que la dirección de ascenso permite la realización de este potencial. Por lo tanto, nuestro algoritmo de Punto Interior, debe satisfacer al menos estos dos requerimientos.

## 5. CENTRO ANALÍTICO

En esta sección vamos a introducir el concepto de centro analítico para poliedros limitados con interior no vacío, veamos esto con el poliedro limitado con interior no vacío mas simple que es un intervalo cerrado en  $\mathbb{R}$

Consideremos  $P = \{y \in \mathbb{R} / a \leq y \leq b\}$  es decir el intervalo  $[a, b]$ , al centro de este segmento,  $y$  lo llamaremos centro analítico. Esto lo podemos notar en la figura (9).



Ahora veamos una forma de como expresar ese centro analítico como un modelo de optimización.

Supongamos que  $y \in P$ , recordemos que los hiperplanos en  $\mathbb{R}$  son puntos, es decir  $\{a\}$ ,  $\{b\}$  son hiperplanos de  $\mathbb{R}$ , ahora la distancia de  $y$  a esos hiperplanos, las cuales son  $|y - a|$  y  $|b - y|$ . Observe que:

- Si  $y < a$  entonces  $(y - a)(b - y) < 0$
- Si  $y > b$  entonces  $(y - a)(b - y) < 0$
- Si  $a \leq y \leq b$  entonces  $(y - a)(b - y) \geq 0$

Luego, calculemos el valor de  $y$  que maximiza el valor de  $(y - a)(b - y)$  para  $y \in \mathbb{R}$ , según lo observado anteriormente, el máximo valor de  $y$  obtiene en el intervalo  $[a, b]$ . El centro analítico en este caso simple es resolver el siguiente problema de optimización.

$$(P) \begin{cases} \text{maximizar } (y - a)(b - y) \\ \text{sujeto a } y \in \mathbb{R} \end{cases}$$

Como la función logaritmo natural (considerando que  $\text{Ln}(x) = -\infty$  si  $x \leq 0$ ) es creciente, entonces el problema (P) es equivalente a

$$\begin{cases} \text{maximizar } \text{Ln}(y - a) + \text{Ln}(b - y) \\ \text{con } y \in \mathbb{R} \end{cases}$$

notemos que la solución  $y$  esta en el intervalo  $\langle a, b \rangle$  por lo tanto  $y - a$  es la distancia de  $y$  al hiperplano  $a$  y  $b - y$  es la distancia de  $y$  al hiperplano  $b$ . si las distancias fueran cercanas a cero, el logaritmo daría valores cada vez mas negativos, con lo cual

no sería máximo, es por ello que el centro analítico  $\bar{y}$  debe estar lo mas alejado de los hiperplanos  $a$  y  $b$ .

Esto lo vamos a generalizar para cualquier hiperplano en  $\mathbb{R}^n$  con interior no vacío

Definamos el centro analítico del poliedro  $P = \{y \in \mathbb{R}^m / A^t y \leq c\}$  con interior no vacío y acotado, (donde  $A$  es un matriz de orden  $m \times n$  y  $c \in \mathbb{R}^n$ ), al único punto  $y$  solución del problema de optimización

$$\underset{y \in \mathbb{R}^m}{\text{maximizar}} \sum_{i=1}^n \text{Ln}(c_i - \langle a^i, y \rangle)$$

Como el problema anterior es un problema concavo diferenciable entonces existe un resultado en la teoría de la optimización conocido como las condiciones de optimalidad de primer orden.

**Ejemplo 5.1** Consideremos  $f(y) = (y - a)(b - y)$  del ejemplo anterior. Derivando  $f'(y) = 2y - a + b = 0$  de donde obtenemos  $y = \frac{a+b}{2}$  es decir la función  $f$  toma su máximo valor para  $y = \frac{a+b}{2}$  el cual es su centro analítico

## 6. MÉTODO DE PLANOS DE CORTE

El método de planos de corte y centro analítico tiene como puntos de partida el llamado *Método de centros* de Huard que aparece a mediados de la década del 60 y el método elipsoidal de Khachian.

En esta sección describiremos dicho método para resolver problemas de optimización convexa, debido a que:

- Ofrecen una manera diferente de aprovechar la estructura de los problemas grandes y complejos. Un método de planos de corte explota la estructura y puede ser más rápido que un método de punto interior de uso general para el mismo problema.
- No requiere la diferenciabilidad de la función objetivo y de las restricciones.
- Este método no requieren la evaluación de la función objetivo ni de todas las restricciones en cada iteración.
- El método de planos de corte hace posible descomponer los problemas de optimización en problemas más pequeños que pueden ser resueltos de forma secuencial o en paralelo.
- Estos métodos construyen una aproximación lineal de la región factible  $P$  (politopo) reduciéndola en cada iteración.

El objetivo en esta sección es dado un poliedro limitado con interior no vacío, cortarlo usando un hiperplano de tal forma que las partes resultantes (ambas partes son también poliedros limitados con interior no vacío) tengan un volumen inferior al

poliedro original.

Una herramienta para poder ver esto y garantizar que un poliedro existe es el trazado de elipsoide en el interior de un poliedro. Sabemos que en el plano la ecuación de una elipse es  $(\frac{x}{a})^2 + (\frac{y}{b})^2 = 1$  entonces un elipsoide es toda la región en el interior de la elipse incluyendo la elipse

$$E = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 / \left( \frac{x}{a} \right)^2 + \left( \frac{y}{b} \right)^2 \leq 1 \right\}$$

Pero la ecuación se puede escribir matricialmente de esta manera

$$z^t M z \leq 1$$

donde  $M = \begin{pmatrix} \frac{1}{a^2} & 0 \\ 0 & \frac{1}{b^2} \end{pmatrix}$  con  $z = \begin{pmatrix} x \\ y \end{pmatrix}$

Esta idea se generaliza a  $\mathbb{R}^n$  de la forma siguiente

**Definición 6.1** Toda matriz  $M \in \mathbb{R}^{m \times m}$  simétrica y definida positiva (es decir que  $x^t M x > 0 \quad \forall x \neq 0, \quad x \in \mathbb{R}^m$ ) define una elipsoide de la forma siguiente:

Para  $y \in \mathbb{R}^m$  y  $M$  de dimension  $m$  tenemos

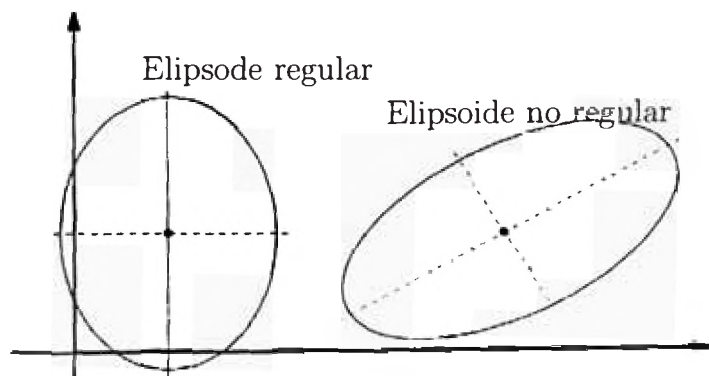
$$E = \{x \in \mathbb{R}^m / (x - y)^t M (x - y) \leq 1\}$$

lo cual es una elipsoide con centro en  $y$ . Cuando  $M$  es diagonal  $E$  es llamado elipsoide regular.

**Ejemplo 6.1** En  $\mathbb{R}^2$  tenemos la elipsoide regular  $3x^2 + 2y^2 \leq 1$  lo cual se puede escribir como

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq 1$$

cuyo centro es  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$



**Comentario 6.1** Usualmente se suele denotar  $s = c - A^t y$ , donde  $S$  es la matriz diagonal generada por el vector  $s$  cuyas componentes son  $s_i = \frac{1}{y_i}$ , es decir que las componentes de la matriz  $S$  son las componentes del vector  $s$ . También se suele utilizar el vector  $e$  cuyas componentes son 1

En la sección anterior se estudio el centro analítico de un poliedro limitado con interior no vacío, el siguiente resultado relaciona el centro analítico con un elipsoide inscrito en dicho poliedro

**Proposición 6.1** Sea  $\bar{y}$  el centro analítico del poliedro  $P$  y  $s = c - A^t y$ , entonces

$$E(\bar{y}) = \{y \in R^m / \|S^{-1} A^t (y - \bar{y})\| \leq 1\}$$

es un elipsoide contenido en  $P$  (ver [11]).

### Demostración

Notemos que  $\|S^{-1} A^t (y - \bar{y})\| \leq 1$  es equivalente a  $\|S^{-1} A^t (y - \bar{y})\|^2 \leq 1$

Pero  $\|S^{-1} A^t (y - \bar{y})\|^2 = (S^{-1} A^t (y - \bar{y}))^t (S^{-1} A^t (y - \bar{y}))$

esto implica que

$$\|(S^{-1} A^t (y - \bar{y}))\|^2 = (y - \bar{y})^t A S^{-2} A^t (y - \bar{y}) \leq 1$$

La simetría de  $A S^{-2} A^t$  es inmediata. Para probar que  $A S^{-2} A^t$  sea definida positiva, tomamos  $w \in R^m$  no nulo, entonces con el mismo argumento anterior, se puede verificar que  $w^t A S^{-2} A^t w = \|S^{-1} A^t w\|^2 > 0$ , pues  $A$  tienen rango  $m$  y  $s > 0$ .

Por lo tanto  $E(\bar{y})$  es un elipsoide. Ahora veamos que  $E(\bar{y}) \subset P$ .

En efecto sea  $y \in E(\bar{y})$  entonces  $\|(S^{-1} A^t (y - \bar{y}))\| \leq 1$ . Haciendo  $s = c - A^t y$  tenemos que

$$\begin{aligned} \|(S^{-1} A^t (y - \bar{y}))\| &= \|(S^{-1} (A^t y - A^t \bar{y}))\| \\ &= \|(S^{-1} (A^t y - c + c - A^t \bar{y}))\| \\ &= \|(S^{-1} (c - A^t \bar{y}) - S^{-1} (c - A^t y))\| \\ &= \|S^{-1} s - S^{-1} s\| = \|e - S^{-1} s\| \leq 1 \end{aligned}$$

Esto implica que  $s \geq 0$  y por lo tanto  $y \in P$ .

Ahora como nuestro objetivo es resolver el siguiente problema

$$\begin{cases} \text{maximizar } b^t y \\ \text{sujeto a } A^t y \leq c \end{cases}$$

donde  $P = \{y \in R^m / A^t y \leq c\}$  es un poliedro limitado con interior no vacío. Si denotamos por  $\bar{y}$  al centro analítico de  $P$  entonces de acuerdo a la proposición anterior existe un elipsoide  $E$  inscrito en  $P$  con centro en  $\bar{y}$ . nuestro objetivo ahora es cortar al poliedro  $P$  con un hiperplano que pase por  $\bar{y}$  y tenga vector normal a  $u = \frac{b}{\|b\|}$ . Este poliedro es definido como  $H(u, \lambda)$  donde  $\lambda = \langle u, \bar{y} \rangle$  este poliedro es dividido en dos

$$P_S = \{y \in R^m / A_s^t y \leq c_s\}$$

$$P_I = \{y \in \mathbb{R}^m / A_I^t \leq c_I\}$$

donde  $A_S = [A_1, \dots, A_n, -u]$ ,  $c_S = -\lambda$ ,  $A_I = [A_1, \dots, A_n, -u]$ ,  $c_I = \lambda$  Sabemos por la sección de programación lineal que el conjunto de soluciones esta en  $P_S$

Consideremos el politopo

$$P = \{y \in \mathbb{R}^m / A^t y \leq c\}$$

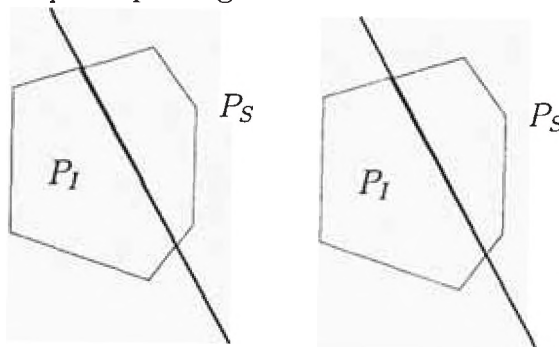
donde  $c \in \mathbb{R}^m$  y  $A \in \mathbb{R}^{m \times n}$  una matriz de máximo rango  $m$ . dado un vector  $a \in \mathbb{R}^m$  y un escalar  $k$ , el nuevo politopo resultante de adicionar el corte  $a^t y \leq k$  es

$$\bar{P} = \{y \in \mathbb{R}^m / a^t y \leq k, \quad y \in P\}$$

Este corte hace que se divida la región factible en dos. El hiperplano  $a^t y = k$  es llamado *plano de corte*, ya que nos da la opción de eliminar el semiespacio  $P_S = \{y / a^t y > k\}$  o el semiespacio  $P_I = \{y / a^t y < k\}$  de nuestra búsqueda.

En general cuando uno traza un corte a un politopo tiene varias opciones para realizar el corte y también para elegir el nuevo politopo como lo vemos en la figura 10

Figura 10: Luego del corte tenemos dos opciones para elegir el nuevo politopo. En ambos casos reducimos el politopo original



**Ejemplo 6.2** Para

$$A^t = \begin{pmatrix} 2 & 1 \\ -2 & 1 \\ 0 & -1 \end{pmatrix} \quad c = \begin{pmatrix} 4 \\ 5 \\ -1 \end{pmatrix} \quad a^t = (2, 1) \quad k = 3$$

tenemos el siguiente politopo

$$P = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 / \begin{pmatrix} 2 & 1 \\ -2 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} 4 \\ 5 \\ -1 \end{pmatrix} \right\}$$

al trazar el corte definido por

$$(2 \ 1) \begin{pmatrix} x \\ y \end{pmatrix} \leq 3$$

obtenemos un nuevo politopo como se muestra en la figura 11

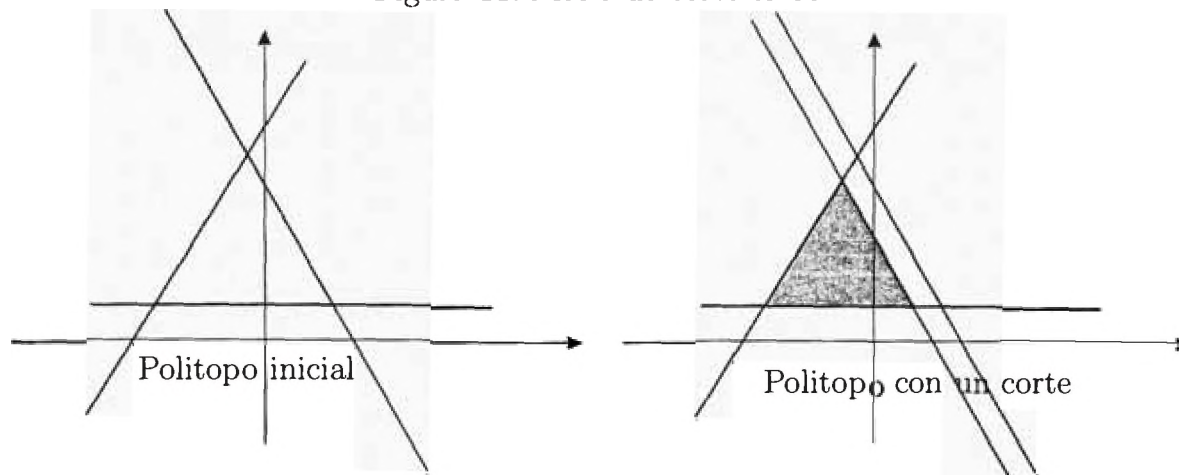


Lo anterior también se puede colocar como

$$\begin{pmatrix} 2 & 1 \\ -2 & 1 \\ 0 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \leq \begin{pmatrix} 4 \\ 5 \\ -1 \\ 3 \end{pmatrix}$$

**Comentario 6.2** *Agregar un corte equivale a agregar un restricción.*

Figura 11: Plano de corte en  $\mathbb{R}^2$



**Ejemplo 6.3** *Para el problema de maximizar*

$$f(x, y) = 4x + y$$

*sujeto a*

$$3x + y \leq 12$$

$$2x - y \leq -2$$

*obtenemos de forma equivalente*

$$\text{Maximizar } z = \begin{pmatrix} 4 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \text{ sujeto a:}$$

$$\begin{pmatrix} 3 & 1 \\ 2 & -1 \\ -1 & 0 \end{pmatrix} \leq \begin{pmatrix} 12 \\ -2 \end{pmatrix}$$

*Si además trazamos el hiperplano definido por  $H = \{x \in \mathbb{R}^n / c^t x = 0\}$  obtenemos la figura (12),*

*Si movemos el hiperplano en sentido del gradiente de la función objetivo  $\nabla f = (4, 1)$  llegaremos a encontrar la solución óptima.*

*Movamos el hiperplano 6 unidades según el gradiente y como podemos notar en la figura (13) hay dos opciones para elegir el poliedro a trabajar, elegimos el que indique el gradiente de la función objetivo. Es decir tracemos el plano de corte  $H = \{x \in \mathbb{R}^2 / c^t x \geq 6\}$*

En el presente trabajo vamos a realizar un corte por un punto que este lo mas alejado de los hiperplanos pero que sea un punto interior, ya que de esta manera podremos reducir mas rápidamente el polítopo original.

Figura 12: Gráfico del poliedro y la función objetivo

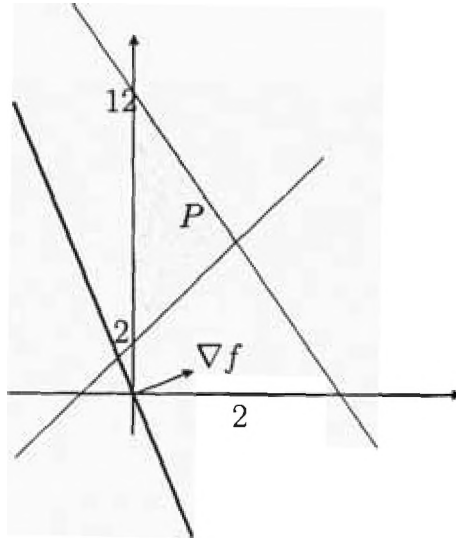
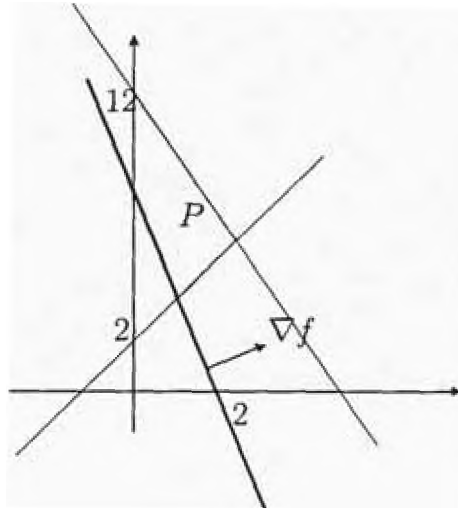


Figura 13: Función objetivo desplazada 6 unidades en sentido del gradiente  $\nabla f$



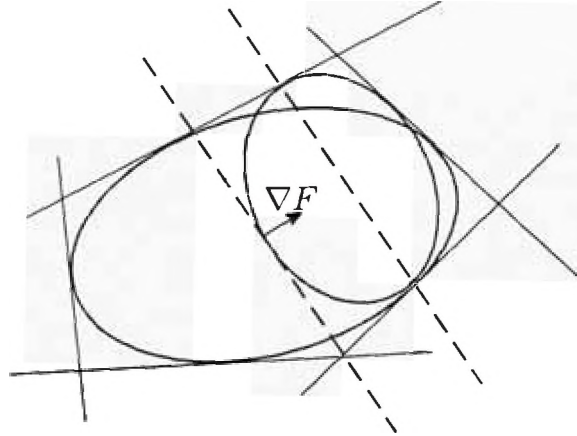
## 7. CENTRO ANALÍTICO EN $\mathbb{R}^n$

Existen varios algoritmos para resolver problemas de optimización que utilizan en cada iteración un punto considerado como centro, o también llamado centro analítico. En general los métodos basados en los centros, realizan un corte en una región convexa acotada que contiene una solución óptima. Cada corte utiliza un hiperplano, que separa la región en dos partes, elegimos la parte donde la solución del problema está contenida y continuamos realizando operaciones de acuerdo con el método que se utiliza para determinar el próximo centro.

Los diversos métodos existentes se diferencian en cómo se determinan los centros y la velocidad de convergencia de estos para determinar dicho centro (véase [13]).

En este trabajo presentaremos una versión simplificada del método ACCPM para

Figura 14: Trazado de dos elipsoides y dos cortes en el poliedro



resolver el problema canónico de la programación lineal el cual se formula como:

$$(DL) \quad \begin{cases} \text{maximizar } b^t y \\ \text{sujeto a} \\ A^t y \leq c \end{cases}$$

Donde  $A \in \mathbb{R}^{m \times n}$ , con rango igual a  $m$  ( $m < n$ ),  $c \in \mathbb{R}^n$ ,

$$P = \{y \in \mathbb{R}^m / A^t y \leq c\}$$

acotado y de interior no vacío.

Definamos el centro analítico del poliedro  $P$ , (definido por  $A^T$  y  $c$ ), como el único punto  $\bar{y}$  solución del problema de optimización

$$\text{máximizarse}_{y \in \dot{P}} \left( \sum_{i=1}^n \text{Ln}(c_i - \langle (a^i)^t, y \rangle) \right)$$

Como se supone que  $\dot{P} = \{y \in \mathbb{R}^{m \times 1} / A^t y < c\} \neq \emptyset$ , entonces  $c_i - \langle (a^i)^t, y \rangle$  siempre es positivo y el logaritmo está bien definido.

A la función

$$\phi(P) = \sum_{i=1}^n \text{Ln}(c_i - \langle (a^i)^t, y \rangle)$$

la llamaremos **función potencial**, la cual no está definida para puntos en la frontera de  $P$  (para algún  $i$  tal que  $\langle (a^i)^t, y \rangle = c_i$ ). Cuando un punto interior está cerca de la frontera, entonces  $c_i - \langle (a^i)^t, y \rangle$  es positivo pero cercano a cero y la función potencial tiene un valor cada vez más negativo. Como el centro analítico maximiza la función potencial, debe estar "lo más alejado posible" de la frontera.

Más adelante se verá que el centro analítico depende específicamente de  $A^t$  y de  $c$  pero no depende exactamente del poliedro  $P$ . Es decir, puede haber dos definiciones que den lugar al mismo conjunto  $P$  pero den lugar a dos centros analíticos diferentes.

## 7.1. Condiciones de optimalidad

Sea  $s \in \mathbb{R}^n$  el vector de variables de holgura, es decir,

$$A^t y + s = c \quad (2)$$

En  $\dot{P}$  se cumple  $s > 0$ . Sea  $x \in \mathbb{R}^n$  definido por

$$x_i = \frac{1}{s_i}$$

Algunas veces, se denota simplemente  $x = s^{-1}$ . Como  $s > 0$ , entonces  $x > 0$ . Sea  $e$  el vector columna de unos de tamaño adecuado. En métodos de punto interior es usual la siguiente notación: si  $x \in \mathbb{R}^n$ ,  $X$  es una matriz diagonal cuyos elementos son los  $x_i$ ,

$$X = \begin{pmatrix} x_1 & 0 & \dots & 0 \\ 0 & x_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & x_n \end{pmatrix} \quad e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Así la definición de  $x$  se puede presentar como  $x = S^{-1}e$ , o también,

$$Xs = e \quad (3)$$

Calculemos el gradiente de la función potencial:

$$\phi(P) = \max_{y \in \dot{P}} \left( \sum_{i=1}^n \ln(c_i - \langle (a^i)^t, y \rangle) \right)$$

Derivando sucesivamente

$$\frac{\partial \phi(P)}{\partial y_j} = \sum_{i=1}^n \frac{1}{c_i - \langle (a^i)^t, y \rangle} (-1)(a^i)^t$$

$$\frac{\partial \phi(P)}{\partial y_j} = \sum_{i=1}^n \frac{1}{s_i} (-1)(a^i)^t$$

$$\frac{\partial \phi(P)}{\partial y_j} = - \sum_{i=1}^n x_i (a^i)^t$$

$$\frac{\partial \phi(P)}{\partial y_j} = - \sum_{i=1}^n A_i x_i$$

Por el comentario 2.4

$$\phi'(y) = -Ax$$

Como el centro analítico (maximizador) es un punto interior, entonces su gradiente debe ser nulo

$$Ax = 0 \quad (4)$$

Agrupando las condiciones de factibilidad y optimalidad, en las variables  $x$ ,  $y$  y  $s$ , se tienen  $m + n + n$  ecuaciones con  $n + m + n$  incógnitas:

De (4):  $Ax = 0, \quad x > 0$

De (2):  $A^t y + s - c = 0, \quad s > 0$

De (3):  $Xs - e = 0$

Se denotará con  $(x^a, y^a, s^a)$  a la tripla que cumpla estas condiciones.

Realmente el centro analítico es  $y^a$ . Para saber si un punto  $\bar{y}$  es el centro analítico, se pueden seguir los siguientes pasos:

1. calcular  $s = c - A^t \bar{y}$
2. verificar que  $s > 0$
3. calcular  $x = S^{-1} e$
4. verificar que  $Ax = 0$

Supongamos ahora que se desea saber si a partir de un vector  $\bar{x}$  se puede obtener el centro analítico. Dicho de otra forma,  $\bar{x}$  ¿Es centro analítico?. Se pueden seguir los siguientes pasos:

1. verificar que  $\bar{x} > 0$
2. verificar que  $A\bar{x} = 0$
3. calcular  $s = \bar{X}^{-1} e$
4. verificar que el sistema  $A^t y = c - s$  tiene solución
5. La solución del sistema es el centro analítico.

**Ejemplo 7.1** Para el poliedro  $P$  limitado por

$$\begin{pmatrix} 2 & 1 \\ -2 & 1 \\ 0 & -1 \end{pmatrix} y \leq \begin{pmatrix} 1 \\ 5 \\ -1 \end{pmatrix}$$

obtenemos  $\bar{y} = \begin{pmatrix} -1 \\ \frac{5}{3} \end{pmatrix}$  el cual es el centro analítico de  $P$  el cual es un triángulo con vértices  $(0, 1)$ ,  $(-2, 1)$  y  $(-1, 3)$ .

$$s = c - A^t \bar{y}$$

$$s = \begin{bmatrix} 4 & 4 & 2 \\ 3 & 3 & 3 \end{bmatrix}^t$$

$$x = S^{-1} e$$

$$x = \begin{bmatrix} 3 & 3 & 3 \\ 4 & 4 & 2 \end{bmatrix}^t$$

$$Ax = [0 \quad 0].$$

luego  $y = \begin{pmatrix} -1 \\ \frac{5}{3} \end{pmatrix}$  es el centro analítico.

**Ejemplo 7.2** Hallemos el centro analítico de la región definida por las restricciones

$$\begin{pmatrix} 2 & 1 \\ -2 & 1 \\ 0 & -1 \\ 0 & 1 \end{pmatrix} y \leq \begin{pmatrix} 1 \\ 5 \\ -1 \\ 3 \end{pmatrix}$$

De nuevo, el conjunto  $P$  es el triángulo con vértices  $(0,1)$ ,  $(-2,1)$  y  $(-1,3)$ . Sea  $\bar{y} = [-1 \frac{3}{2}]^t$ ,

$$\begin{aligned} s &= c - A^t \bar{y} \\ s &= \begin{bmatrix} 3 & 3 & 1 & 3 \\ 2 & 2 & 2 & 2 \end{bmatrix}^t \\ x &= S^{-1} e \\ x &= \begin{bmatrix} 2 & 2 & 2 & 2 \\ 3 & 3 & 2 & 3 \end{bmatrix}^t \\ Ax &= [0 \quad 0]. \end{aligned}$$

luego  $\bar{y} = [-1 \frac{3}{2}]^t$  es el centro analítico. En estos dos ejemplos, el conjunto  $P$  es el mismo, las restricciones son diferentes y los centros analíticos también.

## 7.2. Algoritmo ACCPM

La idea de este algoritmo es construir una sucesión de poliedros que contienen al conjunto de soluciones de (DL) y que en volumen tienden a cero. Para construir esta sucesión de poliedros en la primera iteración se calcula el centro analítico exacto (o aproximado de FD) y luego se hace pasar por ese punto un plano de corte

$$H_0 = \{y \in R^m / b^t y = b^t \bar{y}\}$$

donde  $\bar{y}$  es el centro exacto o aproximado de  $FD$ . Luego se escoge el poliedro que resulta de intersectar  $FD$  con

$$H_0^+ = \{y \in R^m / b^T y \geq b^T \bar{y}\}$$

denotando a esta intersección como  $FD^1$ . A partir de la segunda iteración ( $k \geq 1$ ) se calcula el nuevo centro analítico del nuevo poliedro  $FD^k$  y por ese nuevo centro se pasa un nuevo plano de corte

$$H^k = \{y \in R^m / b^T y = b^T y^k\}$$

siendo  $y^k$  el centro analítico exacto (o aproximado) de  $FD^k$ , luego se considera el poliedro que resulta de intersectar  $FD^k$  con  $H_k^+$ . Lógicamente este método converge a una solución de (DL) ya que las elipses contenidas en el poliedro son cada vez mas pequeñas tendiendo a cero.

### ALGORITMO ACCPM

Datos de entrada.

$m, n \in \mathbb{N}, b \in \mathbb{R}^m, c \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, k = 0, A^k = A^T, c^k = c$  y  
 $FD^k = \{y \in \mathbb{R}^m / A^k y \leq c^k\}$ .

**Paso 1.**

Calcular  $y^k$  (centro analítico aproximado de  $FD^k$ ).

**Paso 2.**

Actualizar.

$$A^{k+1} = [A - b]^T,$$

$$c^{k+1} = \begin{pmatrix} c^k \\ -b^t y^t \end{pmatrix}$$

$$FD^{k+1} = \{y \in \mathbb{R}^m / A^{k+1} y \leq c^{k+1}\}.$$

**Paso 3.**

Hacer  $k = k + 1$

Si  $\|y^k - y^{k-1}\| < \varepsilon$  terminar y mostrar  $y^k$

sino ir al **paso 1.**

## 8. RESULTADOS NUMÉRICOS

A continuación vamos a mostrar un resultado del programa elaborado en Scilab, vamos a resolver el problema

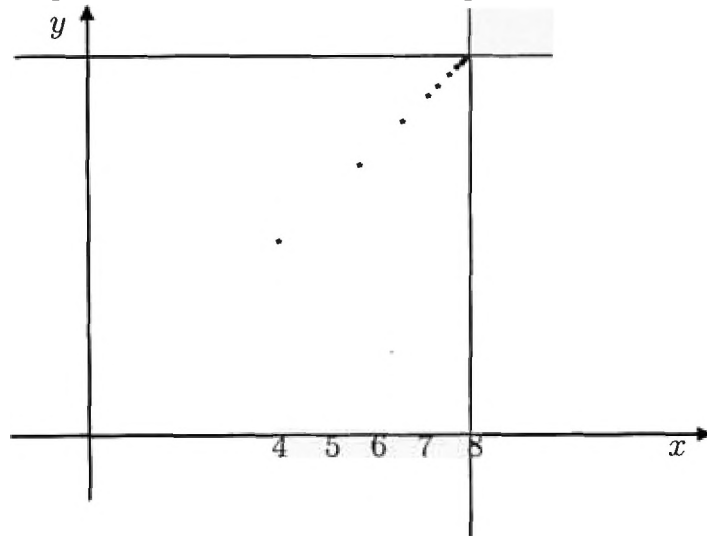
$$\begin{aligned} & \text{maximizar } z = x + y \\ & \text{sujeto a} \\ & x \leq 8 \\ & y \leq 8 \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

El cual sólo es un ejemplo para poder observar lo que ocurre al 'correr' el programa elaborado con el algoritmo planteado.

Iteración	Centro Analítico $\bar{y}$
1	(4.0000 , 4.0000)
2	(5.7888 , 5.7888)
3	(6.6421 , 6.6421)
4	(7.1346 , 7.1346)
5	(7.4383 , 7.4683)
6	(7.6317 , 7.6317)
7	(7.7571 , 7.7571)
8	(7.8392 , 7.8392)
9	(7.8933 , 7.8933)
10	(7.9290 , 7.9290)

Como vemos el centro analítico se va aproximando a  $(8,8)$  el cual es la solución para el problema planteado. Dichos resultados lo podemos apreciar en la figura (15)

Figura 15: 10 iteraciones del algoritmo ACCPM





## 9. CONCLUSIONES Y RECOMENDACIONES

El presente trabajo: 'Métodos de plano de corte y centro analítico' ha sido desarrollado con una doble finalidad, por un lado definir el problema lo mejor posible para su buen entendimiento y por el otro mostrar una alternativa de resolución de problemas de Programación Lineal.

La esencia de este método es la simplicidad, el contenido geométrico del algoritmo es notable y en su construcción se emplean nociones elementales de Álgebra Lineal.

Se recomienda que sean desarrollados en tesis de universidades para una implementación de acuerdo a problemas particulares, es en este sentido que esta monografía fue desarrollada y orientada a cualquier estudiante de ciencias e ingeniería.

En el presente trabajo se explota la geometría propia del concepto de convexidad que ayuda a entender mejor el método en cuestión. Cualquier alumno de ciencias e ingeniería basándose en esta monografía es capaz de desarrollar los códigos en cualquier lenguaje de programación con la idea de adaptar este método a casos particulares.

La base matemática es enunciada de forma sistemática, para hacer comprensible la presente monografía. Es decir este trabajo no es un tratado matemático sino una presentación matemática del método para que luego pueda ser aplicado a los problemas que crean conveniente.

Se recomienda usar el método ACCPM cuando la cantidad de variables sea muy grande ya que es allí donde se aprovechará dicho método.

El centro analítico depende específicamente de las restricciones ( $A^t y \leq c$ ) pero no depende exactamente del poliedro  $P$  que se forma de las restricciones. Es decir, puede haber dos definiciones que den lugar al mismo poliedro  $P$  pero den lugar a dos centros analíticos diferentes.

Al tener restricciones repetidas o redundantes, el centro analítico es diferente ya que esta restricción redundante 'empuja' el centro analítico hacia la solución óptima.

## 10. APENDICE

### Programa ACCPM

```
function [x,fval]=mcapc(b,A,c)
// Input:  b      vector de la función objetivo a maximizar
//         A,c    poliedro acotado con int. no vacío (Ax<=c)
// Output: fval  el valor que maximiza la función objetivo
//         x     vector que maximiza la función

e=1;
[fil col]=size(A);

while e<11
    [m n]=size(A);

    if m>fil+1
        A(m-1,:)=[];
        c(m-1)=[];
    end

    [y,l]=can(c,A)
    if l==1 then
        disp('Poliedro vacío')
    end
    [y]=grad(A',c,x)  %c ax<=c
    A=[A ;-b'];
    c=[c ;-b'*y];
    e=e+1;
end
x=y;
fval=b'*y;
endfunction
```

Programa para realizar las iteraciones de Newton

```
function y=itnewton(x,c,A)
//Input      x  matriz columna
//           a  matriz columna
//           A  matriz de coeficientes
//Output     y  valor que resulta de una iteración de Newton

    s=c-A*x;
    is=s^(-1);
    S=diag(is);
    B=A';
    g=B*is;      // Gradiente de la función potencial
    H=B*S*S*A   //Hessiano de la F.P.
    C=inv(H)
    y=x-C*g;
endfunction
```

Programa para hallar el centro analítico de  $Ax \leq c$

```
function [z,l]=can(c,A)
// Halla el centro analítico de
    [m n]=size(A);
    v=0;
    Aleat=rand(n,1)*10;
    [d,l]=pint(Aleat,c,A);
    if l<100 then
        for i=1:10
            Aleat=itnewton(d,c,A);
            d=Aleat;
        end
    else
        v=1
    end
    z=d;
    l=v;
endfunction
```

Programa para hallar el punto interior de un poliedro acotado de interior no vacío

```
function [z,l]=pint(x,a,A)
// input    x    Un punto cualesquiera
//          a    Lado derecho
//          A    Matriz de coeficientes de restricciones
// Output   z    Punto Interior
//          l    valor que indica si se encontró un punto int.

k=1;
iter=0;
[m n]=size(A);
while k==1 do
    k=0;
    iter=iter+1;
    for j=1:m
        f=A(j,:);
        d=a(j)-f*x;
        if d>0 then
            ta(j)=a(j)
        else
            k=1;
            ta(j)=f*x+0.1
        end
    end
    if k==1 then
        for i=1:6
            y=itnewton(x,ta,A);
            x=y;
        end
    end
    if iter > 10000 then
        k=0
    end
end
z=x;
l=iter
endfunction
```

# Bibliografía

- [1] Berkovitz L.D. *Convexity and Optimization in  $\mathbb{R}^n$* . Wiley 2002
- [2] Boldrini José Luis. *Álgebra linear*. tercera ed. Harper e Row do Brasil 1980.
- [3] Boyd Stephen, Lieven Vandenberghe . *Convex Optimization*. Cambridge University Press
- [4] Callioli Carlos, Domingues Hygino, Costa Roberto. *Álgebra linear e aplicações*. Atual Editora.
- [5] Jean Pierre Crouzeix, Abdelkrim Keraghel, Wilfredo Sosa Sandoval. *Programación Matemática Diferenciable*. IMCA 2010
- [6] Domingues Hygino. *Espaços Métricos e introdução à topologia*. Atual Editora.
- [7] Gonçalves Adilson. *Introdução à álgebra linear*.
- [8] Izmailov Alexey, Mikhail Solodov. *Otimização*, volume 1. IMPA 2005
- [9] Lay David. *Algebra lineal y sus aplicaciones* segunda edición. Addison Wesley.
- [10] Luenberger David, Yinyu Ye. *Linear and Nonlinear Programming* tercera ed. Springer.
- [11] Sosa S. Wilfredo. *Introducción a la Optimización Programación Lineal*. Sociedad Matemática Peruana 2010.
- [12] Stojanoff Demetrio. *Análisis Funcional vs. Matricial*.
- [13] Yinyu Ye. *Interior Point Algorithm: Theory and Analysis*. Wiley - Interscience Series in Discrete Mathematics Optimization, John Wiley and Sons, New York.