UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE CIENCIAS

TESIS

Relación de Inercias entre dos Representaciones de Subespacios Lineales Monótonos

PARA OBTENER EL GRADO DE MAESTRO EN CIENCIAS
CON MENCIÓN EN MATEMÁTICA APLICADA

ELABORADO POR:

Luis Ernesto Flores Luyo

ASESOR:

Dr. Eladio Ocaña Anaya

Lima - Perú

2015

Dedicatoria

A mis Padres
Luis Flores e Isabel Luyo

AGRADECIMIENTOS

Agradezco profundamente a mi asesor, Dr. Eladio Ocaña Anaya no solamente por ayudarme y aconsejarme pacientemente a través de esta Tesis, sino que también por sus aclaraciones y valiosos consejos.

Además agradezco a los profesores de la FC-UNI por brindarme la gran oportunidad de ser parte de esta institución, y también a todos los profesores y compañeros de estudios del IMCA y de la Universidad Nacional de Ingeniería.

RESUMEN

Sabemos que un subespacio vectorial $E \subset \mathbb{R}^n \times \mathbb{R}^n$ puede ser expresado de dos formas distintas, como núcleo y como imagen de transformaciones lineales, esto es

$$E = \{(x, x^*) : Ax + Bx^* = 0\} \quad \text{y} \quad E = \{(x, x^*) : x = Pu,\ x^* = Qu,\ u \in \mathbb{R}^r\}$$

para algunas matrices $A, B \in \mathbb{R}^{p\times n}$ y $P, Q \in \mathbb{R}^{n\times r}$.

En este trabajo estudiaremos la relación entre ambas representaciones cuando el subespacio E es considerado monótono. Se establecerá esta relación por medio de las inercias de las matrices simétricas $(AB^t + BA^t)$ y $(P^tQ + Q^tP)$.

Contenido

Introducción

Dado un operador $\Gamma : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ (para cada $x \in \mathbb{R}^n$, el valor correspondiente $\Gamma(x)$ es un subconjunto de $\mathbb{R}^n$), se dice que este es monótono si satisface la siguiente propiedad:

$$\langle x^* - y^*, x - y\rangle \geq 0 \;\text{ para todo } (x, x^*), (y, y^*) \in \operatorname{graph}(\Gamma),$$

donde $\operatorname{graph}(\Gamma)$ denota el gráfico de Γ definido por

$$\operatorname{graph}(\Gamma) := \{(x, x^*) \in \mathbb{R}^n \times \mathbb{R}^n : x^* \in \Gamma(x)\}.$$

Se deduce inmediatamente de la definición que un operador Γ es monótono si y solo si el operador (inverso) $\Gamma^{-1} : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ definido por $x \in \Gamma^{-1}(x^*)$ si y solo si $x^* \in \Gamma(x)$, es monótono.

Esta equivalencia nos permite deducir que la monotonía es una propiedad no solamente sobre el operador Γ sino más bien sobre el gráfico de este.

De otro lado, debido a todo subconjunto E de $\mathbb{R}^n \times \mathbb{R}^n$ se puede expresar como el gráfico de un cierto operador, este nos permite extender la propiedad de monotonía a estos subconjuntos. De forma natural diremos que E es monótono si

$$\langle x^* - y^*, x - y\rangle \geq 0 \;\text{ para todo } (x, x^*), (y, y^*) \in E.$$

La monotonía está relacionada con la convexidad de funciones en el sentido siguiente: una función $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ es convexa si y solo si el operador $\Gamma : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ definido por

$$\Gamma(x) := \{x^* : f(x) + \langle x^*, y - x\rangle \leq f(y) \;\text{ para todo } y \in \mathbb{R}^n\},$$

es monótono.

Así como la convexidad de funciones es de relevancia importante en los problemas de optimización, la monotonía tiene el mismo efecto en los problemas de Desigualdad Variacional (PDV): Dado $\Gamma : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ y $C \subset \mathbb{R}^n$, el PDV asociado consiste en

$$\text{Hallar } x \in C \;\text{ y }\; x^* \in \Gamma(x) \;\text{ tal que }\; \langle x^*, y - x \rangle \geq 0 \;\text{ para todo } y \in C.$$

Similar al problema de optimización convexa, cuando el operador Γ es monótono y el conjunto C convexo, el conjunto solución (eventualmente vacío) del PDV correspondiente, es convexo.

No obstante, reconocer la monotonía de un operador es tan complicado (o tal véz más) como reconocer la convexidad de una función.

Al menos en el caso lineal, es decir cuando el gráfico del operador Γ es un subspacio lineal, existe una manera explícita de reconocer la monotonía del operador en términos de los autovalores de las matrices asociadas al operador. Ver por ejemplo [1], [2].

El desarrollo de la tesis se divide en tres capítulos:

En el primer capítulo presentamos el problema que trataremos en este trabajo, también damos algunos preliminares que serán usados posteriormente.

En el segundo capítulo desarrollamos el tema central de la tesis, obtenemos una una matriz inyectiva $(P^t, Q^t)^t$ la cual está relacionada con la matriz proyección sobre E, $(P^t, Q^t)^t = \Pi K$, para cierto K, y a continuación encontramos la inercia de $P^tQ + Q^tP = K\Pi F \Pi K$ por medio de la inercia de la matriz $\Pi F \Pi$ y observaremos que está relacionado con la inercia de la matriz $(AB^t + BA^t)(AA^t + BB^t)^{-1}$ y esta última a su vez con la inercia de la matriz $(AB^t + BA^t)$.

En el tercer capítulo presentamos un ejemplo en el que corroboramos cada resultados obtenido en el capítulo anterior.

Capítulo 1

Preliminares y Planteamiento del Problema

A través de este trabajo asumiremos que $\mathbb{R}^n$ ($n \in \mathbb{N} = \{1, 2, 3, \ldots\}$) es un espacio euclideano con producto interno $\langle \cdot, \cdot \rangle$, y $\| \cdot \|$ la norma euclideana inducida.

Se sabe que un subespacio vectorial $E \subset \mathbb{R}^n \times \mathbb{R}^n$ puede ser representado de dos maneras: como imagen y como núcleo de transformaciones lineales, esto es

$$E = \operatorname{img}(T_1) \quad \text{y} \quad E = \ker(T_2),$$

donde $T_1 : \mathbb{R}^r \to \mathbb{R}^n \times \mathbb{R}^n$ y $T_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^p$ son transformaciones lineales.

En términos matriciales, la primera puede representarse como

$$E = \{(x, x^*) : x = Pu,\ x^* = Qu,\ u \in \mathbb{R}^r\} \tag{Im}$$

para ciertas matrices $P, Q \in \mathbb{R}^{n \times r}$ donde, sin pérdida de generalidad, podemos asumir que la matriz $(P^t, Q^t)^t$ es inyectiva.

La segunda puede representarse como

$$E = \{(x, x^*) : Ax + Bx^* = 0\} \tag{Ker}$$

para ciertas matrices $A, B \in \mathbb{R}^{p \times n}$ donde, sin pérdida de generalidad, podemos asumir que la matriz (A, B) es sobreyectiva.

Con los supuestos arriba, tenemos que $\dim(E) = 2n - p = r$.

De la definición de monotonía, reconocer esta propiedad cuando el subespacio E viene representado por la primera expresión (Im) es mucho más fácil que con la segunda (Ker). En efecto, con la expresión (Im), E es monótono si y solo si

$$\langle (P^tQ + Q^tP)u, u\rangle = \langle Pu, Qu\rangle \geq 0 \quad \text{para todo } u \in \mathbb{R}^r$$

el cual significa que la matriz $P^tQ + Q^tP$ es semidefinida positiva.

Referente a la expresión (Ker), en [1] se establece que E es monótono si y solo si $p \geq n$ y la matriz $AB^t + BA^t$ tiene exactamente $p - n$ autovalores positivos (contando las multiplicidades).

Estas dos caracterizaciones se pueden expresar usando la inercia de una matriz.

Definición 1.0.1 (Inercia de una matriz) *Sea $M \in \mathbb{R}^{m\times m}$ una matriz simétrica. La inercia de M, denotada por $\mathrm{In}\,(M)$, se define como la terna*

$$\mathrm{In}\,(M) := (\mu_+(M), \mu_-(M), \mu_0(M)),$$

donde $\mu_+(M)$, $\mu_-(M)$ y $\mu_0(M)$ denotan respectivamente las cantidades de autovalores positivos, negativos y ceros (contando las multipicidades) de M.

Se deduce de la definición que $\mu_+(M) + \mu_-(M) + \mu_0(M) = m$.

Se deduce también directamente las siguientes equivalencias:

$$E \ \text{ es monótono} \iff \mu_-(P^tQ + Q^tP) = 0 \iff \mu_+(AB^t + BA^t) = p - n.$$

En contraste a estas caracterizaciones de la monotonía en la que las relaciones de las inercias de las matrices $P^tQ + Q^tP$ y $AB^t + BA^t$ sólo expresan las conecciones entre μ_- de la primera matriz y μ_+ de la segunda, en este trabajo estableceremos las relaciones de todas la componentes de ambas inercias

$\mathrm{In}\,(P^tQ + QtP)$ e $\mathrm{In}\,(AB^t + B^tA)$ cuando E es considerado monótono. En contraste a lo establecido en [1], mostramos también las relaciones de $\mu_+(P^tQ+Q^tP)$ y $\mu_0(P^tQ + Q^tP)$ con $\mu_-(AB^t + BA^t)$ y $\mu_0(AB^t + BA^t)$. En efecto, mostraremos que

$$\mathrm{In}\,(AB^t + BA^t) = (p-n, p-n+l, 2n-p-l) \iff \mathrm{In}\,(P^tQ+Q^tP) = (l, 0, 2n-p-l),$$

donde l es la dimensión del espacio generado por los autovectores correspondientes a autovalores distintos de cero de la matriz $\Pi F \Pi$, siendo Π la proyección ortogonal de $\mathbb{R}^n \times \mathbb{R}^n$ sobre E, y $F = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix}$.

Algunos resultados muy importantes son los siguientes teoremas:

Teorema 1.0.1 *Los autovalores de una matriz Hermítica (en particular simétrica) son todos reales y autovectores correspondientes a dos autovalores diferentes son ortogonales.*

Teorema 1.0.2 *(Ley de Inercia Lagrange-Sylvester). Para dos matrices* M *y* Q $m \times m$ *con* M *simétrica y* Q *no singular, tenemos*

$$\mathrm{In}\,(\mathrm{M}) = \mathrm{In}\,(\mathrm{Q^tMQ}).$$

Para la demostración de este Teorema es necesario las siguientes definiciones y los siguientes resultados:

1.1 Formas Bilineales

Definición 1.1.1 *Una forma bilineal de* $\mathbb{R}^n$ *es una aplicación* $B : \mathbb{R}^{\ltimes} \times \mathbb{R}^n \to \mathbb{R}$ *que verifica que, para todo* $\lambda, \mu \in \mathbb{R}$ *y para todo* $u, v, w \in \mathbb{R}^n$*,*

$$B(\lambda u + \mu v, w) = \lambda B(u, w) + \mu B(v, w);$$

$$B(u, \lambda v + \mu w) = \lambda B(u, v) + \mu B(u, w)$$

Es decir, que es lineal en cada una de las dos variables vectoriales, lo que significa que, fijando arbitrariamente el valor de una de ellas y permitiendo variar la otra, se obtiene una forma lineal de $\mathbb{R}^n$.

Observación 1.1.1 *Para todo* $u \in \mathbb{R}^n$*,* $B(u, 0) = B(0, u) = 0$.

Observación 1.1.2 *Si* $\nu = (u_1, \ldots, u_n)$ *es una base de* $\mathbb{R}^n$*, entonces, para cualesquiera vectores* $u = \sum_{j=1}^n x_j u_j$ *y* $v = \sum_{j=1}^n y_j u_j$*, se tiene que*

$$B(u, v) = \sum_{i=1}^n x_j B(u_i, \sum_{j=1}^n y_j u_j) = \sum_{i=1}^n x_i \sum j = 1^n y_j B(u_i, u_j),$$

es decir, los n^2 *números reales* $B(u_i, u_j), i, j = 1, 2, \ldots, n$ *determinan completamente la forma bilineal.*

Podemos, pues, expresar matricialmente la forma bilineal de la siguiente manera:

$$B(u,v) = (x_1,\dots,x_n)\begin{pmatrix} B(u_1,u_1) & \dots & B(u_1,u_n) \\ \vdots & \ddots & \vdots \\ B(u_n,u_1) & \dots & B(u_n,u_n)\end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n\end{pmatrix} = ([u]^\nu)^t[B]_\nu[v]^\nu$$

La matriz cuadrada $[B]_\nu$ de la expresión anterior se denomina matriz de la forma bilineal B respecto a la base ν.

Definición 1.1.2 *Una forma bilineal B de $\mathbb{R}^n$ se dice simétrica si para todo $u,v \in \mathbb{R}^n, \quad B(u,v) = B(v,u)$.*

Observación 1.1.3 *Dada cualquier base ν de $\mathbb{R}^n$, la forma bilineal B es simétrica si y solo si su matriz respecto a ν lo es.*

Observación 1.1.4 *Si B es simétrica, queda completamente determinada por los $\dfrac{n(n+1)}{2}$ números reales $B(u_i,u_j),\ 1 \le j \le i \le n$*

1.2 Formas cuadráticas y congruencia de matrices

1.2.1 Formas cuadráticas

Definición 1.2.1 *Dada una forma bilineal B de $\mathbb{R}^n$, se define su* ***forma cuadrática*** *asociada como la aplicación $f_B : \mathbb{R}^n \to \mathbb{R}$ definida por*

$$\forall u \in \mathbb{R}^n, \quad f_B(u) = B(u,u).$$

Si $(x_1,\dots,x_n)$ representa el vector de coordenadas de u respecto a una determinada base de $\mathbb{R}^n$, la forma cuadrática puede expresarse como

$$f_B(u) = \sum_{i=1}^{n} a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j,$$

donde a_{ij} son números reales, y que matricialmente se escribe

$$f(x) = x^tAx,$$

Siendo $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}\end{pmatrix}$ una matriz simétrica.

Observación 1.2.1 *Si f es una forma cuadrática de $\mathbb{R}^n$, para todo $x \in \mathbb{R}^n$, para todo $\alpha \in \mathbb{R}$, $f(\alpha x) = \alpha^2 f(x)$; $f(0) = 0$.*

Observación 1.2.2 *Dos formas bilineales B_1, B_2 que verifican que, para cada par de $i, j = 1, 2, \ldots, n,\ B_1(u_i, u_j) + B_1(u_j, u_i) = B_2(u_i, u_j) + B_2(u_j, u_i)$, siendo $(u_1, \ldots, u_n)$ una base de $\mathbb{R}^n$, determinan la misma forma cuadrática.*

Por lo tanto, una forma cuadrática queda determinada por infinitas formas bilineales. Sin embargo, solo una de ellas es simétrica.

*Dada una forma cuadrática f de $\mathbb{R}^n$, la forma bilineal **simétrica** cuya forma cuadrática asociada es f está definida por $B(u_i, u_j) = \frac{1}{4}(f(u_i + u_j) - f(u_i - u_j))$, para $i, j \in \{1, 2, \ldots, n\}$ siendo $(u_1, \ldots, u_n)$ cualquier base de $\mathbb{R}^n$.*

Definición 1.2.2 *Dada una forma cuadrática f y una base ν de $\mathbb{R}^n$, definimos la matriz f respecto a ν como la matriz respecto a dicha base de la forma bilineal simétrica B tal que $f(x) = B(x, x)$ para todo $x \in \mathbb{R}^n$.*

1.2.2 Matrices congruentes

A continuación estudiamos cómo se relacionan las expresiones matriciales de las formas bilineales y cuadráticas en diferentes bases.

Sea $f(x) = x^t A x$ una forma cuadrática de $\mathbb{R}^n$, expresada en coordenadas respecto a una determinada base.

Si P es una matriz invertible y hacemos el cambio lineal de coordenadas en $\mathbb{R}^n$ $Py = x$, tenemos que

$$f(x) = (Py)^t A P y = y^t P^t A P y,$$

de donde se deduce que la matriz f respecto a la nueva base es $A' = P^t A P$.

Definición 1.2.3 *Dos matrices A, A' reales y simétricas se dicen **congruentes** entre sí, si existe una matriz P invertible, tal que $A' = P^t A P$.*
Equivalentemente, dos matrices reales y simétricas son congruentes si representan la misma forma cuadrática respecto a diferentes bases de $\mathbb{R}^n$.

Observación 1.2.3 *Matrices congruentes entre sí tienen el mismo rango. Ello motiva la sigguiente definición:*

Definición 1.2.4 *Se define el* ***rango*** *de una forma cuadrática, como el rango de su matriz (simétrica) respecto a cualquier base de* $\mathbb{R}^n$.

Observación 1.2.4 *La simetría de la matriz es imprescindible en la definición anterior. Por ejemplo. la forma cuadrática de* $\mathbb{R}^2$ *definida por* $f(x, y) = x^2+y^2+2xy$, *puede escribirse como*

$$f(x,y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

siendo el rango de f *igual a* 1, *que es el de la matriz simétrica. El rango* 2 *de la segunda matriz no da información sobre la forma cuadrática, por no ser aquella simétrica.*

1.3 Diagonalización de formas cuadráticas

1.3.1 Mediante operaciones elementales

Dada una matriz real y simétrica A, si realizamos una operación elemental sobre sus filas y a continuación la misma operación sobre sus columnas, se conserva la simetría de la matriz.
Ello equivale a la post-y premultiplicación por una matriz elemental y su transpuesta: $E_1^tAE_1$. Reiteramos este proceso hasta obtener una matriz diagonal.

$$E_k^t \dots E_2^t E_1^t A E_1 E_2 \dots E_k = D \text{ es decir } (E_1E_2 \dots E_k)^t A E_1 E_2 \dots E_k = D.$$

Asi pues, tomando $P = E_1E_2 \dots E_k$, es decir, la matriz resultante de realizar sobre la identidad las mismas operaciones de columnas efectuadas a A, tenemos que $P^tAP = D$.

1.3.2 Completando cuadrados

Vamos a ver que toda forma cuadrática no nula de $\mathbb{R}^n$ puede reducirse a una suma de cudrados de formas lineales, más una forma cuadrática "residual" que depende de, a lo sumo $n-1$ variables, sobre la cual se reiterará el procedimiento. Así, en $n-1$ pasos como maáximo, la forma cuadrática original quedará reducida a una suma de cuadrados.

Primer caso: Si $f(x) = \sum_{i=1}^n a_{ii}x_i^2 + 2\sum_{i<j} a_{ij}x_ix_j$, con algún coeficiente diagonal no nulo, por ejemplo a_{11}, agrupamos los términos de la siguiente manera:

$$f(x) = a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n\right)^2 - \sum_{j=2}^n \frac{a_{1j}^2}{a_{11}}x_j^2 - 2\sum_{2\leq i<j} \frac{a_{1i}a_{1j}}{a_{11}}x_ix_j +$$

$$+ \sum_{i=2}^n a_{ii}x_i^2 + \sum_{2\leq i<j} a_{ij}x_ix_j = a_{11}y_1^2 + \sum_{i=2}^n b_{ii}x_i^2 + \sum_{2\leq i<j} b_{ij}x_ix_j,$$

en donde hemos introducido la nueva variable $y_1 = x_1 + \frac{a_{12}}{a_{11}}x_2 + \ldots + \frac{a_{1n}}{a_{11}}x_n$, y los dos últimos sumandos constituyen una forma cuadrática de $\mathbb{R}^{n-1}$ en las variables $x_2, \ldots, x_n$.

Segundo Caso: Si son nulos todos los coeficientes diagonales a_{ii}, existirá algún a_{ij} no nulo, con $i \neq j$; por ejemplo, $a_{12} \neq 0$. Entonces procederemos de la siguiente manera:

$$f(x) = 2\sum_{i\leq j} a_{ij}x_ix_j = 2a_{12}x_1x_2 + 2x_1(a_{13}x_3 + \ldots + a_{1n}x_n) + 2x_2(a_{23}x_3 + \ldots + a_{2n}x_n) +$$

$$+2\sum_{3\leq i<j} a_{ij}x_ix_j = 2a_{12}\left(x_1 + \frac{1}{2a_{12}}(a_{13}x_3 + \ldots + a_{1n}x_n)\right)\left(x_2 + \frac{1}{2a_{12}}(a_{23}x_3 + \ldots + a_{2n}x_n)\right)$$

$$+2\sum_{3\leq i<j} a_{ij}x_ix_j - \frac{1}{2a_{12}}(a_{13}x_3 + \ldots + a_{1n}x_n)(a_{23}x_3 + \ldots + a_{2n}x_n)$$

Hacemos ahora el cambio de variables definido por

$$y_1 + y_2 = x_1 + \frac{1}{2a_{12}}(a_{23}x_3 + \ldots + a_{2n}x_n)$$

$$y_1 - y_2 = x_2 + \frac{1}{2a_{13}}(a_{23}x_3 + \ldots + a_{1n}x_n)$$

y la forma cuadrática puede escribirse de la forma:

$$f(x) = 2a_{12}(y_1 + y_2)(y_1 - y_2) - \frac{1}{2a_{12}}(a_{13}x_3 + \ldots + a_{1n}x_n)(a_{23}x_3 + \ldots + a_{2n}x_n) =$$

$$= 2a_{12}y_1^2 - 2a_{12}y_2^2 + \sum i = 3^n b_{ii}x_i^2 + \sum_{3\leq i<j} b_{ij}x_ix_j,$$

en donde los dos últimos sumandos representan una forma cuadrática de $\mathbb{R}^{n-2}$ en las variables $x_3, \ldots, x_n$.

1.3.3 Mediante diagonalización ortogonal

Como se sabe, dada A real y simétrica, podemos encontrar P ortogonal tal que P^tAP sea diagonal, siendo los elementos diagonales de ésta los autovalores de A, y constituyendo las columnas de P una base ortonormal de $\mathbb{R}^n$ formada por autovectores de A.

Observación 1.3.1 *Existen infinitas matrices diagonales congruentes con una matriz dada. Salvo que hayamos utilizado el tercer método, los elementos de una matriz diagonal congruente con A no tienen por qué ser los autovalores de A.*

Sin embargo, como veremos a continuación, lo que si es invariable en todas las formas diagonales de una misma forma cuadrática son los signos de los coeficientes (y que, por tanto, coinciden con los signos de los autovalores de la matriz).

1.4 Ley de inercias de Sylvester

Teorema 1.4.1 ***(Ley de inercia de Sylvester)*** *Todas las matrices diagonales congruentes con una misma matriz real y simétrica tienen el mismo número de elementos positivos, el mismo número de elementos negativos, y el mismo número de elementos nulos*

Demostración Sean dos matrices diagonales congruentes con una misma matriz A, que expresaremos de la siguiente manera, siendo todos los α_j y β_j positivos:

$$D = diag(\alpha_1, \ldots, \alpha_p, -\alpha_{p+1}, \ldots, -\alpha_r, 0, \ldots, 0);$$

$$D' = diag(\beta_1, \ldots, \beta_q, -\beta_{q+1}, \ldots, -\beta_r, 0, \ldots, 0),$$

donde r representa el rango de A.
Queremos demostrar que $p = q$.
Sean $(u_1, \ldots, u_n)$ y $(v_1, \ldots, v_n)$ las bases en las que la forma cuadrática de A se expresa, respectivamente, mediante D y D'.
Tenemos que

$$f(u_j) = \alpha_j > 0 \text{ para } j = 1, \ldots, p \quad f(u_j) = -\alpha_j < 0 \text{ para } j = p+1, \ldots, r;$$

$$f(v_j) = \beta_j > 0 \text{ para } j = 1, \ldots, q \quad f(v_j) = -\beta_j < 0 \text{ para } j = q+1, \ldots, r;$$

Consideremos los subespacios de $\mathbb{R}^n$

$$L = \mathcal{L}\{u_1, \ldots, u_p\} \text{ y } M = \mathcal{L}\{v_{q+1}, \ldots, v_n\}.$$

Para todo $x \in L\backslash\{0\}$, se verifica que $f(x) > 0$; para todo $x \in M$, se tiene que $f(x) \leq 0$; por tanto, el único elemento común a ambos es el 0, luego $L \cap M = \{0\}$ y, por consiguiente,

$$dim(L \oplus M) = dimL + dimM = p + n - q.$$

Ahora bien, como esta dimensión no puede ser mayor que n, se deduce que $p - q \leq 0$, es decir, $p \leq q$. Intercambiando los papeles de D y D', se deduce que $q \leq p$, y se concluye que $p = q$.

Capítulo 2

Relación entre las expresiones de E

Comencemos estudiando la expresión (Im),

$$E = \{(x, x^*) : x = Pu,\ x^* = Qu,\ u \in \mathbb{R}^r\}$$

con $P, Q \in \mathbb{R}^{n\times r}$ y $\begin{pmatrix} P \\ Q \end{pmatrix}$ inyectiva. Con esta expresión la monotonía de E es equivalente a decir que la matriz $(P^tQ + Q^tP)$ es semidefinida positiva.

Es claro que P y Q en la expresión (Im) no es única, no obstante la siguiente proposición muestra que la inercia de $(P^tQ + Q^tP)$ es independiente de los representantes P y Q que se elijan en E.

Proposición 2.0.1 *Sean $P_i, Q_i \in \mathbb{R}^{n\times r}$ con $\begin{pmatrix} P_i \\ Q_i \end{pmatrix}$ inyectiva ($i = 1, 2$), satisfaciendo*

$$\{(x, x^*) : x = P_1u,\ x^* = Q_1u,\ u \in \mathbb{R}^r\} = \{(x, x^*) : x = P_2u,\ x^* = Q_2u,\ u \in \mathbb{R}^r\}.$$

Entonces

$$\mathrm{In}\,(\mathrm{P}_1^t\mathrm{Q}_1 + \mathrm{Q}_1^t\mathrm{P}_1) = \mathrm{In}\,(\mathrm{P}_2^t\mathrm{Q}_2 + \mathrm{Q}_2^t\mathrm{P}_2)$$

Demostración Sea $\{u_1, u_2, \ldots, u_{2n-p}\}$ una base de $\mathbb{R}^{2n-p}$. Para cada $i = 1, \cdots, 2n-p$, definamos $\begin{pmatrix} P_2 \\ Q_2 \end{pmatrix} v_i = \begin{pmatrix} P_1 \\ Q_1 \end{pmatrix} u_i$. Como $\begin{pmatrix} P_1 \\ Q_1 \end{pmatrix}$ y $\begin{pmatrix} P_2 \\ Q_2 \end{pmatrix}$ son inyectivos, entonces $\{v_1, v_2, \ldots, v_{2n-p}\}$ es una base de $\mathbb{R}^{2n-p}$. Sea $R \in \mathbb{R}^{(2n-p)\times(2n-p)}$ tal que $R(u_i) = v_i$. Entonces

$$\begin{pmatrix} P_1 \\ Q_1 \end{pmatrix} u_i = \begin{pmatrix} P_2 \\ Q_2 \end{pmatrix} Ru_i,\ \text{para todo } i = 1, 2, \cdots, 2n-p.$$

Se deduce que $\begin{pmatrix} P_1 \\ Q_1 \end{pmatrix} = \begin{pmatrix} P_2R \\ Q_2R \end{pmatrix}$ y por lo tanto

$$(P_1^tQ_1 + Q_1^tP_1) = R^t(P_2^tQ_2 + Q_2^tP_2)R.$$

Por la propiedad de inercia (1.0.2), se cumple el resultado. ∎

Gracias a la propiedad anterior sólo consideraremos un par de matrices P y Q, para ello denotemos por $\Pi \in \mathbb{R}^{2n\times 2n}$ a la matriz proyección sobre E

$$\Pi \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} x \\ x^* \end{pmatrix}$$

si y solo si

$$\begin{pmatrix} z \\ w \end{pmatrix} - \begin{pmatrix} x \\ x^* \end{pmatrix} \in \left[\mathrm{Ker}\begin{pmatrix} A & B \end{pmatrix}\right]^{\perp} = \mathrm{Im}\begin{pmatrix} A^t \\ B^t \end{pmatrix}$$

el cual es equivalente a

$$\begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} x \\ x^* \end{pmatrix} + \begin{pmatrix} A^t \\ B^t \end{pmatrix} u \;\text{ para algún } u \in \mathbb{R}^p. \tag{2.1}$$

De donde

$$\begin{pmatrix} A & B \end{pmatrix}\begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} A & B \end{pmatrix}\begin{pmatrix} x \\ x^* \end{pmatrix} + \begin{pmatrix} A & B \end{pmatrix}\begin{pmatrix} A^t \\ B^t \end{pmatrix} u$$

si y solo si

$$(AA^t + BB^t)^{-1}\begin{pmatrix} A & B \end{pmatrix}\begin{pmatrix} z \\ w \end{pmatrix} = u$$

Reemplazando en (2.1) obtenemos

$$\begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} x \\ x^* \end{pmatrix} + \begin{pmatrix} A^t \\ B^t \end{pmatrix}(AA^t + BB^t)^{-1}\begin{pmatrix} A & B \end{pmatrix}\begin{pmatrix} z \\ w \end{pmatrix}$$

de donde

$$\Pi\begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} x \\ x^* \end{pmatrix} = \left[I_{2n} - \begin{pmatrix} A^t \\ B^t \end{pmatrix}(AA^t + BB^t)^{-1}\begin{pmatrix} A & B \end{pmatrix}\right]\begin{pmatrix} z \\ w \end{pmatrix}.$$

Se deduce que

$$\Pi = \left[I_{2n} - \begin{pmatrix} A^t \\ B^t \end{pmatrix} (AA^t + BB^t)^{-1} \begin{pmatrix} A & B \end{pmatrix} \right]. \tag{2.2}$$

La matriz $\Pi \in \mathbb{R}^{2n\times 2n}$ es simétrica con imagen E, pero no es inyectiva. Sin embargo, existe una matriz $K \in \mathbb{R}^{2n\times(2n-p)}$ con $K^tK = I$ tal que ΠK es inyectiva.

Sean P y Q tal que

$$\begin{pmatrix} P \\ Q \end{pmatrix} = \Pi K. \tag{2.3}$$

y denotemos por F a la siguiente matriz

$$F = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix} \tag{2.4}$$

Se cumple la siguiente equivalencia:

Proposición 2.0.2 *E es monótono si y solo si $\Pi F \Pi$ es semidefinida positiva.*

Demostración Es claro que

$$\mathrm{Im}\,(\Pi K) = \mathrm{Im}\,(\Pi) = E.$$

El resultado se deduce de esta expresión. □

Ahora daremos una expresión explícita de $\Pi F \Pi$. Como

$$F = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix} \qquad y \qquad \Pi = \left[I_{2n} - \begin{pmatrix} A^t \\ B^t \end{pmatrix} (AA^t + BB^t)^{-1} \begin{pmatrix} A & B \end{pmatrix} \right]$$

Entonces

$$F\Pi = \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix} - \begin{pmatrix} B^t \\ A^t \end{pmatrix} (AA^t + BB^t)^{-1} \begin{pmatrix} A & B \end{pmatrix}$$

y por lo tanto

$$\begin{aligned} \Pi F \Pi \;&=\; \begin{pmatrix} 0 & I_n \\ I_n & 0 \end{pmatrix} - \begin{pmatrix} B^t \\ A^t \end{pmatrix} (AA^t + BB^t)^{-1} \begin{pmatrix} A & B \end{pmatrix} - \begin{pmatrix} A^t \\ B^t \end{pmatrix} (AA^t + BB^t)^{-1} \begin{pmatrix} B & A \end{pmatrix} \\ &\cdots + \begin{pmatrix} A^t \\ B^t \end{pmatrix} (AA^t + BB^t)^{-1}(AB^t + BA^t)(AA^t + BB^t)^{-1} \begin{pmatrix} A & B \end{pmatrix}. \end{aligned}$$

Proposición 2.0.3 *Se cumple la siguiente relación:*

$$E^\perp \subset \ker(\Pi F \Pi).$$

Demostración Sabemos que $E^\perp = \ker \begin{pmatrix} A & B \end{pmatrix}^\perp = Im \begin{pmatrix} A^t \\ B^t \end{pmatrix}$. Consideremos $v \in G^\perp$, entonces $v = \begin{pmatrix} A^t u \\ B^t u \end{pmatrix}$ para algún $u \in \mathbb{R}^p$ y por lo tanto

$$\Pi v = \Pi \begin{pmatrix} A^t u \\ B^t u \end{pmatrix} = \left[I_{2n} - \begin{pmatrix} A^t \\ B^t \end{pmatrix} (AA^t + BB^t)^{-1} \begin{pmatrix} A & B \end{pmatrix} \right] \begin{pmatrix} A^t u \\ B^t u \end{pmatrix}$$

Se deduce que $v \in \ker(\Pi F \Pi)$. ■

Proposición 2.0.4 *Sea v un autovector de $\Pi F \Pi$ correspondiente a un autovalor distinto de cero, entonces $v \in E$.*

Demostración Sean $v \neq 0$ y $\lambda \neq 0$ tal que $\Pi F \Pi v = \lambda v$. Como $\mathrm{Im}\,(\Pi) = E$, entonces $\lambda v \in E$, de donde $v \in E$. ■

Definición 2.0.1 (Diagonal de E) *Llamaremos diagonal de E, y lo denotaremos por D_E, al siguiente subespacio vectorial de E*

$$D_E = \{(x, x^*) \in E : x^* = x\}.$$

Independientemente de la expresión de Π dada anteriormente, este cumple con las siguientes propiedades conocidas

Lema 2.0.1 *Se tiene*

$$|\Pi(x)| \leq |x| \quad \textit{para todo } x \in \mathbb{R}^{2n}$$

y

$$|\Pi(x)| = |x| \quad \textit{si y solo si } x \in E.$$

Demostración i) Sea $x \in \mathbb{R}^{2n}$, entonces existen $u \in E$ y $w \in E^\perp$ tal que $x = u + w$. Se deduce que $\Pi(x) = u$ y como $|x|^2 = |u + w|^2 = \langle u + w, u + w \rangle = |u|^2 + |w|^2 \geq |u|^2 = |\Pi(x)|^2$, tenemos

$$|\Pi(x)| \leq |x| \quad \text{para todo } x \in \mathbb{R}^{2n}.$$

ii) Si $x \in E$ entonces $\Pi(x) = x$ y $|\Pi(x)| = |x|$. Recíprocamente, supongamos que $|\Pi(x)| = |x|$. Debido a que $x = u + w$ con $u \in E$ y $w \in E^{\perp}$, $|x|^2 = |u|^2 + |w|^2$ y $\Pi(x) = u$, concluimos que $w = 0$ y por lo tanto $x \in E$. ■

Corolario 2.0.1 *Sea λ un autovalor de la matriz $\Pi F \Pi$, entonces $|\lambda| \leq 1$.*

Demostración Se deduce inmediatamente del lema anterior y del hecho que F es una isometría. ■

Proposición 2.0.5 *Sea $(x, x^*) \in E \setminus \{0\}$, entonces (x, x^*) es un autovector de $\Pi F \Pi$ correspondiente al autovalor 1 si y solo si $(x, x^*) \in D_E$.*

Demostración Asumamos en primer lugar que (x, x^*) es un autovector de $\Pi F \Pi$ correspondiente al autovalor 1, entonces

$$\begin{pmatrix} x \\ x^* \end{pmatrix} = \Pi F \Pi \begin{pmatrix} x \\ x^* \end{pmatrix} = \Pi F \begin{pmatrix} x \\ x^* \end{pmatrix} = \Pi \begin{pmatrix} x^* \\ x \end{pmatrix}.$$

Debido a que $\left\| \Pi \begin{pmatrix} x^* \\ x \end{pmatrix} \right\| = \left\| \begin{pmatrix} x \\ x^* \end{pmatrix} \right\|$, entonces $(x^*, x) \in E$ y $\Pi \begin{pmatrix} x^* \\ x \end{pmatrix} = \begin{pmatrix} x^* \\ x \end{pmatrix}$.
Se deduce que $\begin{pmatrix} x \\ x^* \end{pmatrix} = \begin{pmatrix} x^* \\ x \end{pmatrix}$, es decir $(x, x^*) \in D_E$.

Ahora asumamos que $(x, x^*) \in D_E$, entonces $x = x^*$ y por lo tanto

$$\Pi F \Pi \begin{pmatrix} x \\ x \end{pmatrix} = \Pi F \begin{pmatrix} x \\ x \end{pmatrix} = \Pi \begin{pmatrix} x \\ x \end{pmatrix} = \begin{pmatrix} x \\ x \end{pmatrix}$$

Se concluye el resultado. ■

Proposición 2.0.6 *Si E es monótono, entonces $(B-A)$ es inyectivo.*

Demostración Sea $x \in \mathbb{R}^n$ tal que $(B-A)x = 0$, entonces $(x, -x) \in E$ y por lo tanto $-\|x\|^2 = \langle x - 0, -x - 0\rangle \geq 0$. De donde $x = 0$ y así $(B-A)$ es inyectiva. ■

Proposición 2.0.7 *Existe una base ortonormal $\mathcal{B}$ de $\mathbb{R}^{2n}$ formado por autovectores de $\Pi F \Pi$ tal que $\mathcal{B} \subset E \cup E^{\perp}$.*

Demostración Sea $\mathcal{B}_1 = \{v_1, \ldots, v_l, v_{l+1}, \ldots, v_{2n}\}$ una base ortonormal de $\mathbb{R}^{2n}$ formado por autovectores de $\Pi F \Pi$ donde v_j (para $j = 1, \cdots, l$) son los autovectores correspondientes a autovalores distintos de cero y los otros (para $j = l+1, \cdots, 2n$) corresponden al autovalor cero.

Por la Proposición 2.0.4 tenemos que $v_j \in E$ para todo $1 \leq j \leq l$, además

$$(\langle v_1, \ldots, v_l\rangle)^{\perp} = \langle v_{l+1}, \ldots, v_{2n}\rangle = Ker(\Pi F \Pi).$$

Completando $\{v_1, \ldots, v_l\}$ ortonormalmente a una base de E obtenemos

$$\mathcal{B}_2 = \{v_1, \ldots, v_l, \hat{v}_{l+1}, \ldots, \hat{v}_{2n-p}\}$$

de donde $\quad \langle \hat{v}_{l+1}, \ldots, \hat{v}_{2n-p}\rangle \subset (\langle v_1, \ldots, v_l\rangle)^{\perp} = Ker(\Pi F \Pi).$

Entonces $\hat{v}_j \in E$ $(l+1 \leq j \leq 2n-p)$ son autovectores de $\Pi F \Pi$ correspondientes al autovalor cero.

Seguimos completando ortonormalmente $\mathcal{B}_2$ a una base de $\mathbb{R}^{2n}$ y obtenemos

$$\mathcal{B} = \{v_1, \ldots, v_l, \hat{v}_{l+1}, \ldots, \hat{v}_{2n-p}, \hat{v}_{2n-p+1}, \ldots, \hat{v}_{2n}\}$$

Se deduce que $\hat{v}_j \in G^{\perp}$ para todo $j = 2n-p, \cdots, 2n$. Por lo tanto $\mathcal{B}$ es una base ortonormal de $\mathbb{R}^{2n}$ formado por autovectores de $\Pi F \Pi$ satisfaciendo $\mathcal{B} \subset E \cup E^{\perp}$. ■

Proposición 2.0.8 *Sea $v = (x, x^*) \in E \backslash E_G$ un autovector de $\Pi F \Pi$ correspondiente al autovalor λ. Entonces $y = Ax^* + Bx$ es un autovector de $(AB^t + BA^t)(AA^t + BB^t)^{-1}$ correspondiente al autovalor $-\lambda$.*

Demostración De $\Pi F \Pi v = \lambda v$ obtenemos

$$\begin{pmatrix} x^* \\ x \end{pmatrix} - \begin{pmatrix} A^t \\ B^t \end{pmatrix} (AA^t + BB^t)^{-1}(Bx + Ax^*) = \begin{pmatrix} \lambda x \\ \lambda x^* \end{pmatrix}$$

de donde

$$-\begin{pmatrix} A^t \\ B^t \end{pmatrix}(AA^t+BB^t)^{-1}(Ax^*+Bx) = \begin{pmatrix} \lambda x - x^* \\ \lambda x^* - x \end{pmatrix}.$$

Multiplicando por $\begin{pmatrix} B & A \end{pmatrix}$ a la izquierda obtenemos

$$-\begin{pmatrix} B & A \end{pmatrix}\begin{pmatrix} A^t \\ B^t \end{pmatrix}(AA^t+BB^t)^{-1}(Ax^*+Bx) = B(\lambda x - x^*) + A(\lambda x^* - x)$$

de donde

$$-(AB^t+BA^t)(AA^t+BB^t)^{-1}(Ax^*+Bx) = \lambda(Ax^*+Bx) - (Ax+Bx^*).$$

Sea $y = Ax^* + Bx$. Como (x, x^*) esta en E, se tiene

$$(AB^t+BA^t)(AA^t+BB^t)^{-1}y = -\lambda y.$$

Para terminar la prueba sólo falta demostrar que $y \neq 0$. Asumamos por contradicción que $y = 0$, entonces

$$0 = y = Ax^* + Bx \quad \text{ y } \quad Ax + Bx^* = 0.$$

Sumando estas dos ecuaciones tenemos

$$0 = (B - A)(x - x^*)$$

y por lo tanto, por ser $(B-A)$ inyectiva, obtenemos $x = x^*$ lo cual es una contradicción debido a que $(x, x^*) \notin D_E$. ■

La siguiente Proposición muestra la relación entre las inercias de las matrices $AB^t + BA^t$ y $(AB^t+BA^t)(AA^t+BB^t)^{-1}$.

Proposición 2.0.9 $\mathrm{In}\,(\mathrm{AB^t+BA^t}) = \mathrm{In}\,[(\mathrm{AB^t+BA^t})(\mathrm{AA^t+BB^t})^{-1}]$

Demostración Como $(AA^t+BB^t)^{-1}$ es simétrica y definida positiva entonces existe una matriz inversible R tal que

$$(AA^t+BB^t)^{-1} = R^tR$$

entonces

$$(AB^t+BA^t)(R^tR) = (AB^t+BA^t)(AA^t+BB^t)^{-1}$$

$$(AB^t + BA^t)(R^t) = (AB^t + BA^t)(AA^t + BB^t)^{-1}R^{-1}$$

$$R(AB^t + BA^t)R^t = R(AB^t + BA^t)(AA^t + BB^t)^{-1}R^{-1}$$

De donde se deduce el resultado. ■

Proposición 2.0.10 *u es un autovector de $(AB^t + BA^t)(AA^t + BB^t)^{-1}$ asociado al autovalor 1 si y solo si $u \in (AA^t + BB^t)(\ker(B - A)^t) \setminus \{0\}$.*

Demostración Sea $u \neq 0$ tal que

$$(AB^t + BA^t)(AA^t + BB^t)^{-1}u = u.$$

Entonces

$$(AB^t + BA^t)(AA^t + BB^t)^{-1}u - (AA^t + BB^t)(AA^t + BB^t)^{-1}u = 0$$

de donde

$$-(B - A)(B - A)^t(AA^t + BB^t)^{-1}u = 0.$$

Sea $w = (AA^t + BB^t)^{-1}u$, entonces $w \neq 0$ y $-(B - A)(B - A)^t w = 0$. Por lo tanto

$$-\langle (B - A)(B - A)^t w, w\rangle = 0$$

de donde $w \in \mathrm{Ker\,(B - A)^t} \setminus \{0\}$.

Recíprocamente, asumamos ahora que $u = (AA^t + BB^t)w$ con $w \in \ker(B - A)^t \setminus \{0\}$, entonces

$$A^t w = B^t w$$

y por lo tanto

$$AA^t w = AB^t w \quad \text{y} \quad BB^t w = BA^t w.$$

Sumando las dos últimas ecuaciones, obtenemos

$$(AA^t + BB^t)w = (AB^t + BA^t)w$$

de donde

$$u = (AB^t + BA^t)(AA^t + BB^t)^{-1}u.$$

Por lo tanto la demostración de la proposición. ■

Observación 2.0.1 *Como* $(B-A)^t \in \mathbb{R}^{n\times p}$ *es sobreyectiva, entonces*

$$\mathrm{Ker}\,(B-A)^t + \mathrm{Im\,(B-A)^t = p.}$$

de donde $\mathrm{Ker}\,(B-A)^t = p-n$.

Corolario 2.0.2 *Se cumple*

$$(AA^t+BB^t)(ker(B-A)^t) = V_1[(AB^t+BA^t)(AA^t+BB^t)^{-1}]$$

y

$$\dim(V_1[(AB^t+BA^t)(AA^t+BB^t)^{-1}]) = p-n$$

donde $V_\lambda(C)$ *es el espacio generado por los autovectores de la matriz* C *asociados al autovalor* λ.

Demostración Inmediato de la proposición anterior. ■

Proposición 2.0.11 u *es un autovector de* $(AB^t+BA^t)(AA^t+BB^t)^{-1}$ *asociado al autovalor* -1 *si y solo si* $u \in (AA^t+BB^t)(\ker(A+B)^t)\setminus\{0\}$.

Demostración Análogo a la demostración de la Proposición 2.0.11. ■

Observación 2.0.2 *Como* $(A+B)\in\mathbb{R}^{p\times n}$ *y*

$$\dim(\ker(A+B)) = \dim\{u\in\mathbb{R}^n | Au+Bu=0\} =$$

$$= \dim\{(u,u)\in\mathbb{R}^{2n} | Au+Bu=0\} = \dim D_E =: k,$$

entonces del hecho que

$$\dim(\mathrm{Ker}\,(A+B)) + \dim(\mathrm{Im\,(A+B)) = n}$$

se tiene que $\dim(\mathrm{Im}\,(A+B)) = n-k$ *y también* $\dim(\mathrm{Im}\,(A+B)^t) = n-k$.

Luego

$$\dim(\mathrm{Ker}\,(A+B)^t) + \dim(\mathrm{Im}\,(A+B)^t) = p$$

y por lo tanto

$$\dim(\mathrm{Ker}\,(A+B)^t) = p-n+k.$$

Corolario 2.0.3 *Se cumple*

$$(AA^t + BB^t)(ker(A+B)^t) = V_{-1}[(AB^t + BA^t)(AA^t + BB^t)^{-1}]$$

y

$$\dim(V_{-1}[(AB^t + BA^t)(AA^t + BB^t)^{-1}]) = p - n + k.$$

Demostración Inmediato de la Proposición anterior. ■

Proposición 2.0.12

$$E \cap \operatorname{Ker}(\Pi F \Pi) = \begin{pmatrix} B^t \\ A^t \end{pmatrix} (\ker(AB^t + BA^t))$$

Demostración Asumamos en primer lugar que $(x, x^*) \in E \cap \ker(\Pi F \Pi)$.

$$\Pi F \Pi \begin{pmatrix} x \\ x^* \end{pmatrix} = 0 \Rightarrow \Pi F \begin{pmatrix} x \\ x^* \end{pmatrix} = 0 \;\Rightarrow\; \Pi \begin{pmatrix} x^* \\ x \end{pmatrix} = 0$$

$$\begin{pmatrix} x^* \\ x \end{pmatrix} \in E^\perp = (\ker \begin{pmatrix} A & B \end{pmatrix})^\perp = \operatorname{Im} \begin{pmatrix} A^t \\ B^t \end{pmatrix}.$$

Por lo tanto existe $u \in \mathbb{R}^p$ tal que

$$\begin{pmatrix} x \\ x^* \end{pmatrix} = \begin{pmatrix} B^t \\ A^t \end{pmatrix} u.$$

Como $Ax + Bx^* = 0$, entonces $AB^t u + BA^t u = 0$, es decir $u \in \operatorname{Ker}(AB^t + BA^t)$.

Recíprocamente, asumamos que $\begin{pmatrix} x \\ x^* \end{pmatrix} \in \begin{pmatrix} B^t \\ A^t \end{pmatrix} (\operatorname{Ker}(AB^t + BA^t))$ entonces existe $u \in \operatorname{Ker}(AB^t + BA^t)$ tal que $x = B^t u$ y $x^* = A^t u$. Por lo tanto

$$Ax + Bx^* = AB^t u + BA^t u = (AB^t + BA^t)u = 0,$$

de donde $(x, x^*) \in E$.

Por otro lado

$$\Pi F \Pi \begin{pmatrix} x \\ x^* \end{pmatrix} = \Pi F \Pi \begin{pmatrix} B^t u \\ A^t u \end{pmatrix} = \Pi F \begin{pmatrix} B^t u \\ A^t u \end{pmatrix} = \Pi \begin{pmatrix} A^t u \\ B^t u \end{pmatrix} = 0$$

donde la segunda y la última igualdad se cumplen porque $(B^t u, A^t u) \in E$ e $\operatorname{Im} \begin{pmatrix} A^t \\ B^t \end{pmatrix} = E^\perp$, con lo cual queda demostrado esta Proposición. ■

Corolario 2.0.4 *Se cumple que* $\dim(\ker(\Pi F\Pi)) = p + \dim(\ker(AB^t + BA^t))$.

Demostración De la Proposición 2.0.12 tenemos $\dim(E \cap \ker(\Pi F\Pi)) = \dim(\ker(AB^t + BA^t))$. El resultado se deduce del hecho que $\dim(E^\perp) = p$ y $E^\perp \subset \ker(\Pi F\Pi)$. ■

Proposición 2.0.13 *Asumamos que E es monótono, entonces*

1. $In(\Pi F\Pi) = (l, 0, 2n - l)$;
2. $\mathbb{R}^p = (AA^t + BB^t)(\ker(B - A)^t) \oplus (AA^t + BB^t)(\ker(A + B)^t) \oplus W$;
3. $In(AB^t + BA^t) = (p - n, p - n + l, 2n - p - l)$.

donde W será definido en la demostración.

Demostración Sabemos de la Proposición 2.0.7 que existe una base ortonormal $\mathcal{B}$ de $\mathbb{R}^{2n}$ formado por autovectores de $\Pi F\Pi$ tal que

$$\mathcal{B} = \{v_1, \ldots, v_l, v_{l+1}, \ldots, v_{2n-p}, v_{2n-p+1}, \ldots, v_{2n}\} \subset (E \cup E^\perp)$$

donde los l primeros autovectores corresponden a los autovalores distintos de cero, estos son positivos debido a que E es monótono; y los $2n - l$ vectores siguientes son autovectores correspondientes al autovalor cero. De este se deduce que

$$\mathrm{In}\,(\Pi \mathrm{F} \Pi) = (\mathrm{l}, 0, 2\mathrm{n} - \mathrm{l}).$$

Sea k la dimensión de D_E. Por la Proposición 2.0.5 los únicos autovectores en E asociados al autovalor 1 son los que pertenecen a D_E. Cambiando de índice si fuese necesario, tenemos

$$\mathcal{B} = \{v_1, \ldots, v_k, v_{k+1}, \ldots, v_l, v_{l+1}, \ldots, v_{2n-p}, v_{2n-p+1}, \ldots, v_{2n}\}$$

donde los k primeros autovectores forman una base de D_E. Sea $V = \langle v_{k+1}, \ldots, v_l\rangle$ y $W = T(V)$ con $T = \begin{pmatrix} B & A \end{pmatrix}$. Se deduce que

$$\mathbb{R}^p = (AA^t + BB^t)(Ker(B - A)^t) \oplus (AA^t + BB^t)(Ker(A + B)^t) \oplus W.$$

En efecto, por los Corolarios 2.0.2 y 2.0.3, los dos primeros conjuntos son respectivamente los espacios generados por los autovectores de $(AB^t + BA^t)(AA^t +$

$BB^t)^{-1}$ asociados a los autovalores 1 y -1, y W, gracias a la Proposición 2.0.8 y del hecho que el autoespacio $V \subset E$ no contiene autovectores correspondientes al autovalor 1, es el autoespacio generado por autovectores de $(AB^t+BA^t)(AA^t+BB^t)^{-1}$ asociados a autovalores negativos distintos de -1. Por lo tanto los tres espacios tienen al $\{0\}$ como única intersección, y como

$$\dim[(AA^t + BB^t)(Ker(B - A)^t)] + \dim[(AA^t + BB^t)(\ker(A + B)^t)] + \dim(W)$$
$$= (p - n) + (p - n + k) + (2n - p - k) = p,$$

se concluye el segundo resultado.

Como el único autovalor positivo de la matriz $(AB^t + BA^t)(AA^t + BB^t)^{-1}$ es 1, entonces posee $p-n$ autovalores positivos contando multiplicidad, similarmente posee $p - n + k$ autovalores negativos -1 más los que vienen de V via T, es decir existen $(p-n+k)+(l-k) = p-n+l$ autovalores negativos contando multiplicidad y por el Corolario 2.0.4 concluimos que existen $(2n - l) - p$ autovalores iguales a 0 contando multiplicidad. Entonces $\mathrm{In}\left((\mathrm{AB^t}+\mathrm{BA^t})(\mathrm{AA^t}+\mathrm{BB^t})^{-1}\right) = (\mathrm{p-n, p-n+l, 2n-l-p})$ y gracias a la Proposición 2.0.9 se concluye el tercer resultado. ∎

Proposición 2.0.14 *Asumamos que E es monótono, entonces*

$$\mathrm{In}\,(\mathrm{P^tQ + Q^tP}) = (\mathrm{l, 0, 2n - p - l})$$

donde P y Q son como en (2.3).

Demostración Sea $\mathcal{B} = \{v_1, \ldots, v_{2n-p}, v_{2n-p+1}, \ldots, v_{2n}\} \subset E \cup E^{\perp}$ una base ortonormal de $\mathbb{R}^{2n}$ de autovectores de $\Pi F \Pi$ como en la Proposición 2.0.7 donde los $(2n - p)$ primeros vectores están en E y los p restantes se encuentran en $E^{\perp}$.

Sea $K : \mathbb{R}^{2n-p} \to \mathbb{R}^{2n}$ la transformación lineal tal que

$$K(e_i) = v_i \quad \text{para todo } i = 1, \cdots, 2n - p$$

donde $\{e_1, \ldots, e_{2n-p}\}$ es la base canónica de $\mathbb{R}^{2n-p}$. Como

$$P^tQ + Q^tP = \begin{pmatrix} P \\ Q \end{pmatrix}^t F \begin{pmatrix} P \\ Q \end{pmatrix} = K^t\Pi\ F\ \Pi K$$

y del hecho que

$$K^t\Pi F\Pi K(e_i) = K^t\Pi F\Pi v_i = K^t\lambda_i v_i = \lambda_i K^t K(e_i) = \lambda_i e_i$$

tenemos que $\mathrm{In}\,(\mathrm{P^tQ + Q^tP}) = (\mathrm{l, 0, 2n - p - l})$. ∎

Capítulo 3

Resultados Numéricos

Sean

$$A = \begin{pmatrix} 1 & -7 & 0 \\ -1 & 2 & 0 \\ -3 & 4 & 1 \\ 1 & 0 & 1 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} -4 & 1 & -2 \\ 1 & -1 & -8 \\ 2 & 1 & 1 \\ -2 & 0 & 1 \end{pmatrix}$$

matrices de $\mathbb{R}^{4\times 3}$ entonces $p = 4$ y $n = 3$

$$(A\ B) = \begin{pmatrix} 1 & -7 & 0 & -4 & 1 & -2 \\ -1 & 2 & 0 & 1 & -1 & -8 \\ -3 & 4 & 1 & 2 & 1 & 1 \\ 1 & 0 & 1 & -2 & 0 & 1 \end{pmatrix}$$

$E = \mathrm{Ker}\,(\mathrm{A\ B})$ es un espacio vectorial de dimensión $2n - p = 2$

$$AB^t + BA^t = \begin{pmatrix} -22 & 14 & 9 & 8 \\ 14 & -6 & -15 & -5 \\ 9 & -15 & -2 & 10 \\ -8 & -5 & 10 & -2 \end{pmatrix}$$

Calculando los autovalores de $AB^t + BA^t$ se obtiene

$$\lambda_1 = -38.7673, \lambda_2 = -7.9948, \lambda_3 = -3.4406, \lambda_4 = 18.2028$$

Es decir

$$\mathrm{In}\,(\mathrm{AB^t + BA^t}) = (1, 3, 0)$$

y como $\mu_+(AB^t + BA^t) = p - n = 1$ entonces tenemos que E es un subespacio lineal monótono y si E es la imagen de alguna matriz $\begin{pmatrix} P \\ Q \end{pmatrix}$ vimos que $\mu_-(P^tQ + Q^tP)$ es cero, pero antes no se tenía información sobre la relación del resto de elementos de la inercia, sin embargo por lo expuesto en este trabajo se completa la relación entre las inercias y vamos a aprovechar este ejemplo para comprobar cada resultado visto.
La matriz proyección

$$\Pi = I_{2n} - \begin{pmatrix} A^T \\ B^T \end{pmatrix} (AA^T + BB^T)^{-1} \begin{pmatrix} A & B \end{pmatrix}$$

$$\Pi = \begin{pmatrix} 0.2886269 & 0.0128178 & 0.1264848 & 0.1776969 & 0.3924491 & -0.0597180 \\ 0.0128178 & 0.0866834 & -0.2189779 & -0.1084791 & 0.1384538 & -0.0107980 \\ 0.1264848 & -0.2189779 & 0.6411975 & 0.3813788 & -0.1436644 & -0.0049247 \\ 0.1776969 & -0.1084791 & 0.3813788 & 0.2666588 & 0.0780686 & -0.0257581 \\ 0.3924491 & 0.1384538 & -0.1436644 & 0.0780686 & 0.7037069 & -0.0926475 \\ -0.0597180 & -0.0107980 & -0.0049247 & -0.0257581 & -0.0926475 & 0.0131264 \end{pmatrix}$$

La matriz $\Pi F \Pi$ seria

$$\Pi F \Pi = \begin{pmatrix} 0.0975300 & 0.0184723 & 0.0058591 & 0.0409361 & 0.1524866 & -0.0215170 \\ 0.0184723 & 0.0259514 & -0.0574494 & -0.0225882 & 0.0604363 & -0.0061993 \\ 0.0058591 & -0.0574494 & 0.1530805 & 0.0815932 & -0.0731387 & 0.0042467 \\ 0.0409361 & -0.0225882 & 0.0815932 & 0.0581841 & 0.0213607 & -0.0061611 \\ 0.1524866 & 0.0604363 & -0.0731387 & 0.0213607 & 0.2827580 & -0.0366264 \\ -0.0215170 & -0.0061993 & 0.0042467 & -0.0061611 & -0.0366264 & 0.0049480 \end{pmatrix}$$

Como E es un subespacio vectorial de $\mathbb{R}^6$ de dimensión 2 entonces $E^\perp$ es un subespacio vectorial de dimensión 4 entonces los autovalores de $\Pi F \Pi$ son $\lambda_1 = 0$ con multiplicidad 4, $\lambda_2 = 0.40481$ y $\lambda_3 = 0.21764$. Asi mismo la matriz $(AB^t + BA^t)(AA^t + BB^t)^{-1}$ es

$$(AB^t + BA^t)(AA^t + BB^t)^{-1} = \begin{pmatrix} -0.482078 & 0.055642 & -0.4703686 & -0.909331 \\ -0.152356 & -0.301440 & -0.881740 & -1.665435 \\ 0.253283 & 0.015827 & 0.494882 & 1.553647 \\ 0.232926 & -0.141082 & 0.569134 & -0.333816 \end{pmatrix}$$

Los autovalores de $(AB^t + BA^t)(AA^t + BB^t)^{-1}$ son $\overline{\lambda}_1 = 1$, $\overline{\lambda}_2 = -1$, $\overline{\lambda}_3 = -\lambda_2 = -0.40481$ y $\overline{\lambda}_4 = -\lambda_3 = -0.21764$.

Como podemos observar, algunos de los autovalores de $(AB^t + BA^t)(AA^t + BB^t)^{-1}$ son los negativos de los autovalores de la matriz $\Pi F \Pi$ y el resto de autovalores de $(AB^t + BA^t)(AA^t + BB^t)^{-1}$ son -1 y 1, es decir se cumple la Proposición 2.0.8. Sea

$$K = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{entonces} \qquad K^t K = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

La matriz ΠK es inyectiva y tiene como imagen a E, entonces:

$$\begin{pmatrix} P \\ Q \end{pmatrix} = \Pi K = \begin{pmatrix} 0.2886269 & 0.0128178 \\ 0.0128178 & 0.0866834 \\ 0.1264848 & -0.2189779 \\ 0.1776969 & -0.1084791 \\ 0.3924491 & 0.1384538 \\ -0.0597180 & -0.0107980 \end{pmatrix}$$

$$P^t Q + Q^t P = K^t \Pi F \Pi K = \begin{pmatrix} 0.097530 & 0.018472 \\ 0.018472 & 0.025951 \end{pmatrix}$$

Los autovalores de $P^t Q + Q^t P$ son

$$\mathrm{eig}\,(\mathrm{P^tQ + Q^tP}) = (0.102016, 0.021465)$$

La inercia es

$$\mathrm{In}\,(\mathrm{P^tQ + Q^tP}) = (2, 0, 0)$$

No es la única manera de obtener una matriz inyectiva $\begin{pmatrix} P \\ Q \end{pmatrix}$ cuya imagen es E. Tomemos la matriz escalonada reducida por fila de la matriz $\begin{pmatrix} A & B \end{pmatrix}$.

$$rref(\begin{pmatrix} A & B \end{pmatrix}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 7/12 & 26/3 \\ 0 & 1 & 0 & 0 & -5/4 & -8 \\ 0 & 0 & 1 & 0 & 43/12 & 77/3 \\ 0 & 0 & 0 & 1 & 25/12 & 50/3 \end{pmatrix}$$

Entonces

$$\begin{pmatrix} P_1 \\ Q_1 \end{pmatrix} = \begin{pmatrix} -7/12 & -26/3 \\ 5/4 & 8 \\ -43/12 & -77/3 \\ -25/12 & -50/3 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$P_1^t Q_1 + Q_1^t P_1 = \begin{pmatrix} 4.9306 & 32.1944 \\ 32.1944 & 237.5556 \end{pmatrix}$$

Los autovalores de $P_1^t Q_1 + Q_1^t P_1$ son

$$\mathrm{eig}\left(\mathrm{P}_1^{\mathrm{t}} \mathrm{Q}_1 + \mathrm{Q}_1^{\mathrm{t}} \mathrm{P}_1\right) = (0.55718, 241.92893)$$

Y la inercia seria nuevamente

$$\mathrm{In}\left(\mathrm{P}_1^{\mathrm{t}} \mathrm{Q}_1 + \mathrm{Q}_1^{\mathrm{t}} \mathrm{P}_1\right) = (2, 0, 0).$$

Conclusiones

Dado un subespacio lineal monótono E expresado a la vez como núcleo de la transformación lineal sobreyectiva $\begin{pmatrix} A & B \end{pmatrix}$ y como imagen de la tranformación lineal inyectiva $\begin{pmatrix} P \\ Q \end{pmatrix}$, se ha establecido la relación de estas expresiones por medio de las inercias de las partes simétricas de las transformaciones lineales AB^t y P^tQ:

$$\mathrm{In}\,(\mathrm{P^tQ + Q^tP}) = (\mu_-(\mathrm{AB^t + BA^t}) - \mu_+(\mathrm{AB^t + BA^t}),\, 0,\, \mu_0(\mathrm{AB^t + BA^t})).$$

Perspectivas

- Establecer las relaciones de las inercias de las dos expresiones de E cuando este no necesariamente es considerado monótono. De hecho ya se tiene bastante avanzado esta parte, pronto será sometido a una editorial de revista indexada.
- Estudiar la monotonía de subespacios lineales cuando este está definido sobre el octante positivo de $\mathbb{R}^n$.
- La caracterización de la p-monotonía de subespacios lineales por medio de los autovalores de la matrices asociadas.

Bibliografía

[1] J.-P Crouzeix, E. Ocaña Anaya, *Monotone and maximal monotone affine subspaces.* Operations Research Letters 38, 139–142, 2010.

[2] Xianfu Wang, Liangjin Yao, *Maximally Monotone Linear Subspace Extensions of Monotone Subspaces: Explicit Constructions and Characterizations.* Journal of mathematical Programming, Springer Volume 139, Issue 1-2, 327–352.