I

Extremos y formas cuadráticas

I.1. Extremos y formas cuadráticas

En este apéndice se consideran algunas aplicaciones interesantes de la teoría de extremos condicionados al estudio de las formas cuadráticas. Se recomienda comenzar con los ejercicios 9.17, 9.18 que contienen casos particulares de los resultados generales que se exponen aquí. El siguiente teorema es un resultado bien conocido del álgebra lineal del que ofrecemos una demostración alternativa basada en optimizaciones sucesivas de una forma cuadrática sobre la intersección de la esfera unidad con una sucesión decreciente de subespacios vectoriales. La idea clave es que el mayor y el menor autovalor de una matriz simétrica real proporcionan el máximo y el mínimo absoluto, sobre la esfera unidad, de la forma cuadrática asociada a la matriz.

Teorema I.1 *Sea $A = (\alpha_{ij})_{1\le i,j\le n}$ una matriz simétrica, $L : \mathbb{R}^n \to \mathbb{R}^n$ la aplicación lineal asociada y $Q : \mathbb{R}^n \to \mathbb{R}$ la forma cuadrática asociada:*

$$L(\mathbf{x}) = (L_1(\mathbf{x}), L_2(\mathbf{x}), \cdots L_n(\mathbf{x})), \ \textit{donde } L_k(\mathbf{x}) = \sum_{j=1}^{n} \alpha_{kj} x_j$$

$$Q(\mathbf{x}) = \ \langle\ L(\mathbf{x}) \mid \mathbf{x}\ \rangle = \sum_{i,j=1}^{n} \alpha_{ij} x_1 x_j$$

Entonces se verifica: Todos los autovalores de A son reales. Si $\mu_1 \ge \mu_2 \ge \cdots \mu_n$ son los autovalores de A se cumple

$$\mu_1 = \text{máx}\{Q(\mathbf{x}) : \|\mathbf{x}\|_2 = 1\}, \quad \mu_n = \text{mín}\{Q(\mathbf{x}) : \|\mathbf{x}\|_2 = 1\}$$

Existe una base ortonormal de $\mathbb{R}^n$ $\{\mathbf{u}_1, \mathbf{u}_2, \cdots \mathbf{u}_n\}$ formada por vectores propios

$$L(\mathbf{u}_j) = \mu_j \mathbf{u}_j, \quad i \le j \le n$$

y respecto a esta base la matriz de Q es diagonal: Si $\mathbf{x} = \sum_{i=1}^{n} t_i \mathbf{u}_i$ entonces

$$Q(\mathbf{x}) = \mu_1 t_1^2 + \mu_2 t_2^2 + \cdots + \mu_n t_n^2$$

Dem: La función continua $Q(\mathbf{x})$ alcanza un máximo y un mínimo absoluto sobre la esfera compacta $S = \{\mathbf{x} \in \mathbb{R} : \|\mathbf{x}\|_2 = 1\}$. Sea $\mathbf{y} \in S$ tal que

$$Q(\mathbf{y}) = \text{máx}\{Q(\mathbf{x}) : \mathbf{x} \in S\}$$

$S = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{g}(\mathbf{x}) = 0\}$ donde $g(\mathbf{x}) = x_1^2 + x_2^2 + \cdots x_n^2 - 1$ con $\nabla g(\mathbf{x}) = 2\mathbf{x} \neq \mathbf{0}$ para cada $\mathbf{x} \in S$ luego $\mathbf{y} \in S$ es un punto estacionario para $Q|_S$, es decir

$$\nabla Q(\mathbf{y}) = \lambda \nabla g(\mathbf{y})$$

para algún $\lambda \in \mathbb{R}$. Como $D_k Q(\mathbf{x}) = 2\sum_{j=1}^n \alpha_{kj} x_j = 2L_k(\mathbf{x})$, $1 \leq k \leq n$, resulta $2L(\mathbf{y}) = 2\lambda\mathbf{y}$, luego $\mathbf{y}$ es un vector propio de L y λ es el autovalor asociado. Nótese que

$$Q(\mathbf{y}) = \langle L(\mathbf{y}) \mid \mathbf{y} \rangle = \lambda \|\mathbf{y}\|^2 = \lambda$$

Empezamos la construcción con $\mathbf{u}_1 = \mathbf{y}$, y $\mu_1 = \lambda$, que cumplen $Q(\mathbf{u}_1) = \mu_1$. Tenemos demostrado así que L tiene un autovalor real. (De hecho hemos detectado el mayor autovalor, pues si ν es otro autovalor, $L(\mathbf{v}) = \nu\mathbf{v}$ para algún $\mathbf{v} \in S$, luego

$$\nu = \langle \nu\mathbf{v} \mid \mathbf{v}\rangle = \langle L(\mathbf{v}) \mid \mathbf{v}\rangle = Q(\mathbf{v}) \leq Q(\mathbf{y}) = \mu_1).$$

Continuamos la construcción considerando

$$S_1 = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1, \quad \langle \mathbf{u}_1 \mid \mathbf{x} \rangle = 0\}$$

y un punto $\mathbf{z} \in S_1$ tal que

$$Q(\mathbf{z}) = \text{máx}\{Q(\mathbf{x}) : \mathbf{x} \in S_1\} \leq Q(\mathbf{y}) = \mu_1$$

$S_1 = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0, g_1(\mathbf{x}) = 0\}$ donde $g_1(\mathbf{x}) = \langle \mathbf{u}_1 \mid \mathbf{x}\rangle$. Como los vectores $\nabla g(\mathbf{z}) = 2\mathbf{z}$, $\nabla g_1(\mathbf{z}) = \mathbf{u}_1$ son independientes (por ser ortogonales) podemos asegurar que $\mathbf{z} \in S_1$ es estacionario para $Q|_{S_1}$, luego existen $\lambda', \lambda'' \in \mathbb{R}$ tales que

$$\nabla Q(\mathbf{z}) = \lambda' \nabla g(\mathbf{z}) + \lambda'' \nabla g_1(\mathbf{z})$$

es decir

$$2L(\mathbf{z}) = 2\lambda' 2\mathbf{z} + \lambda'' \mathbf{u}_1$$

Como $\mathbf{u}_1$ y $\mathbf{z}$ son ortogonales, usando la simetría de A se obtiene que $\mathbf{u}_1$ y $L(\mathbf{z})$ también lo son:

$$\langle L(\mathbf{z}) \mid \mathbf{u}_1\rangle = \langle \mathbf{z} \mid L(\mathbf{u}_1) \rangle = \langle \mathbf{z} \mid \mu_1\mathbf{u}_1 \rangle = 0.$$

Se sigue de esto que $\lambda'' = 0$, que $\mathbf{z}$ es un vector propio y que λ' es el autovalor correspondiente. Si tomamos $\mu_2 = \lambda'$ y $\mathbf{u}_2 = \mathbf{z}$ es claro que

$$\langle \mathbf{u}_1 \mid \mathbf{u}_2 \rangle = 0; \quad \mu_2 = \langle L(\mathbf{u}_2) \mid \mathbf{u}_2 \rangle = Q(\mathbf{u}_2) \leq \mu_1.$$

La construcción continúa ahora considerando el compacto $S_2 \subset S_1$ definido por

$$S_2 = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_2 = 1, \ \langle \mathbf{u}_1 \mid \mathbf{x}\rangle = 0, \quad \langle \mathbf{u}_2 \mid \mathbf{x}\rangle = 0\}$$

y un punto $\mathbf{w} \in S_2$ tal que

$$Q(\mathbf{w}) = \text{máx}\{Q(\mathbf{x}) : \mathbf{x} \in S_2\}$$

Ahora $S_2 = \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) = 0,\ g_1(\mathbf{x}) = 0,\ g_2(\mathbf{x}) = 0\}$ donde $g_2(\mathbf{x}) = \langle\ \mathbf{u}_2 \mid \mathbf{x}\rangle$.

Como los vectores $\nabla g(\mathbf{w}) = 2\mathbf{w}$ y $\nabla g_1(\mathbf{w}) = \mathbf{u}_1$, $\nabla g_2(\mathbf{w}) = \mathbf{u}_2$ son independientes (por ser ortogonales) el punto $\mathbf{w} \in S_2$ es estacionario para $Q|_{S_2}$ y existen $\mu', \mu'', \mu''' \in \mathbb{R}$ tales que

$$\nabla Q(\mathbf{w}) = \mu' \nabla g(\mathbf{w}) + \mu'' \nabla g_1(\mathbf{w}) + \mu''' \nabla g_2(\mathbf{w})$$

es decir

$$2L(\mathbf{w}) = 2\mu' \mathbf{w} + \mu'' \mathbf{u}_1 + \mu''' \mathbf{u}_2$$

Como $\mathbf{w}$ es ortogonal a $\mathbf{u}_1$, y $\mathbf{u}_2$, usando la simetría de A se obtiene que $L(\mathbf{w})$ también lo es:

$$\langle\ L(\mathbf{w}) \mid \mathbf{u}_i\ \rangle = \langle\ \mathbf{w} \mid L(\mathbf{u}_i)\ \rangle = \langle\ \mathbf{w} \mid \mu_i \mathbf{u}_i\ \rangle = 0$$

Se sigue de esto que $\mu'' = \mu''' = 0$, que $\mathbf{w}$ es un vector propio, y que μ' es el autovalor asociado. Si tomamos $\mu_3 = \mu'$ y $\mathbf{u}_3 = \mathbf{w}$ es claro que se cumple

$$\langle\ \mathbf{u}_1 \mid \mathbf{u}_3\rangle = 0; \quad \langle\ \mathbf{u}_2 \mid \mathbf{u}_3\rangle = 0; \quad \mu_3 = \langle\ L(\mathbf{u}_3) \mid \mathbf{u}_3\ \rangle = Q(\mathbf{u}_3) \le \mu_2.$$

La construcción sigue de esta forma hasta que acaba en un número finito de pasos. ■

A la hora de aplicar la condición suficiente de extremo condicionado dada en el apartado b) del teorema 9.11 se plantea el problema de saber cuando la restricción a un subespacio $T \subset \mathbb{R}^n$ de la forma cuadrática $Q(\mathbf{u}) = \sum_{ij=1}^n \alpha_{ij} u_i u_j$, asociada a una matriz simétrica, es definida positiva o definida negativa.

Proposición I.2 *Sea $Q(\mathbf{u}) = \sum_{ij=1}^n \alpha_{ij} u_i u_j$ la forma cuadrática asociada a una matriz simétrica y $T \subset \mathbb{R}^n$ un subespacio vectorial k-dimensional de ecuaciones implícitas*

$$\sum_{j=1}^n \beta_{ij} u_j = 0, \quad 1 \le i \le m.$$

donde $m = n - k$ y los vectores $(\beta_{i1}, \beta_{i2}, \cdots \beta_{in})$, $1 \le i \le m$ son linealmente independientes. Si todas las raíces del polinomio

$$\Delta(\sigma) = \begin{vmatrix} \alpha_{11} - \sigma & \alpha_{12} & \cdots & \alpha_{1n} & \beta_{11} & \cdots & \beta_{m1} \\ \alpha_{21} & \alpha_{22} - \sigma & \cdots & \alpha_{2n} & \beta_{12} & \cdots & \beta_{m2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} - \sigma & \beta_{1n} & \cdots & \beta_{mn} \\ \beta_{11} & \beta_{12} & \cdots & \beta_{1n} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \beta_{2n} & 0 & \cdots & 0 \\ \beta_{m1} & \beta_{m2} & \cdots & \beta_{mn} & 0 & \cdots & 0 \end{vmatrix}$$

son positivas entonces la restricción de Q al subespacio $T \subset \mathbb{R}^n$ es definida positiva.

Dem: Razonando como en el lema 6.14 es claro que $Q(\mathbf{u}) > 0$ para todo $\mathbf{u} \in T \setminus \{\mathbf{0}\}$ si y sólo si el mínimo absoluto de la función continua Q sobre el compacto $S \cap T = \{\mathbf{x} \in T : \|\mathbf{x}\|_2 = 1\}$, que se alcanza en algún $\mathbf{h} \in S \cap T$, es positivo:

$$\text{mín}\{Q(\mathbf{u}) : \mathbf{u} \in T, \|\mathbf{u}\|_2 = 1\} = Q(\mathbf{h}) > 0$$

$S \cap T$ está definido mediante las $m+1$ condiciones de ligadura:

$$\beta_{11}x_1 + \beta_{12}x_2 + \cdots + \beta_{1n}x_n = 0$$

$$\beta_{21}x_1 + \beta_{22}x_2 + \cdots + \beta_{2n}x_n = 0$$

$$\cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots$$

$$\beta_{m1}x_1 + \beta_{m2}x_2 + \cdots + \beta_{mn}x_n = 0$$

$$x_1^2 + x_2^2 + \cdots + x_n^2 - 1 = 0$$

Para cada $\mathbf{x} \in S \cap T$ los gradientes de las condiciones de ligadura

$$(\beta_{11}, \beta_{12}, \cdots \beta_{1n}),\ (\beta_{21}, \beta_{22}, \cdots \beta_{2n}),\ (\beta_{m1}, \beta_{m2}, \cdots \beta_{mn}),\ (2x_1, 2x_2, \cdots 2x_n)$$

son independientes (los m primeros vectores son independientes y el último es ortogonal a todos ellos) luego, en virtud del teorema 9.10, $\mathbf{h}$ es un punto estacionario de Q sobre $S \cap T$, es decir, existen coeficientes $\sigma, \lambda_1, \lambda_2, \cdots \lambda_m$ tales que

$$D_iQ(\mathbf{h}) = 2\sigma h_i + 2\lambda_1\beta_{1i} + \cdots + 2\lambda_m\beta_{mi} = 0, \quad 1 \le i \le n.$$

es decir

$$\sum_{j=1}^{n} \alpha_{ij}h_j = \sigma h_i + \sum_{r=1}^{m} \lambda_r\beta_{ri} = 0, \quad 1 \le i \le n.$$

Multiplicando le ecuación i-ésima por h_i, sumando, y utilizando que las componentes de $\mathbf{h} = (h_1, h_2, \cdots, h_n) \in T \cap S$ satisfacen las ecuaciones

$$\beta_{i1}h_1 + \beta_{i2}h_2 + \cdots + \beta_{in}h_n = 0, \quad h_1^2 + h_2^2 + \cdots + h_n^2 = 1$$

se concluye que

$$Q(\mathbf{h}) = \sum_{ij} \alpha_{ij}h_ih_j = \sigma$$

Hemos demostrado así que si el mínimo absoluto de Q sobre $S \cap T$ se alcanza en $\mathbf{h} \in S \cap T$ entonces el mínimo absoluto $Q(\mathbf{h})$ es el valor del multiplicador σ asociado a la condición de ligadura $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$.

Sabemos que $h_1, h_2, \cdots h_n, -\lambda_1, -\lambda_2, \cdots - \lambda_m$ son soluciones del sistema homogéneo de $n+m$ ecuaciones con $n+m$ incógnitas

$$(\alpha_{11} - \sigma)h_1 + \alpha_{12}h_2 + \cdots + \alpha_{1n}h_n + \beta_{11}\lambda_1 + \beta_{21}\lambda_2 + \cdots + \beta_{m1}\lambda_m = 0$$

$$\alpha_{21}h_1 + (\alpha_{22} - \sigma)h_2 + \cdots + \alpha_{2n}h_n + \beta_{12}\lambda_1 + \beta_{22}\lambda_2 + \cdots + \beta_{m2}\lambda_m = 0$$

$$
\begin{array}{l}
\cdots\cdots\cdots \qquad \cdots\cdots\cdots \qquad \cdots\cdots\cdots \\
\alpha_{n1}h_1 + \alpha_{n2}h_2 + \cdots + (\alpha_{nn} - \sigma)h_n + \beta_{1n}\lambda_1 + \beta_{2n}\lambda_2 + \cdots + \beta_{mn}\lambda_m = 0 \\
\beta_{11}h_1 + \beta_{12}h_2 + \cdots + \beta_{1n}h_n = 0 \\
\cdots\cdots\cdots \qquad \cdots\cdots\cdots \\
\beta_{m1}h_1 + \beta_{m2}h_2 + \cdots + \beta_{mn}h_n = 0
\end{array}
$$

Como este sistema admite soluciones no triviales, su determinante $\Delta(\sigma)$ se anula, es decir σ es solución de la ecuación $\Delta(\sigma) = 0$. Podemos afirmar entonces que si el polinomio $\Delta(\sigma)$ tiene todas sus raíces positivas entonces la forma cuadrática $Q|_T$ es definida positiva. ■