El Problema de Procrusto y Sumas de Kronecker. ¹

María Gabriela Eberle*

* Departamento de Matemática. Universidad Nacional del Sur.

Resumen

Se llama problema de Procrusto a la clase de todos los problemas de la forma

$$\begin{cases}
\min_{s.a} & ||AX - B||_{\mathbf{F}}^2 \\
X \in \mathcal{D}
\end{cases}$$
(1)

dónde $A, B \in \mathbb{R}^{m \times n}, m > n, \mathcal{D} \subset \mathbb{R}^{n \times n}$. Este problema ha sido resuelto para distintas estructuras de matrices en \mathcal{D} , simétricas [4], simétricas y definidas positivas [3], persimétricas [1], Toeplitz [2]. Van Loan y Pitsianis [7] han resuelto una variante sobre este empleando productos de Kronecker en lugar de producto habitual de matrices. En este trabajo nos proponemos resolver un problema de Procrusto en particular, consiste en hallar una matriz $Z \in \mathbb{R}^{n \times n}$ que minimice la norma $\|K - (Z^T \oplus Z)\|_F^2$ para $K \in \mathbb{R}^{n^2 \times n^2}$ matriz simétrica dada. Una vez presentado el problema se caracterizarán las soluciones del mismo.

Keywords: Problema de Procrusto, Sumas de Kronecker.

1. Introducción

Nos proponemos resolver el siguiente problema de Procrusto

$$\begin{cases}
\min & ||K - (Z^T \oplus Z)||_F^2 \\
Z \in \mathcal{D}
\end{cases}$$
(2)

 $^{^1\}mathrm{Este}$ trabajo es realizado con el apoyo del Departamento de Matemática de la UNS geberle@uns.edu.ar.

Se quiere hallar $Z\in \mathbb{R}^{n\times n}$ que minimice la norma para una matriz simétrica dada $K\in \mathbb{R}^{n^2\times n^2}$.

Definición 1 (Producto de Kronecker) [5] Dadas $X \in \mathbb{R}^{m \times n}$ e $Y \in \mathbb{R}^{p \times q}$ se define el producto de Kronecker de X por Y como $X \otimes Y = (x_{ij}Y) \in \mathbb{R}^{mp \times nq}$

El producto de Kronecker no es conmutativo, pero existe una matriz P tal que $P^T = P^{-1}$ [6] que satisface

$$(X \otimes Y) = P^T(Y \otimes X)P.$$

Ello, junto a la invariancia de la norma de Frobenius para el producto por matrices ortogonales, serán herramientas fundamentales para caracterizar las soluciones del problema (2).

Definición 2 (Suma de Kronecker) [5] Dadas $X \in \mathbb{R}^{m \times m}$ e $Y \in \mathbb{R}^{n \times n}$ se define la suma de Kronecker de X e Y como

$$X \oplus Y = (X \otimes I_n) + (I_m \otimes Y).$$

Finalmente, sea vec(X), transformación que aplicada a una matriz apila sus columnas, de modo que

$$si \ X \in \mathbb{R}^{n \times n} \ entonces \ vec(X) = x \in \mathbb{R}^{n^2}.$$

2. Caracterización de las soluciones.

Interesa resolver el siguiente problema de optimización

$$\begin{cases}
\min \|K - (Z^T \oplus Z)\|_{\mathbf{F}}^2, \\
\end{cases} (3)$$

donde $K \in \mathbb{R}^{n^2 \times n^2}$, $Z \in \mathbb{R}^{n \times n}$. Escribiremos la función objetivo en términos de los elementos z_{ij} de la matriz Z. Para ello los ordenaremos en la misma

forma en que resultan ordenados al vectorizar la matriz Z, como componentes del vector vec(Z).

$$\begin{split} \|K - (Z^T \oplus Z)\|_{\mathbf{F}}^2 &= \|K - ((Z^T \otimes I) + (I \otimes Z))\|_{\mathbf{F}}^2 \\ &= \left\| \left(\frac{1}{2} K - (Z^T \otimes I) \right) + \left(\frac{1}{2} K - (I \otimes Z) \right) \right\|_{\mathbf{F}}^2 \\ &= \left\| \frac{1}{2} K - (Z^T \otimes I) \right\|_{\mathbf{F}}^2 + \left\| \frac{1}{2} K - (I \otimes Z) \right\|_{\mathbf{F}}^2 \\ &+ 2Tr \left(\left(\frac{1}{2} K - (Z^T \otimes I) \right)^T \left(\frac{1}{2} K - (I \otimes Z) \right) \right) \\ &= \left\| \frac{1}{2} K - (Z^T \otimes I) \right\|_{\mathbf{F}}^2 + \left\| \frac{1}{2} K - P^T (Z \otimes I) P \right\|_{\mathbf{F}}^2 \\ &+ 2Tr \left(\left(\frac{1}{2} K - (Z^T \otimes I) \right)^T \left(\frac{1}{2} K - P^T (Z \otimes I) P \right) \right) \\ &= \left\| \frac{1}{2} K - (Z^T \otimes I) \right\|_{\mathbf{F}}^2 + \left\| P^T \left(\frac{1}{2} P K P^T - (Z \otimes I) \right) P \right\|_{\mathbf{F}}^2 \\ &+ 2Tr \left(\left(\frac{1}{2} K - (Z^T \otimes I) \right)^T \left(\frac{1}{2} K - P^T (Z \otimes I) P \right) \right) \\ &= \left\| \frac{1}{2} K - (Z^T \otimes I) \right\|_{\mathbf{F}}^2 + \left\| \left(\frac{1}{2} \hat{K} - (Z \otimes I) \right) \right\|_{\mathbf{F}}^2 \\ &+ 2Tr \left(\left(\frac{1}{2} K - (Z^T \otimes I) \right)^T \left(\frac{1}{2} K - P^T (Z \otimes I) P \right) \right). \end{split}$$

Desarrollamos cada uno de los términos en la última expresión. Para ello consideraremos a la matriz $K \in \mathbb{R}^{n^2 \times n^2}$, particionada en bloques K_{ij} , cada uno de los cuales es una matriz $n \times n$, de modo que

$$\left\| \frac{1}{2}K - (Z^T \otimes I) \right\|_{\mathbf{F}}^2 = \sum_{i=1}^n \sum_{j=1}^n \left\| \frac{1}{2}K_{ij} - (Z^T \otimes I)_{ij} \right\|_{\mathbf{F}}^2 = \sum_{i=1}^n \left\| \frac{1}{2}K_{ij} - z_{ji}I_n \right\|_{\mathbf{F}}^2.$$

Y cada uno de los sumandos

$$\left\| \frac{1}{2} K_{ij} - z_{ji} I_n \right\|_{F}^{2} = Tr \left(\left(\frac{1}{2} K_{ij} - z_{ji} I_n \right)^{T} \left(\frac{1}{2} K_{ij} - z_{ji} I_n \right) \right)$$

$$= \frac{1}{4} Tr \left(K_{ij}^{T} K_{ij} \right) - \frac{1}{2} z_{ji} Tr \left(K_{ij}^{T} \right) - \frac{1}{2} z_{ji} Tr \left(K_{ij} \right)$$

$$+ (z_{ji})^{2} Tr (I_n) = \frac{1}{4} \left\| K_{ij} \right\|_{F}^{2} - z_{ji} Tr \left(K_{ij} \right) + n \left(z_{ji} \right)^{2}.$$

Por lo tanto

$$\begin{aligned} \left\| \frac{1}{2}K - (Z^{T} \otimes I) \right\|_{F}^{2} &= \sum_{j=1}^{n} \sum_{i=1}^{n} \left\| \frac{1}{2}K_{ij} - (Z)_{ji}I_{n} \right\|_{F}^{2} \\ &= \sum_{j=1}^{n} \sum_{i=1}^{n} \left(\frac{1}{4} \|K_{ij}\|_{F}^{2} - (Z)_{ji}Tr(K_{ij}) + n((Z)_{ji})^{2} \right) \\ &= \frac{1}{4} \sum_{j=1}^{n} \sum_{i=1}^{n} \|K_{ij}\|_{F}^{2} - \sum_{j=1}^{n} \sum_{i=1}^{n} (Z)_{ji}Tr(K_{ij}) + n \sum_{j=1}^{n} \sum_{i=1}^{n} ((Z)_{ji})^{2} \\ &= \frac{1}{4} \sum_{j=1}^{n} \sum_{i=1}^{n} \|K_{ij}\|_{F}^{2} - \sum_{j=1}^{n} \sum_{i=1}^{n} (Z)_{ji}Tr(K_{ij}) + n \|Z\|_{F}^{2}. \end{aligned}$$

De manera similar

$$\left\| \frac{1}{2} \hat{K} - (Z \otimes I) \right\|_{\mathbf{F}}^{2} = \frac{1}{4} \sum_{j=1}^{n} \sum_{i=1}^{n} \|\hat{K}_{ij}\|_{\mathbf{F}}^{2} - \sum_{j=1}^{n} \sum_{i=1}^{n} z_{ij} Tr\left(\hat{K}_{ij}\right) + n \|Z\|_{\mathbf{F}}^{2},$$

ya que

$$\sum_{j=1}^{n} \sum_{i=1}^{n} ||K_{ij}||_{F}^{2} = ||K||_{F}^{2}, \qquad \sum_{j=1}^{n} \sum_{i=1}^{n} ||\hat{K}_{ij}||_{F}^{2} = ||\hat{K}||_{F}^{2},$$

У

$$\|\hat{K}\|_{\mathbf{F}}^2 = \|PKP^T\|_{\mathbf{F}}^2 = \|K\|_{\mathbf{F}}^2.$$

Entonces

$$\left\| \frac{1}{2}K - (Z^T \otimes I) \right\|_{\mathbf{F}}^2 = \frac{1}{4} \|K\|_{\mathbf{F}}^2 - \sum_{i=1}^n \sum_{j=1}^n z_{ji} Tr(K_{ij}) + n \|Z\|_{\mathbf{F}}^2, \tag{4}$$

$$\left\| \frac{1}{2}\hat{K} - (Z \otimes I) \right\|_{F}^{2} = \frac{1}{4} \|K\|_{F}^{2} - \sum_{j=1}^{n} \sum_{i=1}^{n} z_{ij} Tr\left(\hat{K}_{ij}\right) + n \|Z\|_{F}^{2}.$$
 (5)

Para el tercer término

$$Tr\left(\left(\frac{1}{2}K - (Z^T \otimes I)\right)^T \left(\frac{1}{2}K - P^T(Z \otimes I)P\right)\right)$$

$$= \frac{1}{4}\|K\|_F^2 - \frac{1}{2}Tr\left(K^T(P^T(Z \otimes I)P)\right) - \frac{1}{2}Tr\left((Z^T \otimes I)^TK\right)\right)$$

$$+Tr\left(((Z \otimes I)^TP)(P^T(Z \otimes I)P)\right).$$

Veamos que

$$\frac{1}{2}Tr\left((Z^{T} \otimes I)^{T}K\right) = \frac{1}{2}Tr\left((Z \otimes I)K\right) = \frac{1}{2}\sum_{j=1}^{n}Tr\left(((Z \otimes I)K)_{jj}\right)
= \frac{1}{2}\sum_{j=1}^{n}Tr\left(\sum_{i=1}^{n}(Z \otimes I)_{ji}K_{ij}\right)
= \frac{1}{2}\sum_{j=1}^{n}Tr\left(\sum_{i=1}^{n}(z_{ji}I)K_{ij}\right) = \frac{1}{2}\sum_{i=1}^{n}\sum_{i=1}^{n}z_{ji}Tr\left(K_{ij}\right).$$

Además

$$\frac{1}{2}tr\left(K^{T}(P^{T}(Z\otimes I)P)\right) = \frac{1}{2}Tr\left(P^{T}PK^{T}P^{T}(Z\otimes I)P\right)$$

$$= \frac{1}{2}Tr\left(P^{T}\left((PK^{T}P^{T})(Z\otimes I)\right)P\right)$$

$$= \frac{1}{2}Tr\left((PK^{T}P^{T})(Z\otimes I)\right) = \frac{1}{2}Tr\left(\hat{K}^{T}(Z\otimes I)\right)$$

$$= \frac{1}{2}Tr\left((Z\otimes I)\hat{K}^{T}\right) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}z_{ij}Tr(\hat{K}_{ji})$$

у

$$Tr\left(((Z \otimes I)^T P^T)((Z \otimes I)P)\right) = Tr\left((Z \otimes I)^T (P^T (Z \otimes I)P)\right)$$
$$= Tr\left((Z^T \otimes I)(I \otimes Z)\right) = Tr\left(Z^T \otimes Z\right)$$
$$= Tr(Z^T).Tr(Z) = (Tr(Z))^2.$$

Luego,

$$Tr\left(\left(\frac{1}{2}K - (Z^{T} \otimes I)\right)^{T} \left(\frac{1}{2}K - P^{T}(Z \otimes I)P\right)\right)$$

$$= \frac{1}{4}||K||_{F}^{2} - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}z_{ji}Tr\left(K_{ij}\right) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}z_{ij}Tr\left(\hat{K}_{ji}\right) + (Tr(Z))^{2}.$$

Finalmente, teniendo en cuenta que la matriz K es simétrica

$$||K - (Z^T \oplus Z)||_{\mathbf{F}}^2 = \left(\frac{1}{4}||K||_{\mathbf{F}}^2 - \sum_{j=1}^n \sum_{i=1}^n z_{ji} Tr(K_{ij}) + n||Z||_{\mathbf{F}}^2\right)$$

11vo Simposio Argentino de Investigacion Operativa, SIO 2013

$$+ \left(\frac{1}{4}\|K\|_{F}^{2} - \sum_{j=1}^{n} \sum_{i=1}^{n} z_{ij} Tr\left(\hat{K}_{ij}\right) + n\|Z\|_{F}^{2}\right)$$

$$+ \frac{1}{2}\|K\|_{F}^{2} - \sum_{i=1}^{n} \sum_{j=1}^{n} z_{ji} Tr\left(K_{ij}\right)$$

$$- \sum_{i=1}^{n} \sum_{j=1}^{n} z_{ij} Tr\left(\hat{K}_{ji}\right) + 2\left(Tr(Z)\right)^{2}$$

$$= \|K\|_{F}^{2} - 2\sum_{i=1}^{n} \sum_{j=1}^{n} \left(z_{ji} Tr\left(K_{ij}\right) + z_{ij} Tr\left(\hat{K}_{ji}\right) - nz_{ij}^{2}\right)$$

$$+ 2\left(\sum_{i=1}^{n} z_{ii}\right)^{2}.$$

El problema (3) queda expresado en términos de los z_{ij} como sigue:

$$\begin{cases}
\min \|K\|_{F}^{2} - 2\sum_{i=1}^{n} \sum_{j=1}^{n} \left(z_{ji} Tr(K_{ij}) + z_{ij} Tr(\hat{K}_{ji}) - n z_{ij}^{2} \right) + 2 \left(\sum_{i=1}^{n} z_{ii} \right)^{2}.
\end{cases}$$
(6)

El siguiente teorema caracteriza las soluciones de (6).

Teorema 1 Para cada matriz simétrica K, la única solución del problema (6) está dada por $Z^* \in \mathbb{R}^{n \times n}$

$$z^{\star}_{ij} = \begin{cases} \frac{Tr(K_{ij}) + Tr(\hat{K}_{ji})}{2n}, & i \neq j \\ \frac{n\left(Tr(K_{ii}) + Tr(\hat{K}_{ii})\right) - Tr(K)}{2n^2}, & i = j \end{cases}$$

 $\mathcal{D}emostraci\'{o}n$

Se trata de minimizar la función definida de ${\rm I\!R}^{n^2}$ en ${\rm I\!R},$ dada por

$$\varphi(z_{11}, z_{21}, \dots, z_{nn}) = \|K\|_{F}^{2} - 2\sum_{i=1}^{n} \sum_{j=1}^{n} \left(z_{ji} Tr(K_{ij}) + z_{ij} Tr(\hat{K}_{ji}) - nz_{ij}^{2}\right) + 2\left(\sum_{i=1}^{n} z_{ii}\right)^{2}.$$

La condición necesaria para este problema es

11vo Simposio Argentino de Investigacion Operativa, SIO 2013

$$\begin{cases}
\frac{\partial \varphi}{\partial z_{ij}} = 4nz_{ij} - 2\left(Tr\left(K_{ij}\right) + Tr\left(\hat{K}_{ji}\right)\right) = 0, & si \quad i \neq j \\
\frac{\partial \varphi}{\partial z_{ii}} = 4nz_{ii} - 2\left(Tr\left(K_{ii}\right) + Tr\left(\hat{K}_{ii}\right)\right) + 4\left(\sum_{i=1}^{n} z_{ii}\right) = 0, & si \quad i = j.
\end{cases}$$
(7)

Luego, para $i \neq j$ es

$$z_{ij} = \frac{Tr(K_{ij}) + Tr(\hat{K}_{ji})}{2n}.$$

Para i = j, veamos que sumando para i = 1, ..., n

$$0 = \sum_{i=1}^{n} \frac{\partial \varphi}{\partial z_{ii}} = \sum_{i=1}^{n} \left(4nz_{ii} - 2\left(Tr(K_{ii}) + Tr(\hat{K}_{ii}) \right) + 4\left(\sum_{i=1}^{n} z_{ii} \right) \right)$$
$$= 4n\sum_{i=1}^{n} z_{ii} - 2\sum_{i=1}^{n} \left(Tr(K_{ii}) + Tr(\hat{K}_{ii}) \right) + 4n\sum_{i=1}^{n} z_{ii},$$

por lo tanto

$$8n\sum_{i=1}^{n} z_{ii} = 2\left(Tr(K) + Tr(\hat{K})\right),$$

y entonces

$$\sum_{i=1}^{n} z_{ii} = \frac{2\left(Tr(K) + Tr(\hat{K})\right)}{8n} = \frac{2\left(Tr(K)\right)}{4n} = \frac{Tr(K)}{2n}.$$
 (8)

Reemplazando (8) en (7) resulta para cada i = 1, ..., n

$$4nz_{ii} - 2\left(Tr(K_{ii}) + Tr(\hat{K}_{ii})\right) + 4\left(\frac{Tr(K)}{2n}\right) = 0,$$
$$z_{ii} = \frac{n\left(Tr(K_{ii}) + Tr(\hat{K}_{ii})\right) - Tr(K)}{2n^2}.$$

Las derivadas parciales segundas son para $i \neq j$

$$\frac{\partial^2 \varphi}{\partial z_{ij}^2} = 4n.$$

Para todo i, j

$$\frac{\partial^2 \varphi}{\partial z_{ij} \partial z_{st}} = 0.$$

Mientras que para i = j,

$$\frac{\partial^2 \varphi}{\partial z_{ii}^2} = 4n + 4.$$

El Hessiano para el problema es una matriz de orden n^2 , con la siguiente estructura

Debe tenerse en cuenta la forma en que fueron ordenados los z_{ij} como variables del problema. Se demostrará que H es definida positiva analizando el signo de los menores de dicha matriz.

Sea $[M_k]$ un menor de orden k y $1 \le t \le n$. Si $(t-1)(1+n) \le k \le t(n+1)-1$, entonces

$$[M_k] = 4^{k+1}(n+t)n^k,$$

que es siempre positivo ya que n y t son valores naturales. Luego el Hessiano es definido positivo y Z^* es un minimizador de (6).

3. Posible Aplicación.

Supongamos que queremos obtener una aproximación cuadrática para una función matricial $\mathbf{F}: \mathbb{R}^{n \times n} \longrightarrow \mathbb{R}^{n \times n}$, es decir

$$\mathbf{Q}(\Delta) = \Delta^T \mathbf{A} \Delta + \mathbf{B}^T \Delta + \Delta^T \mathbf{B},$$

donde Δ , **A** y **B** son matrices en $\mathbb{R}^{n \times n}$. Es necesario, en este caso, contar con matrices que aproximen en algún sentido a las derivadas primera y segunda de **F**, ya que la no conmutatividad del producto de matrices hace imposible aislar las derivadas a partir de los correspondientes diferenciales de Fréchet, $D\mathbf{F}(X,\Delta)$.

De la vectorización de la aproximación cuadrática resulta

$$vec(\mathbf{Q}) = vec(\Delta^T \mathbf{A} \Delta + \mathbf{B}^T \Delta + \Delta^T \mathbf{B})$$
 (9)

$$= vec\left(\Delta^T \mathbf{A} \Delta\right) + vec\left(\mathbf{B}^T \Delta + \Delta^T \mathbf{B}\right)$$
 (10)

$$= vec(\Delta^T \mathbf{A} \Delta) + ((\mathbf{B}^T \otimes \mathbf{I}) + (\mathbf{I} \otimes \mathbf{B})) vec(\Delta).$$
 (11)

Es posible obtener la llamada forma de Kronecker de la derivada de Fréchet. Higham [5] demuestra que

$$vec(D\mathbf{F}(X,\Delta)) = K(X)vec(\Delta),$$
 (12)

para alguna matriz K(X) de dimensión $n^2 \times n^2$, independiente de Δ , llamada forma de Kronecker de la derivada de Fréchet de \mathbf{F} .

Si observamos la vectorización de la parte lineal de \mathbf{Q} , segundo término en (12), es razonable hallar una matriz simétrica Z tal que $\left((Z^T \otimes I) + (I \otimes Z)\right)$ aproxime a K(X), entonces dicha matriz Z sería la aproximación de la derivada primera. Proponemos hallar la matriz que aproxima a la derivada primera como solución de

$$\begin{cases}
\min \|K(X) - (Z^T \oplus Z)\|_{\mathbf{F}}^2.
\end{cases}$$
(13)

4. Conclusiones.

Hemos resuelto el Problema de Procrusto que involucra Sumas de Kronecker y se sugiere una posible aplicación para dichos resultados. Precisamente el estudio de tal problema de optimización ha sido motivado por la necesidad de buscar una aproximación de las derivadas para funciones matriciales, es decir como parte de la solución de un problema más general. Aún resta analizar

la cuestión desde el punto de vista numérico y verificar si la aproximación es conveniente, así como también encontrar la forma de aproximar la derivada segundo orden.

Bibliografía

- [1] M.G. EBERLE. Alternating projections methods for persymmetric optimization problems. Master's thesis, Departmento de Matemática, Universidad Nacional del Sur, Bahía Blanca, Buenos Aires, Argentina, 1999.
- [2] M.G. EBERLE and M.C. MACIEL. Finding the closest Toeplitz matrix.

 Computational and Applied Mathematics, 22(1):1–18, 2003.
- [3] R. ESCALANTE and M. RAYDAN. Dykstra's algorithm for a constrained least-squares matrix problem. *Numerical Linear Algebra with applications*, 9(6):459–471, 1996.
- [4] N. J. HIGHAM. The symmetric Procrustes problem. BIT, 28:133–143, 1988.
- [5] N. J. HIGHAM. Function of Matrices. Theory and computation. SIAM, Philadelphia, Pennsylvania, 2008.
- [6] R. A. HORN and C. R. JOHNSON. Matrix Analysis. Cambridge University Press, New York, 1992.
- [7] C.F. VAN LOAN and N. PITSIANIS. Approximation with Kronecker products. Linear Algebra for Large Scale and Real-Time Applications, pages 293–314, 1993.