RAIRO. RECHERCHE OPÉRATIONNELLE

A. QANNARI
E. M. QANNARI

Approximation d'une matrice non inversible par une matrice inversible et D-inverse

RAIRO. Recherche opérationnelle, tome 27, nº 3 (1993), p. 319-335
<http://www.numdam.org/item?id=RO_1993__27_3_319_0>

Recherche opérationnelle/Operations Research
(vol. 27, n° 3, 1993, p. 319 à 335)

APPROXIMATION D'UNE MATRICE NON INVERSIBLE PAR UNE MATRICE INVERSIBLE ET D-INVERSE (*)

par A. QANNARI ([1]) et E. M. QANNARI ([1])

Communiqué par Jean-Yves JAFFRAY

Résumé. – *Nous nous proposons d'approcher, au sens de la norme de Hilbert-Schmidt, une matrice non inversible par une matrice inversible de déterminant donné. Nous démontrons l'existence et l'unicité d'une telle matrice. Nous définissons la D-inverse d'une matrice à partir de la solution du problème précédent en choisissant pour déterminant le produit des valeurs singulières non nulles de la matrice considérée. Nous discutons l'application de cette approche pour améliorer le conditionnement d'une matrice.*

Mots clés : Inverse d'une matrice, théorème de Von Neumann, décomposition spectrale, conditionnement d'une matrice.

Abstract. – *We discuss an approximation (with respect to Hilbert-Schmidt norm) of a singular matrix by a non singular matrix with a given determinant. We prove the existence and unicity of this matrix. We use this result to define the D-inverse of a matrix by requiring the determinant to be the product of the nonzero singular values of the given matrix. We discuss the interest of this approach for improving the conditioning of a given matrix.*

Keywords: Matrix inverse, Von Neumann Theorem, Spectral Decomposition, Condition Number.

INTRODUCTION

La résolution de beaucoup de problèmes en statistique et dans d'autres disciplines met en œuvre l'inversion d'une matrice. Lorsque cette matrice n'est pas inversible, il est d'usage de travailler avec une pseudo-inverse. La pseudo-inverse dite de Moore-Penrose (Moore, 1977; Penrose, 1955) présente, entre autres, l'avantage d'être unique et facile à calculer. L'exemple suivant montre, néanmoins, un inconvénient de cette inverse : pour une matrice (3,3) diagonale non inversible et de rang 2 :

$$D = \operatorname{diag}(d_1,\ d_2,\ 0)$$

l'inverse de Moore-Penrose est donnée par :

(*) Received January 1992.

([1]) E.N.I.T.I.A.A., La Géraudière, 44072 Nantes Cedex 03.

$$D^{+} = \operatorname{diag}\left(d_1^{-1}, d_2^{-1}, 0\right).$$

Il s'ensuit que lorsque d_2, par exemple, est proche de zéro, d_2^{-1} sera relativement grand. En d'autres termes d_2 et 0 qui sont, *a priori*, de même ordre de grandeur interviennent dans le calcul de l'inverse avec des « poids » très disproportionnés. Il n'en est pas de même dans la démarche que nous proposons.

1. DONNÉES

Soit G une matrice diagonale d'ordre p et de rang r ($r < p$) :

$$G = \operatorname{diag}(g_1, g_2, \ldots, g_p).$$

Nous supposerons que ses éléments sont non négatifs et, sans perte de généralité, qu'ils sont classés par ordre décroissant. La matrice G étant de rang r ($r < p$), nous avons :

$$g_j = 0, \qquad \forall j \in \{r+1, r+2, \ldots, p\}.$$

Nous considérons c un réel strictement positif et désignons par $\mathcal{M}_c$ l'ensemble des matrices diagonales d'ordre p d'éléments diagonaux positifs et de déterminant c.

Nous nous intéressons au problème d'approximation de la matrice G par un élément de $\mathcal{M}_c$. En désignant par $\|.\|$ la norme de Hilbert-Schmidt dans l'ensemble des matrices (p, p), nous cherchons à résoudre le problème d'optimisation suivant :

$$(\mathcal{P}) \qquad \min_{F \in \mathcal{M}_c} \|G - F\|^2$$

2. EXISTENCE ET UNICITÉ DE LA SOLUTION

2.1. Proposition

Le problème $(\mathcal{P})$ *admet une solution unique.*

En effet, nous avons :

$$\|G - F\|^2 = \|G\|^2 + \|F\|^2 - 2 \, \mathit{Trace}\,(GF).$$

où Trace (F) désigne la trace de la matrice F.

Le problème $(\mathcal{P})$ se ramène à chercher le minimum de la fonction :

$$h(F) = \|F\|^2 - 2 \operatorname{Trace}(GF)$$

ou encore, puisque F est une matrice diagonale :

$$h(F) = \text{Trace}(F^2) - 2\,\text{Trace}(GF)$$

avec la contrainte Det $(F) = c$.

Soit 2λ le multiplicateur de Lagrange associé à la contrainte Det $(F) = c$; la fonction de Lagrange s'écrit :

$$\Lambda(F, \lambda) = \text{Trace}(F^2) - 2\,\text{Trace}(GF) - 2\lambda\,[Det\,(F) - c].$$

En annulant les dérivées partielles de Λ par rapport à F et λ, nous obtenons le système d'équations suivant :

$$\left.\begin{array}{c} F - G - \lambda\,\text{Det}(F)\,F^{-1} = 0 \\ \text{Det}(F) = c \end{array}\right\}$$

ou encore, en notant I la matrice identité d'ordre p :

$$\left.\begin{array}{c} F^2 - GF - \lambda\,c\,I = 0 \\ \text{Det}(F) = c \end{array}\right\} \qquad (\star)$$

En notant $F = \text{Diag}(f_1, f_2, \ldots, f_p)$ le système d'équations $(\star)$ s'écrit pour tout $j = 1, \ldots, p$:

$$\left.\begin{array}{c} f_j^2 - g_j f_j - \lambda\,c = 0 \\ \displaystyle\prod_{i=1}^{p} fi = c \end{array}\right\} \qquad (\star\star)$$

En particulier, nous avons :

$$f_j^2 = \lambda\,c, \qquad \forall j \in \{r+1, \ldots, p\}.$$

Ceci prouve que λ est une valeur positive et que les équations du second degré données dans $(\star\star)$ ont pour solutions :

$$\left.\begin{array}{l} f_j = \dfrac{g_j + \sqrt{g_j^2 + 4\,\lambda\,c}}{2} \quad \text{ou} \quad f_j = \dfrac{g_j - \sqrt{g_j^2 + 4\,\lambda\,c}}{2} \quad \text{pour } j = 1, \ldots, r \\ \text{et} \\ \qquad f_j = \sqrt{\lambda c} \quad \text{pour } j = r+1, \ldots, p \qquad \text{avec} \quad \displaystyle\prod_{i=1}^{p} f_{i=c}. \end{array}\right\}$$

Comme nous avons contraint les valeurs f_j ($j = 1, \ldots, p$) à être positives, nous ne conserverons que les racines positives.

Nous allons maintenant prouver l'existence et l'unicité de la valeur λ.

L'égalité $\prod_{j=1}^{p} f_j = c$ s'écrit :

$$\prod_{j=1}^{p} \frac{g_j + \sqrt{g_j^2 + 4\,\lambda\,c}}{2} = c.$$

Les deux membres de cette équation étant positifs, elle peut aussi s'écrire, en considérant les logarithmes népériens :

$$\sum_{j=1}^{p} \mathrm{Log}\left(g_j + \sqrt{g_j^2 + 4\,\lambda\,c}\right) = \mathrm{Log}\,(2^p c).$$

Soit la fonction numérique définie sur $]0, +\infty[$ par :

$$k\,(x) = \sum_{j=1}^{p} \mathrm{Log}\left(g_j + \sqrt{g_j^2 + 4\,x\,c}\right) - \mathrm{Log}\,(2^p c)$$

Nous avons $\lim_{x \to 0} k\,(x) = -\infty$ car au moins la valeur g_p est nulle.

De même, nous avons $\lim_{x \to +\infty} k\,(x) = +\infty$.

La dérivée de k s'écrit :

$$k'\,(x) = \sum_{j=1}^{p} \frac{2\,c}{\left(g_j + \sqrt{g_j^2 + 4\,x\,c}\right)\sqrt{g_j^2 + 4\,x\,c}}$$

Nous remarquons que k' est positive et nous en déduisons que la fonction k est strictement croissante. Elle s'annule donc en un point unique λ.

Dans la suite, la matrice F^{-1} inverse de la matrice F sera appelée D_c–inverse de G.

2.2. Remarques

1) La valeur λ peut être déterminée par l'algorithme de Newton ou par l'algorithme de dichotomie qui consistent à déterminer le zéro d'une fonction par itérations successives en « cernant » de proche en proche cette valeur :

2) La valeur λ est comprise entre 0 et $c^{(2-p)/p}$. En effet :

$$k\left(c^{(2-p)/p}\right) \geq \sum_{j=1}^{p} \mathrm{Log}\,(2\,c^{1/p}) - \mathrm{Log}\,(2^p\,c)$$

d'où $k\left(c^{(2-p)/p}\right) \geq 0$.

Comme k est croissante et tend vers $-\infty$ quand x tend vers 0, nous en déduisons que $\lambda \in]0,\ c^{(2-p)/p}]$.

3) Nous discutons une autre majoration de λ qui peut être plus adaptée pour certains choix de la valeur c. Son intérêt est de tenir compte du nombre des éléments diagonaux non nuls de G et de leur produit (donc de leur importance numérique).

Désignons par d le produit des éléments diagonaux non nuls de G, $d = \prod_{j=1}^{r} g_j$, nous avons la relation :

$$c = \prod_{j=1}^{p} f_j$$

$$= \prod_{j=1}^{r} \left(\frac{g_j + \sqrt{g_j^2 + 4\,\lambda\, c}}{2} \right) \prod_{j=r+1}^{p} \sqrt{\lambda\, c}$$

$$= \prod_{j=1}^{r} \left(\frac{g_j + \sqrt{g_j^2 + 4\,\lambda\, c}}{2} \right) (\lambda\, c)^{(p-r)/2}$$

En remarquant que, pour tout j, nous avons :

$$\frac{g_j + \sqrt{g_j^2 + 4\,\lambda\, c}}{2} \geq g_j$$

il s'ensuit que :

$$c \geq (\lambda\, c)^{(p-r)/2} \prod_{j=1}^{r} g_j$$

d'où

$$c \geq d\,(\lambda\, c)^{(p-r)/2}$$

et par conséquent :

$$\lambda \leq \frac{1}{c} \left(\frac{c}{d}\right)^{2/(p-r)}$$

3. ERREUR D'APPROXIMATION

A partir de l'égalité ($\star$) donnée au paragraphe 2, nous déduisons l'égalité :

$$F^2 - GF = \lambda\, c\, I$$

Cette égalité peut s'écrire encore :

$$GF^{-1} = I - \lambda\, c\, F^{-2}$$

Ceci peut s'interpréter grossièrement en disant que GF^{-1} est égal à l'identité à $\lambda\, c\, F^{-2}$ près.

Il est aisé de vérifier que :

$$\|G - F\|^2 = \|\lambda\, c\, F^{-1}\|^2$$

Cette égalité permet de trouver une majoration de l'erreur d'approximation En remarquant que $\sqrt{1/\lambda\, c}$ est la plus grande valeur propre de F^{-1}, nous avons :

$$\|G - F\|^2 \leq p\, \lambda\, c$$

Comme nous avons montré que $\lambda \leq c^{(2-p)/p}$ *et* $\lambda \leq \dfrac{1}{c}\left(\dfrac{c}{d}\right)^{2/(p-r)}$ il s'ensuit :

$$\|G - F\|^2 \leq \min\left(p\, c^{2/p}\, ; \quad p\left(\frac{c}{d}\right)^{2/(p-r)}\right)$$

4. D-INVERSE D'UNE MATRICE DIAGONALE

4.1. Définition

Soit G une matrice (p, p) diagonale de rang r et $g_1, \ldots, g_r$ ses éléments diagonaux non nuls.

Nous définissons la D-inverse de la matrice G comme étant l'inverse de la matrice F, solution du problème $(\mathcal{P})$ avec le choix de $c = \prod_{i=1}^{r} g_i$

4.2. Remarque

Cette définition trouve sa justification dans le fait que la notion de déterminant est liée à la notion de volume et en statistique, à la notion de variance généralisée (Anderson, 1971). Nous avons donc proposé cette approximation de la matrice G par une matrice F selon le principe de « toutes choses égales par ailleurs ».

4.3. Exemples

Pour illustrer la démarche proposée, nous donnons des exemples de matrices non inversibles, leur D-inverse et, à titre illustratif, leur inverse selon Moore-Penrose :

G	diag (1, 1, 0)	diag (1, 0.10, 0.001, 0)
F	diag (1.285, 1.285, 0.605)	diag (1.001, 0.108, 0.031, 0.030)
D-inverse	diag (0.785, 0.785, 1.651)	diag (0.999, 9.227, 32.662, 33.208)
Moore-Penrose ..	diag (1, 1, 0)	diag (1, 10, 1 000, 0)

5. D-INVERSE D'UNE MATRICE CARRÉE

Soit A une matrice carrée (p, p), la décomposition spectrale de A est donnée par :

$$A = PG\,{}^tQ$$

avec P et Q deux matrices orthogonales d'ordre p et $G = \text{diag}\,(g_1, \dots, g_p)$ où g_i $(i = 1, \dots, p)$ sont les racines carrées des valeurs propres de ${}^tAA$ et, par conséquent, positives ou nulles.

5.1. Définition

Soit c une valeur positive, nous définissons la D_c-inverse de A comme étant l'inverse de la matrice $B = P\,F\,{}^tQ$, où F est la matrice diagonale qui approche la matrice G au sens du critère $(\mathcal{P})$.

Remarquons que le déterminant de la matrice B est donné par :

$\text{Det}\,(B) = \det\,(P) \det\,(F) \det\,(Q) = \mp c$

5.2. Proposition

La matrice B est solution du problème :

$$\min_B \|A - B\|^2 \qquad (\mathcal{P}^\star)$$

$$\text{sous la contrainte } |\text{Det}\,(B)| = c$$

En effet :

considérons la décomposition spectrale de B, $B = RF^tS$ où R et S sont des matrices orthogonales d'ordre p, Il s'ensuit :

$$\begin{aligned} \|B - A\|^2 &= \|A\|^2 + \|B\|^2 - 2\,\mathrm{trace}\,(A\,{}^tB) \\ &= \|F\|^2 + \|G\|^2 - 2\,\mathrm{trace}\,(PG^tQSF^tR) \end{aligned}$$

la quantité trace (PG^tQSF^tR) peut s'écrire trace (G^tQSF^tRP).

Nous pouvons facilement vérifier que les matrice ${}^tQS$ et ${}^tRP$ sont orthogonales. Nous en déduisons, par le théorème de Von-Neumann (Rao, 1979) que la quantité trace (G^tQSF^tRP) est maximale lorsque ${}^tRP = \mathrm{I}$ et ${}^tQS = \mathrm{I}$ ou encore :

$$P = R \qquad \text{et } Q = S$$

Il s'ensuit que :

$$\|B - A\|^2 \leq \|F\|^2 + \|G\|^2 - 2\,\mathrm{trace}\,(FG).$$

Soit encore :

$$\|B - A\|^2 \leq \|F - G\|^2$$

Nous en déduisons que $B = PF^tQ$ est la matrice qui réalise le minimum.

5.3. Remarque

Il est aisé de vérifier que la matrice B, D_c-inverse de A, vérifie les propriétés d'hermicité des unités, à savoir :

$$AB^{-1} = {}^t(AB^{-1})$$

et

$$B^{-1}A = {}^t(B^{-1}A)$$

6. APPLICATIONS

Dans le calcul numérique, l'étude de la stabilité des solutions présente un intérêt primordial. Pour le calcul numérique matriciel, cette stabilité est, en général, liée au conditionnement des matrices. Dans ce contexte, la démarche que nous proposons permet de substituer à une matrice mal conditionnée une matrice de meilleur conditionnement améliorant ainsi la stabilité des solutions. Nous nous proposons d'illustrer ceci par deux exemples. Dans un premier temps, nous traitons le cas d'une matrice non inversible, et dans un deuxième temps, nous proposons une manière de perturber une matrice afin d'améliorer son conditionnement.

6.1. Application pour l'étude d'un système linéaire

Considérons le système :

$$AX = B \quad (*)$$

où

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \text{et} \quad X = \begin{bmatrix} x \\ y \end{bmatrix}$$

En utilisant l'inverse généralisée de Moore-Penrose, la solution du système est donnée par $x = y = (a + b)/4$.

En particulier dans le cas où $a = b$, nous avons $x = y = a/2$.

Les solutions ainsi obtenues ne sont pas stables. En effet, si nous substituons à la matrice A, la matrice A' définie par :

$$A' = \begin{bmatrix} 1 & 1,001 \\ 1 & 1 \end{bmatrix}$$

la solution devient :

$$x = 1\,000,99\,a - 999,99\,b$$
$$y = -999,99\,a + 999,99\,b$$

La matrice qui approche A selon le programme $(\mathcal{P})$, en choisissant pour c le produit des valeurs propres non nulles de A, est donnée par :

$$E = \begin{bmatrix} 1,59 & 0,73 \\ 0,73 & 1,59 \end{bmatrix}$$

et la solution du système :

$$EX = B \quad (**)$$

est donnée par :

$$x = 0,796\,a - 0,365\,b$$
$$y = 0,796\,b - 0,365\,a$$

En particulier dans le cas où $a = b$ nous avons $x = y = 0{,}43\,a$

Il apparait que nous avons substitué au système (*) le système (**) qui est plus stable car une perturbation de E de même nature que celle effectuée sur A ne modifie que très légèrement les solutions.

6.2. Cas d'une matrice inversible mal conditionnée

Lorsque la matrice A intervenant dans un système linéaire est inversible mais mal conditionnée, divers auteurs ont proposé des démarches pour l'approcher par une matrice de meilleur conditionnement. Nous citons, par exemple, la démarche de Marquardt (Marquardt, 1970) qui introduit la notion de rang fractionnaire d'une matrice.

Pour une matrice A de rang p admettant des valeurs propres proches de zéro, il est proposé de lui associer un rang fractionnaire, f. Le choix de f a été discuté par l'auteur. Par la suite il est proposé de prendre comme inverse généralisée de A la matrice A^+ donnée par :

$$A^+ = \sum_1^g \frac{1}{\mu_k} u_k \, {}^t v_k + \frac{f-g}{\mu_{g+1}} u_{g+1} \, {}^t v_{g+1}$$

où g est la partie entière de f, u_k (resp. v_k) pour $k = 1, \dots, g+1$ sont les vecteurs propres de A^tA (resp. ${}^tAA$) associés aux valeurs propres μ_k^2.

Dans la stratégie de la Ridge régression (Hocking, 1976), il est proposé de substituer à la matrice A, une matrice de la forme $(A+kI)$ où I désigne la matrice identité et k une constante dont le choix a été discuté par plusieurs auteurs. Il est aisé de vérifier que ces deux démarches améliorent le conditionnement de la matrice A.

La stratégie que nous proposons s'inspire de ces deux démarches. Pour la matrice A inversible d'ordre p, nous considérons la décomposition spectrale sous la forme :

$$A = \sum_1^p \mu_k \, u_k \, {}^t v_k$$

La matrice $A' = \sum_1^{p-1} \mu_k \, u_k \, {}^t v_k$ est singulière et nous nous proposons de l'approcher par la matrice B' selon le programme $(\mathcal{P})$. Ceci permet de proposer une approximation de la matrice A par la matrice B définie par :

$$B = B' + \mu_p \, u_p \, {}^t v_p$$

La matrice B ainsi obtenue est de meilleur conditionnement que la matrice A. En effet, le conditionnement de la matrice A est :

$$\text{cond}\,(A) = \frac{\mu_1}{\mu_p}$$

et le conditionnement de la matrice B est cond (B)

$$= \frac{1}{2} \frac{\mu_1 + \sqrt{\mu_1^2 + 4\,\lambda\,c}}{\mu_p + \sqrt{\lambda\,c}} \qquad \text{si} \quad \mu_p + \sqrt{\mu\,c} \leq \frac{\mu_{p-1} + \sqrt{\mu_{p-1}^2 + 4\,\lambda\,c}}{2}$$

$$= \frac{\mu_1 + \sqrt{\mu_1^2 + 4\,\lambda\,c}}{\mu_{p-1} + \sqrt{\mu_{p-1}^2 + 4\,\lambda\,c}} \qquad \text{sinon}$$

Dans le premier cas, nous avons :

$$\sqrt{\mu_1^2 + 4\,\lambda\,c} \leq \sqrt{\mu_1^2} + \sqrt{4\,\lambda\,c}$$

soit

$$\mu_p \sqrt{\mu_1^2 + 4\,\lambda\,c} \leq \mu_p\,\mu_1 + 2\,\mu_p \sqrt{\lambda\,c}$$

d'où, μ_p étant inférieur à μ_1,

$$\mu_p \sqrt{\mu_1^2 + 4\,\lambda\,c} \leq \mu_p\,\mu_1 + 2\,\mu_1 \sqrt{\lambda\,c}$$

soit encore, en ajoutant $\mu_p\,\mu_1$ aux deux membres de l'inégalité :

$$\mu_p \sqrt{\mu_1^2 + 4\,\lambda\,c} + \mu_p\,\mu_1 \leq 2\,\mu_p\,\mu_1 + 2\,\mu_1 \sqrt{\lambda\,c}$$

ce qui entraîne :

$$\mu_p \left(\sqrt{\mu_1^2 + 4\,\lambda\,c} + \mu_1 \right) \leq 2\,\mu_1 \left(\mu_p + \sqrt{\lambda\,c} \right)$$

et par conséquent :

$$\text{cond}\,(B) \leq \text{cond}\,(A).$$

De même, nous pouvons montrer que cond $(B) \leq$ cond (A) dans le second cas.

Remarque : Pour des raisons de simplicité, nous avons « tronqué » la matrice A au niveau de la dernière valeur propre. Il est bien évident que cette démarche peut être étendue aux p-r dernières valeurs singlières qui sont estimées proches de zéro.

Nous nous proposons d'illustrer cette démarche par un cas pratique proposé par R. S. Wilson et cité dans Ciarlet (Ciarlet, 1982).

Le système linéaire $AX = B$ avec :

$$A = \begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix}, \qquad B = \begin{bmatrix} 32 \\ 23 \\ 33 \\ 31 \end{bmatrix} \qquad \text{et} \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

admet pour solution le vecteur $X = {}^t(1,\ 1,\ 1,\ 1)$

Considérons maintenant le système perturbé où B a été modifié à la manière suivante $B' = {}^t(32.1,\ 22.9,\ 33.1,\ 30.9)$, la solution devient $X' = {}^t(9.2,\ 12.6,\ 4.5,\ -1.1)$

Considérons également le système perturbé où, cette fois ce sont les éléments de la matrice A qui sont modifiés :

$$A' = \begin{bmatrix} 10 & 7 & 8,1 & 7,2 \\ 7,08 & 5,04 & 6 & 5 \\ 8 & 5,89 & 9,89 & 9 \\ 6,99 & 4,55 & 9 & 9,89 \end{bmatrix}.$$

La solution devient $X'' = \begin{bmatrix} -81 \\ 137 \\ -34 \\ 22 \end{bmatrix}$

Cette instabilité de la solution s'explique par le fait que la matrice A est mal conditionnée (cond $(A) = 2\,984$).

Nous avons approché cette matrice selon la stratégie indiquée ci-dessus, en prenant comme choix de c, le produit des valeurs singulières. La matrice F obtenue a un conditionnement égal à 44,3 et le système $FX = B$ admet pour solution :

$$X \;=\; {}^t(1.12,\ 0.80,\ 1.05,\ 0.97).$$

Lorsque le second membre est perturbé, comme auparavant, la solution du nouveau système est $X' = {}^t(1.22,\ 0.61,\ 1.16,\ 0.89)$.

Nous avons par la suite, considéré la matrice perturbée, A', et nous lui avons appliqué également la procédure décrite dans ce paragraphe, la solution du nouveau système est $X'' = {}^t(1.00,\ 0.82,\ 1.16,\ 0.94)$.

7. EXTENSIONS

7.1. D_c-inverse pondérée

Le programme $(\mathcal{P})$ énoncé précédemment peut se généraliser en introduisant le produit scalaire par rapport à deux matrices définies positives, Σ et Φ :

Pour deux matrices A et B d'ordre p, le produit scalaire par rapport à Σ et Φ est :

$$\langle A,\ B\rangle_{\Sigma,\,\Phi} = \text{Trace}\,(A\,\Sigma\,{}^tB\,\Phi).$$

En désignant par $\|.\|_{\Sigma,\Phi}$ la norme associée à ce produit scalaire, le problème s'énonce alors, de la manière suivante :

Étant donné une matrice singulière A d'ordre p et un réel c, nous cherchons une matrice B de rang p et de déterminant c telle que :

$$\|A - B\|^2_{\Sigma,\,\Phi} \text{ soit minimum.}$$

Nous avons :

$$\|A - B\|^2_{\Sigma,\,\Phi} = \|A\|^2_{\Sigma,\,\Phi} + \|B\|^2_{\Sigma,\,\Phi} - 2\,\langle A,\ B\rangle_{\Sigma,\Phi}$$

Nous sommes donc ramenés à minimiser la fonction :

$$h\,(B) = \text{Trace}\,(B\,\Sigma\,{}^tB\,\Phi) - 2\,\text{Trace}\,(B\Sigma\,{}^tA\,\Phi)$$

sous la contrainte det $(B) = c$.

Considérons la décomposition spectrale généralisée de A et B sous la forme :

$$\begin{array}{lll} A = PG^tQ & \text{avec} \quad {}^tP\Phi P = I & \text{et} \quad {}^tQ\Sigma Q = I\,; \\ B = RF^tS & \text{avec} \quad {}^tR\Phi R = I & \text{et} \quad {}^tS\Sigma S = I \end{array}$$

Ces représentations s'obtiennent en écrivant la décomposition en valeurs singulières de $\Phi^{1/2}\ A\Sigma^{1/2}$ sous la forme :

$$\Phi^{1/2}\,A\Sigma^{1/2} = MG^tN \qquad \text{avec} \quad {}^tMM = I \quad \text{et} \quad {}^tNN = I$$

nous en déduisons

$$A = \Phi^{-1/2}\,MG\,{}^tN\Sigma^{-1/2}$$

soit par conséquent

$$P = \Phi^{-1/2}M \qquad \text{et} \qquad Q = \Sigma^{-1/2}N.$$

De la même manière, en écrivant la décomposition spectrale de $\Phi^{-1/2} B \Sigma^{-1/2}$ sous la forme $\Phi^{-1/2} B \Sigma^{-1/2} = U F {}^t U$, nous avons :

$$R = \Phi^{-1/2} U, \qquad S = \Sigma^{-1/2} V \quad \text{avec} \quad {}^t U U = I \quad \text{et} \quad {}^t V V = I$$

Il s'ensuit que $h(B)$ s'écrit :

$$\begin{aligned} h(B) &= \text{Trace}\,(R F {}^t S\, \Sigma S F {}^t R \Phi) - 2\,\text{Trace}\,(R F {}^t S\, \Sigma\, Q G {}^t P \Phi) \\ &= \text{Trace}\,({}^t R \Phi R\, F\, {}^t S \Sigma S\, F) - 2\,\text{Trace}\,({}^t P \Phi R\, F\, {}^t S \Sigma Q\, G) \\ &= \text{Trace}\,(F^2) - 2\,\text{Trace}\,({}^t P F R\, F\, {}^t S S Q\, G) \end{aligned}$$

Or

$$\begin{aligned} {}^t P F R &= {}^t M\, \Phi^{-1/2}\, \Phi \Phi^{-1/2}\, U \\ &= {}^t M U \end{aligned}$$

De même

$${}^t S \Sigma Q = {}^t V N$$

Ceci permet de vérifier que ${}^t P \Phi R$ et ${}^t S \Sigma Q$ sont orthogonales. La quantité Trace $({}^t P \Phi R\, F \quad {}^t S \Sigma Q G)$ atteint donc son maximum, d'après le théorème de Von Neumann, pour :

$${}^t P \Phi R = I \qquad \text{et} \quad {}^t S \Sigma Q = I$$

ou encore

$$M = U \qquad \text{et} \qquad N = V$$

et par conséquent

$$P = R \qquad \text{et} \quad S = Q.$$

Il s'ensuit ainsi que :

$$\|A - B\|^2_{\Sigma,\, \Phi} \leq \|F - G\|^2.$$

Nous sommes donc ramenés à approcher la matrice G (à éléments diagonaux non négatifs) par la matrice F de déterminant $c' = |c| \det(\Phi^{1/2}) \det(\Sigma^{1/2})$. La solution de ce problème a été discutée au paragraphe 2.

7.2. Approximation par une *D*-*g*-inverse

Nous avons remarqué (§ 5.2) que la *D*-inverse vérifie deux des propriétés de l'inverse de Moore Penrose, nous imposons, dans ce paragraphe, comme contrainte supplémentaire, la troisième propriété à savoir d'être une g-inverse. Nous nous proposons d'approcher une matrice singulière A par une matrice inversible B selon le programme $(\mathcal{P})$ avec la contrainte supplémentaire que B est une g-inverse de A.

7.2.a. Cas d'une matrice diagonale

Soit c un réel positif et G une matrice diagonale d'ordre p et de rang r à élements non négatifs $G = \text{diag}\,(g_1, \dots, g_p)$ avec $g_{r+1} = \dots = g_p = 0$;

Nous nous proposons de chercher une matrice $F = diag\,(f_1, \dots, f_p)$ à éléments diagonaux strictement positifs telle que :

$$\left.\begin{array}{c} \|F - G\|^2 \text{ minimum} \\ \det(F) = c \\ GF^{-1}G = G. \end{array}\right\} \qquad (\mathcal{P}'')$$

Ce programme s'écrit aussi :

$$\sum_1^p f_i^2 - 2\sum_1^p f_i g_i \text{ minimum}$$

$$\prod_1^p f_i = c$$

$$g_i^2\, f_i^{-1} = g_i \qquad \text{pour} \quad i = 1, \dots, p \quad (*)$$

A partir des contraintes (*), nous déduisons $f_i = g_i$ pour $i = 1, \dots, r$.

En notant $d = \prod_1^r g_i$, le programme $(\mathcal{P}'')$ s'écrit donc :

$$\sum_1^{p-r} f_{r+i}^2 \text{ minimum}$$

$$\prod_1^{p-r} f_{r+i} = \tfrac{c}{d}$$

$$f_i = g_i \qquad \text{pour} \quad i = 1, \dots, r$$

Soit 2λ le multiplicateur de Lagrange associé à la contrainte

$$\prod_1^{p-r} f_{r+i} = \frac{c}{d} ;$$

La fonction de Lagrange s'écrit :

$$\Lambda\,(f_{r+1}, \dots , f_p \,;\, \lambda) = \sum_{1}^{p-r} f_{r+1}^2 - 2\lambda \left(\prod_{1}^{p-r} f_{r+i} - \frac{c}{d} \right)$$

En dérivant cette fonction par rapport aux divers arguments, nous obtenons le système suivant :

$$f_{r+k} - \lambda\, \tfrac{c}{d}\, \tfrac{1}{f_{r+k}} = 0 \qquad \text{pour} \quad k = 1, \dots , p - r$$
$$\prod_{1}^{p-r} f_{r+i} = \tfrac{c}{d}$$

d'où

$$f_{r+k}^2 - \lambda\, \tfrac{c}{d} = 0 \qquad \text{pour} \quad k = 1, \dots , p - r$$
$$\prod_{1}^{p-r} f_{r+i} = \tfrac{c}{d}$$

Les éléments f_k étant supposés positifs, il vient que :

$$f_{r+k} = \sqrt{\lambda\, \tfrac{c}{d}} \qquad \text{pour} \quad k = 1, \dots , p - r$$
$$\prod_{1}^{p-r} f_{r+i} = \tfrac{c}{d}$$

D'où la solution du problème $(\mathcal{P}'')$ est donnée par :

$$f_i = g_i \qquad \text{pour} \quad i = 1, \dots , r$$
$$f_i = \left(\frac{c}{d}\right)^{1/(p-r)} \qquad \text{pour} \quad i = r + 1, \dots , p$$

Remarques :

1) Il est aisé de voir que la matrice F trouvée est aussi solution du programme suivant :

F de norme minimum telle que :

$\|GF^{-1} - I\|^2$ minimum

et det $(F) = c$

2) Si $c = d$, nous avons $f_i = 1$ pour $i = r + 1, \dots , p$. Ainsi, la D-g-inverse de la matrice G d'ordre p, $G = \text{diag}\,(g_1, \dots , g_r, 0, \dots , 0)$, est donnée par $F^{-1} = \text{diag}\,(g_1^{-1}, \dots , g_r^{-1}, 1, \dots , 1)$.

7.2.b. Cas d'une matrice carrée quelconque

Soit A une matrice carrée d'ordre p singulière, de décomposition spectrale PG^tQ. Nous définissons la D-g-inverse de A comme étant l'inverse de la matrice $B = PF^tQ$ où F est la solution du programme $(\mathcal{P}'')$.

Exemple : Revenons à l'exemple du paragraphe 6.1. La solution par la D-g-inverse est donnée par :

$$x = \frac{(3a-b)}{4} \qquad \text{et} \qquad y = \frac{(3b-a)}{4}$$

En particulier si $a = b$, nous avons $x = y = a/2$.

REMERCIEMENTS

Les auteurs tiennent à remercier M. Lenclud et M. Jaffray pour les remarques et les corrections qu'ils ont apportées à ce document.

RÉFÉRENCES

1. T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley & Sons, 1984.
2. J. Bouzitat, F. Fourgeaud et B. Lenclud, Inverses généralisées de matrices, définitions et propriétés, *Cahier du Groupe de Mathématiques Économiques*, 1980, Univ. Paris-I, 3.
3. P. G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, 1982.
4. R. R. Hocking, The Analysis and Selection of Variables in Linear Regression, *Biometrics, 82*, 1976, p. 1-49.
5. D. W. Marquardt, Generalized Inverses, Ridge Regression, Biased Linear Estimation and Nonlinear Estimation, *Technometrics, 12*, 1970, p. 591-612.
6. D. S. Moore, Generalized Inverse, Wald's Method, and the Construction of Chi Squared Tests of Fit, *Journal of the American Statistical Association, 72*, 1937, p. 131–137.
7. R. Penrose, A Generalized Inverse for Matrices, *Proceedings of the Cambridge Phylosophical Society, 51*, 1955, p. 406-413.
8. R. C. Rao, Separation Theorems for Singular Values of Matrices and their Applications in Multivariate Analysis, *Journal of Multivariate Analysis*, 1979, p. 362-377.