

Choleski et déterminant explicite de la matrice (Sibony page I 16-17

Pierre Lissy

April 30, 2010

On considère une matrice SDP A , et on veut l'écrire sous la forme $L^t L$. Démontrons le résultat par récurrence sur n . Pour $n = 1$ il n'y a rien à démontrer. Supposons le résultat vrai pour toute matrice de taille inférieure ou égale à $n-1$ et SDP. On écrit alors A par blocs sous la forme $\begin{pmatrix} B & {}^t v \\ v & a \end{pmatrix}$. On sait que la matrice B reste symétrique définie positive (c'est la matrice de la restriction à un hyperplan d'une forme quadratique déf pos, cela reste déf pos). On lui applique donc l'hypothèse de récurrence et on l'écrit sous la forme $L'^t L'$. Cherchons L sous la forme $L = \begin{pmatrix} L' & 0 \\ {}^t & b \end{pmatrix}$. Calculons $L^t L$, cela vaut $L = \begin{pmatrix} L' L'^t & L w \\ (L w)^t < w^t w + b^2 & L w \end{pmatrix}$. Pour avoir ${}^t L L = A$, il faut donc et il suffit d'avoir $L w = v$ et $w^t w + b^2 = a$, autrement dit $w = L^{-1} v$ et $b^2 = a - L^{-1} v (L^{-1} v)^t$. Pour choisir $b > 0$ qui vérifie cela, il est suffisant de montrer que $a - L^{-1} v (L^{-1} v)^t > 0$. Posons $x = \begin{pmatrix} b^{-1} v \\ -1 \end{pmatrix}$. Alors comme x est non nul et que A est définie positive on a $(Ax, x) > 0$, ie $\begin{pmatrix} B & {}^t v \\ v & a \end{pmatrix} b^{-1} v, b^{-1} v > 0$, ce qui donne exactement $a - L^{-1} v (L^{-1} v)^t > 0$. On remarque de plus que l'on obtient du coup par récurrence que les coefficients diagonaux de L sont strictement positifs.

Pour résoudre un système $Ax = b$ avec A déf positive, on se ramène à résoudre deux systèmes ce coup-ci triangulaires (donc facilement inversibles), à savoir ${}^t L x = y$ et $L y = b$. On aimerait donc pouvoir calculer explicitement L . Expliquons comment procéder. On a pour tout i, j , $a_{ij} = \sum_{k=1}^j l_{ik} l_{jk}$, avec $j \leq i$ (on déduit les autres par symétrie). Notamment, on a nécessairement $a_{11} = l_{11}^2$ et $a_{i1} = l_{i1} l_{11}$. Remarquons déjà que l'on sait que tous les coefficients diagonaux de A sont strictement positifs (ce sont les $(A \varepsilon_i, \varepsilon_i)$). Ainsi, on a $l_{11} = \sqrt{a_{11}}$, et $l_{i1} = a_{i1} / \sqrt{a_{11}}$. Montrons alors que l'on peut par récurrence construire L par colonnes. Supposons construites les colonnes $1, 2, \dots, k-1$. Cherchons à construire la colonne k . d'abord on a $a_{kk} = \sum_{j=1}^k l_{kj}^2 = l_{kk}^2 + \sum_{j=1}^{k-1} l_{kj}^2$, on en déduit $l_{kk} = \sqrt{a_{kk} - \sum_{j=1}^{k-1} l_{kj}^2}$. Enfin, on a $a_{ik} = \sum_{j=1}^k l_{ij} l_{kj} = l_{ik} l_{kk} + \sum_{j=1}^{k-1} l_{ij} l_{kj}$, et donc $l_{ik} = (a_{ik} - \sum_{j=1}^{k-1} l_{ij} l_{kj}) / l_{kk}$ pour peu que le l_{kk} choisi précédemment soit non nul, mais c'est forcément le cas puisqu'on a entre guillemets pas le choix et qu'on sait que la matrice L existe et est inversible (donc a ses coefficients diagonaux non nuls, et même sont strictement positifs comme démontré avant).

Finalement, la méthode de Cholevski est intéressante car on sait très facilement la programmer grâce aux calculs précédents. De plus, elle est d'ordre le même que Gauss (en fait, elle nécessite moins d'opérations, seulement $(m^3 + 6m^2 - 7m)/2$ alors que Gauss en nécessite $m(m^2 - 1)/3$, mais le gain n'est pas énorme).

References

[1]