Cours ANN 203 – Examen écrit, mardi 9 mai 2023

NOTE: Les matrices et les vecteurs considérés dans tous les exercices sont supposés à valeurs réelles. La notation $\|x\|$ désigne la norme 2 (euclidienne) du vecteur x.

Exercice E23-1 *Montrer que la matrice*

$$A = \begin{bmatrix} 9 & 6 \\ 6 & 5 \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

est définie positive, et calculer ses factorisations de Cholesky et LDL^T.

Solution: on a $\mathrm{Tr}(A) = \lambda_1 + \lambda_2 = 14 > 0$ et $\det(A) = \lambda_1\lambda_2 = 9 > 0$, donc $\lambda_1, \lambda_2 > 0$ et A est définie positive. De nombreuses méthodes permettent d'obtenir une factorisation symétrique de A. Par exemple, avec les notations du chapitre 2, la matrice d'élimination associée au pivot a_{11} est $G_1(a_{11}/\ell_1) = \left[\begin{smallmatrix} 1 & 0 \\ -2/3 & 1 \end{smallmatrix}\right]$, et on obtient alors

$$G_1(a_{11}/\ell_1)\, A\, G_1(a_{11}/\ell_1)^{\mathrm{T}} = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}$$

Puisque $G_1^{-1}(z) = G_1(-z)$ pour tout vecteur z, on obtient donc

$$A = LDL^{\mathrm{T}} \qquad \text{avec} \quad D = \begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 \\ 2/3 & 1 \end{bmatrix},$$

et la décomposition de Choleski de A est

$$A = GG^{\mathrm{T}} \qquad \text{avec} \quad G = LD^{1/2} = \begin{bmatrix} 1 & 0 \\ 2/3 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 2 & 1 \end{bmatrix}.$$

Exercice E23-2 (éléments propres de la modification de rang 1 d'une matrice diagonale) *Soit $A \in \mathbb{R}^{n\times n}$ de la forme $A = D + \rho zz^T$, où $D = diag(d_1 \dots d_n) \in \mathbb{R}^{n\times n}$ est telle que $d_1 > d_2 > \dots > d_n$, le vecteur $z \in \mathbb{R}^n$ a toutes ses composantes non nulles, et $\rho \neq 0$. Certains algorithmes de calcul de spectres de matrices utilisent la capacité de calculer le spectre d'une matrice A de la forme ci-dessus (modification de rang 1 d'une matrice diagonale), objet de cet exercice.*

(a) Soit λ une valeur propre de A et $q \in \mathbb{R}^n$ un vecteur propre associé. Montrer que $z^T q \neq 0$ et que $D - \lambda I$ est inversible.

(b) En déduire que λ est solution de l'équation

$$R(\lambda) := \rho z^T (D - \lambda I)^{-1} z + 1 = 0$$

(c) On suppose $\rho > 0$. En étudiant le sens de variation de $\lambda \mapsto R(\lambda)$, montrer que les racines $\lambda_1, \dots, \lambda_n$ sont distinctes et que l'on a

$$\lambda_1 > d_1 > \lambda_2 > d_2 > \dots \lambda_n > d_n.$$

(d) Que devient le résultat ci-dessus pour le cas $\rho < 0$?

(e) Proposer le principe d'une méthode de calcul numérique de $\lambda_1, \dots, \lambda_n$ utilisant les résultats ci-dessus.

Eléments de solution:

(a) Par hypothèse, q et λ vérifient $Aq - \lambda q = 0$, soit

$$(D - \lambda I)q + \rho(z^{\mathrm{T}} q)z = 0. \qquad (\star)$$

L'hypothèse $z^{\mathrm{T}} q = 0$ conduit à une contradiction: en effet, cela implique $(D - \lambda)q = 0$ et donc $q = q_0 e_\ell$, $\lambda = d_\ell$ pour un indice ℓ et $q_0 \in \mathbb{R}$. Le vecteur z ayant par hypothèse toutes ses composantes non nulles, cela entraîne $z^{\mathrm{T}} q = q_0 z_\ell \neq 0$, en contradiction avec l'hypothèse initiale. On doit donc avoir $z^{\mathrm{T}} q \neq 0$. Les composantes de z étant par hypothèse non nulles, l'équation $(\star)$ entraîne que $(D - \lambda)q$ ne peut pas avoir de composante nulle. D étant diagonale, il en résulte que $\lambda \neq d_i$ $(1 \leq i \leq n)$, et donc que $D - \lambda I$ est inversible.

(b) $D-\lambda I$ étant inversible, on peut multiplier à gauche l'équation $(\star)$ par $z^{\mathrm{T}}(D-\lambda I)^{-1}$; cela donne

$$(z^{\mathrm{T}}q)\big(\rho z^{\mathrm{T}}(D-\lambda I)^{-1}z+14\big)=(z^{\mathrm{T}}q)R(\lambda)=0,$$

et donc $R(\lambda)=0$ puisque $z^{\mathrm{T}}q\neq 0$.

(c) l'équation $R(\lambda)=0$ prend la forme plus explicite

$$\rho\sum_{i=1}^{n}\frac{z_i^2}{d_i-\lambda}+1=0,$$

qui montre que la fonction $\lambda\mapsto R(\lambda)$ est strictement croissante sur chacun des $n+1$ intervalles $(-\infty,d_n)$, (d_n,d_{n-1}),... (d_2,d_1), $(d_1,+\infty)$ et vérifie

$$\begin{aligned}R\big((-\infty,d_n)\big)&=(1,+\infty);\\ R\big((d_{i+1},d_i)\big)&=(-\infty,+\infty),\quad 1\le i\le n-1;\\ R\big((d_1,+\infty)\big)&=(-\infty,1).\end{aligned}$$

On déduit facilement de ce qui précède que $R(\lambda)$ a exactement n racines simples vérifiant les encadrements annoncés.

(d) Pour $\rho<0$, le raisonnement précédent se transpose facilement, tous les sens de variation étant inversés, et on trouve encore n racines simples vérifiant

$$\lambda_1>d_1>\lambda_2>\ldots\lambda_{n-1}>d_n>\lambda_n.$$

(e) Les résultats précédents suggèrent d'utiliser une méthode itérative de recherche de zéro, telle que la méthode de Newton, dans chaque intervalle contenant exactement un zéro simple. La suite des approximations λ_k d'un zéro est définie par la récurrence

$$(\lambda_{k+1}-\lambda_k)\,R'(\lambda_k)+R(\lambda_k)=0,$$

son initialisation λ_0 étant choisie dans un des intervalles pertinents; la dérivée $R'(\lambda)$ est donnée par

$$R'(\lambda)=\rho\sum_{i=1}^{n}\Big(\frac{z_i}{d_i-\lambda}\Big)^2.$$

Exercice E23-3 (actualisation de la SVD suite à modification de rang 1) *Soit $A\in\mathbb{R}^{m\times n}$ une matrice dont le rang r est strictement inférieur à m et n. Il est rappelé que A admet une décomposition en valeurs singulières (SVD) réduite de la forme*

$$A=U_rS_rV_r^T$$

où les colonnes de U_r et V_r sont les r vecteurs singuliers à gauche et à droite, respectivement, associés aux r valeurs singulières non nulles $\sigma_1\ge\ldots\ge\sigma_r>0$, et $S_r=diag(\sigma_1,\ldots,\sigma_r)\in\mathbb{R}^{r\times r}$.

On considère une perturbation A_1 de A de la forme

$$A_1=A+ab^T,\qquad a\in\mathbb{R}^m,\ b\in\mathbb{R}^n,$$

où les vecteurs a,b n'appartiennent pas aux sous-espaces $span(U_r)$, $span(V_r)$ respectivement engendrés par les colonnes de U_r et V_r. Il est important, pour diverses applications reposant sur des calculs en temps réel impliquant des matrices modifiées par acquisition répétée de données supplémentaires, de disposer d'algorithmes les plus économiques en temps de calcul permettant de déterminer la SVD de A_1 connaissant celle de A (actualisation de SVD). Cet exercice a pour objet d'établir la base de tels algorithmes.

(a) *On décompose a sous la forme*

$$a = a_r + \alpha a', \qquad \textit{avec} \quad a_r \in span(U_r), \quad a' \perp span(U_r),\ \|a'\| = 1.$$

Exprimer a', a_r, α en fonction de a et U_r. On décompose de la même manière $b = b_r + \beta b'$ avec $b_r \perp span(V_r)$.

(b) *Décomposer A_1 sous la forme*

$$A_1 = \big[U_r\ a'\big] S_1 \big[V_r\ b'\big]^T$$

Donner S_1, préciser ses dimensions, et montrer que S_1 est somme d'une matrice diagonale et d'une matrice de rang 1.

(c) *Proposer le principe d'une méthode permettant de déterminer la SVD de A_1, connaissant celle de A, en s'aidant de la décomposition ci-dessus.*

Eléments de solution:

(a) On a $a_r = UU^{\mathrm{T}}a$ (projection orthogonale de a sur les colonnes de U_r, qui définissent une famille orthonormée), et $\hat{a} := a - a_r$ est bien orthogonal à a $(U^{\mathrm{T}}(a - a_r) = 0)$. Ensuite, $|\hat{a}\|^2 = a^{\mathrm{T}}(I - UU^{\mathrm{T}})a > 0$, et on peut poser $\alpha := \sqrt{a^{\mathrm{T}}(I - UU^{\mathrm{T}})a}$. Ainsi,

$$a_r = UU^{\mathrm{T}}a, \qquad \alpha = \sqrt{a^{\mathrm{T}}(I - UU^{\mathrm{T}})a}, \qquad a' = (a - a_r)/\alpha \tag{$\star$}$$

conviennent, et on procède de même pour obtenir

$$b_r = VV^{\mathrm{T}}b, \qquad \beta = \sqrt{b^{\mathrm{T}}(I - VV^{\mathrm{T}})b}, \qquad b' = (b - b_r)/\beta. \tag{$\star\star$}$$

(b) Supposons que la décomposition $A_1 = \big[U_r\ a'\big] S_1 \big[V_r\ b'\big]^{\mathrm{T}}$ ait lieu. Puisque $\big[U_r\ a'\big]^{\mathrm{T}}\big[U_r\ a'\big] = I_{r+1}$ et $\big[V_r\ b'\big]^{\mathrm{T}}\big[V_r\ b'\big] = I_{r+1}$, S_1 est alors donnée par

$$S_1 = \big[U_r\ a'\big]^{\mathrm{T}} A_1 \big[V_r\ b'\big] = \begin{bmatrix} U_r^{\mathrm{T}}AV_r + (U_r^{\mathrm{T}}a)(b^{\mathrm{T}}V_r) & U_r^{\mathrm{T}}Ab' + (U_r^{\mathrm{T}}a)(b^{\mathrm{T}}b') \\ a'^{\mathrm{T}}AV_r + (a'^{\mathrm{T}}a)(b^{\mathrm{T}}V_r) & a'^{\mathrm{T}}Ab' + (a'^{\mathrm{T}}a)(b^{\mathrm{T}}b') \end{bmatrix}.$$

De plus, on a $A = U_r S V_r^{\mathrm{T}}$ par hypothèse sur A et les relations ($\star$) et ($\star\star$) impliquent

$$a'^{\mathrm{T}}a = \alpha, \quad b^{\mathrm{T}}b' = \beta, \quad a'^{\mathrm{T}}AV_r = 0, \quad U_r^{\mathrm{T}}Ab' = 0, \quad a'^{\mathrm{T}}Ab' = 0$$

et on obtient donc

$$S_1 = \begin{bmatrix} S + (U^{\mathrm{T}}a)(b^{\mathrm{T}}V) & \beta(U^{\mathrm{T}}a) \\ \alpha(b^{\mathrm{T}}V) & \alpha\beta \end{bmatrix} = \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} + \begin{Bmatrix} U_r^{\mathrm{T}}a \\ \alpha \end{Bmatrix} \begin{Bmatrix} V_r^{\mathrm{T}}b \\ \beta \end{Bmatrix}^{\mathrm{T}}$$

qui établit la décomposition annoncée comme somme d'une matrice diagonale et d'une matrice de rang 1.

(c) Supposons connue la SVD de $S_1 \in \mathbb{R}^{(r+1)\times(r+1)}$, qui est de la forme $S_1 = X\Sigma Y^{\mathrm{T}}$, où $X, Y \in \mathbb{R}^{(r+1)\times(r+1)}$ sont orthogonales et $\Sigma \in \mathbb{R}^{(r+1)\times(r+1)}$ est diagonale. On obtient

$$A_1 = \big(\big[U_r\ a'\big]X\big)\,\Sigma\,\big(Y^{\mathrm{T}}\big[V_r\ b'\big]^{\mathrm{T}}\big),$$

et on vérifie facilement que $\big(\big[U_r\ a'\big]X\big)^{\mathrm{T}}\big(\big[U_r\ a'\big]X\big) = I_{r+1}$ et $\big(\big[V_r\ b'\big]Y\big)^{\mathrm{T}}\big(\big[V_r\ b'\big]Y\big) = I_{r+1}$. La formule ci-dessus donne donc la SVD réduite de A_1, les valeurs singulières de A_1 étant portées par la diagonale principale de Σ.

La décomposition $S_1 = X\Sigma Y^{\mathrm{T}}$ peut être obtenue de façon économique par une adaptation de la méthode présentée dans l'exercice E23-2.

Exercice E23-4 (itérations d'Arnoldi pour une matrice tridiagonale non symétrique) *On considère une matrice $A \in \mathbb{R}^{n\times n}$ de la forme*

$$A = I + B, \qquad A \text{ inversible et } B \text{ antisymétrique} \quad (B^T = -B). \tag{$\star$}$$

Les itérations d'Arnoldi (à la base des formes efficaces de GMRES par exemple) consistent, sous leur forme complète, à trouver la matrice orthogonale $Q \in \mathbb{R}^{n\times n}$ et la matrice (a priori Hessenberg supérieure) $H \in \mathbb{R}^{n\times n}$ telles que

$$AQ = QH \tag{$\star\star$}$$

L'objet de cet exercice est d'appliquer les itérations d'Arnoldi à A ayant la forme particulière ($\star$):

(a) Montrer que, pour tout $x \in \mathbb{R}^n$, on a $(Ax, x) = \|x\|^2$.

(b) Montrer que H réalisant ($\star\star$) est tridiagonale de la forme

$$H = \begin{bmatrix} 1 & -\eta_2 & 0 & \dots & 0 \\ \eta_2 & 1 & -\eta_3 & \ddots & \vdots \\ 0 & & \dots & & 0 \\ \vdots & \ddots & -\eta_{n-1} & 1 & -\eta_n \\ 0 & \dots & 0 & \eta_n & 1 \end{bmatrix}$$

et donner la méthode permettant de déterminer $Q = [q_1, \dots, q_n]$ et les scalaires $\eta_2, \dots, \eta_n$ connaissant B.

Eléments de solution:

(a) On multiplie ($\star$) par x^{T} à gauche et par x à droite, ce qui donne $(Ax, x) = (Ax)^{\mathrm{T}}x = x^{\mathrm{T}}Ax = \|x\|^2 + x^{\mathrm{T}}Bx$. De plus, $x^{\mathrm{T}}Bx = (x^{\mathrm{T}}Bx)^{\mathrm{T}} = x^{\mathrm{T}}B^{\mathrm{T}}x = -x^{\mathrm{T}}Bx$ par antisymétrie de B, et donc $x^{\mathrm{T}}Bx = 0$. CEla prouve l'égalité $(Ax, x) = \|x\|^2$.

(b) On sait montrer l'existence de Q et H vérifiant ($\star\star$) pour A donnée, les itérations d'Arnoldi permettant le calcul pratique de ces matrices. Si A vérifie les conditions ($\star$), la multiplication à gauche de ($\star\star$) par Q^{T} donne

$$H = Q^{\mathrm{T}}AQ = Q^{\mathrm{T}}(I + B)Q = Q^{\mathrm{T}}Q + Q^{\mathrm{T}}BQ = I + Q^{\mathrm{T}}BQ$$

De plus, B étant par hypothèse antisymétrique, on a

$$H^{\mathrm{T}} = I + (Q^{\mathrm{T}}BQ)^{\mathrm{T}} = I - Q^{\mathrm{T}}BQ.$$

Par conséquent, $H - I$ est antisymétrique. Etant également Hessenberg supérieure, H est donc nécessairement tridiagonale. L'ensemble de ces propriétés implique qu'il existe des scalaires $\eta_2, \dots, \eta_n$ tels que H soit de la forme annoncée.

On peut retrouver ce résultat directement en particularisant les itérations d'Arnoldi (algorithme 4.3) à une matrice A vérifiant ($\star$).