Analyse Numérique
Corrigé Série 25

28 mai 2020

1. ($\star$, tout l'exercice)(**Différentiabilité des valeurs propres**)
Considérons la matrice

$$A(\epsilon) = \begin{pmatrix} 3\varepsilon & -1+\varepsilon & -1+4\varepsilon \\ \varepsilon & 1+5\varepsilon & -1+9\varepsilon \\ 2\varepsilon & 6\varepsilon & 2+5\varepsilon \end{pmatrix}.$$

(a) **(1 point)** Pour $A = A(0)$, calculer les valeurs propres et les vecteurs propres à gauche et à droite.
Sol.: *On considère la matrice*

$$A = \begin{pmatrix} 0 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Le polynôme caractéristique de cette matrice est donné par $p_\lambda = -\lambda(1-\lambda)(2-\lambda)$. Nous trouvons donc que les valeurs propres de A sont $\{0,1,2\}$. En résolvant un système à 3 inconnues nous pouvons trouver les vecteurs propres à droite associés aux valeurs propres de A, cela nous donne pour t non nul,

$$P_0^d = t\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad P_1^d = t\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \qquad P_2^d = t\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Pour les valeurs propres à gauche le même procédé pour A^T nous fournit

$$P_0^g = t\begin{pmatrix} 1 & 1 & 1 \end{pmatrix},\ P_1^g = t\begin{pmatrix} 0 & 1 & 1 \end{pmatrix},\ P_2^g = t\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}.$$

(b) **(1 point)** En utilisant le théorème de différentiabilité des valeurs propres montrer que, pour ϵ petit, la matrice $A(\epsilon)$ possède une valeur propre unique dans un voisinage de $\lambda = 2$. Déterminer $\lambda'(0)$ dans le développement de Taylor

$$\lambda(\epsilon) = \lambda + \lambda'(0)\epsilon + \mathcal{O}(\epsilon^2).$$

Sol.: *A présent considérons une perturbation de notre matrice*

$$A(\epsilon) = A + \epsilon C = \begin{pmatrix} 0 & -1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix} + \epsilon\begin{pmatrix} 3 & 1 & 4 \\ 1 & 5 & 9 \\ 2 & 6 & 5 \end{pmatrix}.$$

Le théorème de différentiabilité des valeurs propres nous garantit qu'il existe une unique valeur propre de $A(\epsilon)$ dans un voisinage de λ une valeur propre simple de A. Notons cette dernière $\lambda(\epsilon)$. Le théorème nous fournit de plus l'identité

$$\lambda(\epsilon) = \lambda + \epsilon \cdot \frac{u^T C v}{u^T v} + \mathcal{O}(\epsilon^2)$$

où u et v sont des vecteurs propres respectivement à gauche et à droite de A associés à λ.

Utilisons ce théorème pour la valeur propre $\lambda = 2$.
Ce que nous avons fait précédemment nous permet de calculer

$$\frac{u^T C v}{u^T v} = \frac{\begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 1 & 4 \\ 1 & 5 & 9 \\ 2 & 6 & 5 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}}{\begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}} = -1.$$

Donc

$$\lambda(\epsilon) = 2 - \epsilon + \mathcal{O}(\epsilon^2)$$

ce qui revient à dire que $\lambda'(0) = -1$.

2. (**La pseudo-inverse de Moore–Penrose**)
Soit une matrice $A \in \mathbb{R}^{m \times n}$, m, $n \in \mathbb{N}$, on verra en cours que A peut se mettre sous la forme $A = U\Sigma V^\top$ où $U \in \mathbb{R}^{m \times m}$ et $V \in \mathbb{R}^{n \times n}$ sont des matrices orthogonales et

$$\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \mathbf{0} \\ & & \sigma_r & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] \in \mathbb{R}^{m \times n},$$

où les $\mathbf{0}$ désignent des matrices zéros avec des tailles appropriées. L'entier r est le rang de A et les σ_i sont appelées les *valeurs singulières* et satisfont $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$. Cette décomposition généralise la diagonalisation lorsque $m \neq n$ et est appelée la *décomposition en valeurs singulières (SVD)* de A.
Dans cet exercice, on étudie la *pseudo-inverse de Moore–Penrose* $A^+ \in \mathbb{R}^{n \times m}$ définie comme

$$A^+ = V\Sigma^+ U^\top, \qquad \Sigma^+ = \left[\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & \mathbf{0} \\ & & \sigma_r^{-1} & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] \in \mathbb{R}^{n \times m}, \qquad r = \operatorname{rang}(A).$$

On va montrer les résultats suivants.

(a) Soit un vecteur $\boldsymbol{b} \in \mathbb{R}^m$. Si l'équation $A\boldsymbol{x} = \boldsymbol{b}$ n'a pas de solution, montrer que $\boldsymbol{x}^* = A^+\boldsymbol{b}$ est le vecteur qui minimise $\|A\boldsymbol{x} - \boldsymbol{b}\|_2$, c.-à-d., la distance entre $A\boldsymbol{x}^*$ et $\boldsymbol{b}$ par rapport à la norme euclidienne $\|\cdot\|_2$.
Indice : Utiliser les équations normales.
Sol.: *On a démontré en cours (chapitre 6) qu'un vecteur $\boldsymbol{x}^* \in \mathbb{R}^n$ qui minimise $\|A\boldsymbol{x} - \boldsymbol{b}\|_2$ résout aussi les équations normales, et vice-versa. Alors nous montrons que $\boldsymbol{x}^* = A^+\boldsymbol{b}$ résout $A^\top A\boldsymbol{x} = A^\top\boldsymbol{b}$ en utilisant $A = U\Sigma V^\top$ avec*

$$\Sigma = \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \mathbf{0} \\ & & \sigma_r & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] \in \mathbb{R}^{m \times n},$$

et la définition $A^+ = V\Sigma^+U^\top$ du pseudo-inverse de A avec

$$\Sigma^+ = \left[\begin{array}{ccc|c} \sigma_1^{-1} & & & \\ & \ddots & & \mathbf{0} \\ & & \sigma_r^{-1} & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] \in \mathbb{R}^{n \times m}.$$

En effet, cela nous donne

$$A^\top A\boldsymbol{x}^* = A^\top AA^+\boldsymbol{b} = V\Sigma^\top U^\top U\Sigma V^\top V\Sigma^+U^\top\boldsymbol{b} = V\Sigma^\top\Sigma\Sigma^+U^\top\boldsymbol{b} = V\Sigma^\top U^\top\boldsymbol{b} = A^\top\boldsymbol{b},$$

où nous avons utilisé le fait que

$$\Sigma^\top\Sigma\Sigma^+ = \Sigma^\top\left[\begin{array}{c|c} I_r & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0}\end{array}\right] = \Sigma^\top.$$

(b) Montrer que si $A \in R^{n\times n}$ est inversible, alors $A^+ = A^{-1}$.
Sol.: *Si $A = U\Sigma V^\top \in \mathbb{R}^{n\times n}$ est inversible, alors $\Sigma = U^\top AV \in \mathbb{R}^{n\times n}$ est inversible comme produit des matrices inversibles. Cela implique que toutes les valeurs singulières $\sigma_1, \ldots, \sigma_n$ sont non-nulles et*

$$\Sigma^{-1} = \begin{bmatrix}\sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_n^{-1}\end{bmatrix} \equiv \Sigma^+.$$

Alors on a $A^{-1} = V\Sigma^{-1}U^\top = V\Sigma^+U^\top = A^+$, par définition de A^+.

(c) Si l'équation $A\boldsymbol{x} = \boldsymbol{b}$ a une infinité de solutions, alors $\boldsymbol{x}^* = A^+\boldsymbol{b}$ est la solution la plus proche à $\mathbf{0}$. Démontrer cela avec les étapes suivantes :

i. Soient $\boldsymbol{v_1}, \ldots, \boldsymbol{v_r}$ les vecteurs singuliers à droite de A correspondants aux valeurs singulières non nulles $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ de A. Montrer que $\boldsymbol{x}^* \in \operatorname{span}\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_r}\}$.
Sol.: *Par définition de A^+ et de Σ^+ nous avons*

$$\begin{aligned}\boldsymbol{x}^* &= A^+\boldsymbol{b} = V\Sigma^+U^\top\boldsymbol{b} \\ &= [\boldsymbol{v_1}, \ldots, \boldsymbol{v_r}, \boldsymbol{v_{r+1}}, \ldots, \boldsymbol{v_n}]\left[\begin{array}{ccc|c}\sigma_1^{-1} & & & \\ & \ddots & & \mathbf{0} \\ & & \sigma_r^{-1} & \\ \hline & \mathbf{0} & & \mathbf{0}\end{array}\right]U^\top\boldsymbol{b} \\ &= [\sigma_1^{-1}\boldsymbol{v_1}, \ldots, \sigma_r^{-1}\boldsymbol{v_r}, \mathbf{0}, \ldots, \mathbf{0}]\,U^\top\boldsymbol{b} \in \operatorname{span}\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_r}\}.\end{aligned}$$

ii. Soit $\boldsymbol{x} \in \mathbb{R}^n$ une autre solution de $A\boldsymbol{x} = \boldsymbol{b}$. Montrer que $A^\top A(\boldsymbol{x} - \boldsymbol{x}^*) = \mathbf{0}$.
Sol.: *Si $\boldsymbol{x}$ est une autre solution de $A\boldsymbol{x} = \boldsymbol{b}$, alors ce vecteur résout aussi les équations normales $A^\top A\boldsymbol{x} = A^\top\boldsymbol{b}$. Puisque $\boldsymbol{x}^*$ résout les équations normales, alors en utilisant la linéarité on obtient $A^\top A(\boldsymbol{x} - \boldsymbol{x}^*) = A^\top A\boldsymbol{x} - A^\top A\boldsymbol{x}^* = A^\top\boldsymbol{b} - A^\top\boldsymbol{b} = \mathbf{0}$.*

iii. Montrer que la propriété ii. implique $\boldsymbol{x} - \boldsymbol{x}^* \in \operatorname{span}\{\boldsymbol{v_{r+1}}, \ldots, \boldsymbol{v_n}\}$.
Sol.: *En utilisant ii. nous obtenons $\mathbf{0} = (\boldsymbol{x} - \boldsymbol{x}^*)^\top A^\top A(\boldsymbol{x} - \boldsymbol{x}^*) = \|A(\boldsymbol{x} - \boldsymbol{x}^*)\|_2^2 \iff A(\boldsymbol{x}-\boldsymbol{x}^*) = \mathbf{0} \iff \boldsymbol{x}-\boldsymbol{x}^* \in \ker(A)$. D'après le corollaire 8.13, $\ker(A) = \operatorname{span}\{\boldsymbol{v_{r+1}}, \ldots, \boldsymbol{v_n}\}$ et donc $\boldsymbol{x} - \boldsymbol{x}^* \in \operatorname{span}\{\boldsymbol{v_{r+1}}, \ldots, \boldsymbol{v_n}\}$.*

iv. En utilisant les résultats des point i. et iii., montrer que $\boldsymbol{x}^*$ est orthogonal à $\boldsymbol{x} - \boldsymbol{x}^*$ et en déduire $\|\boldsymbol{x}\|_2 \geq \|\boldsymbol{x}^*\|_2$.
Sol.: *Comme $\boldsymbol{x}^* \in \operatorname{span}\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_r}\}$ (point i.) et $\boldsymbol{x} - \boldsymbol{x}^* \in \operatorname{span}\{\boldsymbol{v_{r+1}}, \ldots, \boldsymbol{v_n}\}$ (point iii.), on a $\langle\boldsymbol{x}^*, \boldsymbol{x} - \boldsymbol{x}*\rangle = 0$, car $\{\boldsymbol{v_1}, \ldots, \boldsymbol{v_n}\}$ est une base orthonormale de $\mathbb{R}^n$. Donc $\|\boldsymbol{x}\|_2^2 = \|\boldsymbol{x}^*\|_2^2 + \|\boldsymbol{x} - \boldsymbol{x}^*\|_2^2 \geq \|\boldsymbol{x}^*\|_2^2$.*