Licence de Mathématiques Fondamentales, L3

Examen d'analyse matricielle et modélisation

Soit $n \geq 1$, nous notons, pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, $M_n(\mathbb{K})$ l'ensemble de matrices carrées à éléments dans $\mathbb{K}$ de taille n, $U_n(\mathbb{C})$ celles qui sont unitaires, et enfin, les vecteurs de $\mathbb{K}^n$ sont des vecteurs colonne.

I (3 points) **Question de cours** : Enoncer et démontrer la décomposition en valeurs singulières (SVD) pour les matrices carrées inversibles.

Théorème. Soit $A \in M_n(\mathbb{K})$ une matrice carrée de taille n inversible. Il existe un couple de matrices unitaires, $U, V \in U_n(\mathbb{K})$ telles que UAV soit une matrice diagonale de nombres réels strictement positifs. (Ces nombres sont apellés les valeurs singulières de A.

Démonstration. La matrice A^*A est (toujours) une matrice hermitienne semi-définie positive, elle est donc normale, à valeurs propres réelles positives ou nulles. Puisque A est inversible, aucune de ses valeurs propres est nulle, A^*A est définie positive. Le théorème fondamental de diagonalisation unitaire des matrices normales, nous permet de démontrer qu'il existe une matrice unitaire $P \in U_n(\mathbb{C})$ telle que $A^*A = P\Delta P^*$ où Δ est une matrice diagonale, avec les valeurs propres de A^*A sur la diagonale. Ce sont donc des réels strictement positifs qui sont sur sur la diagonale. Notons δ la matrice diagonale, dont les éléments de la diagonale sont la racine carré de l'élément diagonal correspondant de Δ, ainsi $\delta^2 = \Delta$; ainsi $A^*A = P\delta\delta P^*$. D'où $A = ((A^{-1})^*P\delta)\delta P^*$. Il suffit maintenant de prouver que $(A^{-1})^*P\delta$ est unitaire. Puisqu'elle est inversible, il suffit de calculer $E = ((A^{-1})^*P\delta)^*(A^{-1})^*P\delta = \delta P^*A^{-1}(A^{-1})^*P\delta$. Or $A^{-1}(A^{-1})^* = A^{-1}(A^*)^{-1} = (A^*A)^{-1} = (P\Delta P^*)^{-1} = P\Delta^{-1}P^*$. Ainsi $E = \delta P^*P\Delta^{-1}P^*P\delta = \delta\Delta^{-1}\delta$, qui est l'identité. Ceci prouve que $U = (A^{-1})^*P\delta$ et $V = P^*$ sont unitaires et le théorème est démontré.

II (3 points) La décomposition de Choleski demandée est :

$$\begin{pmatrix} 1 & -1 & 0 & 1 \\ -1 & 10 & 3 & -1 \\ 0 & 3 & 5 & 2 \\ 1 & -1 & 2 & 16 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & 3 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 1 & 0 & 1 & \sqrt{14} \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & \sqrt{14} \end{pmatrix}$$

III - A (4 points)

Soient $A, B \in M_n(\mathbb{C})$, A normale ; notons $\lambda_1, \lambda_2, \ldots, \lambda_n$ ses valeurs propres, δ la matrice diagonale ayant ces valeurs propres sur la diagonale principale. Considérons μ une valeur propre de $A + B$.

1- Soit v un vecteur propre unitaire de la matrice $A + B$, associé à la valeur propre μ. Nous avons $Av + Bv = \mu v = \mu Iv$. Par ailleurs, puisque A est normale, elle est diagonalisable dans une base unitaire, c'est-à-dire, qu'il existe une matrice unitaire U telle que $A = U^*\delta U$. Ainsi, en multipliant par la droite par U, $UAv+UBv = \mu IUv$, or $UA = \delta U$, ainsi $\delta Uv+UBv = \mu IUv$, ou, $UBv = \mu IUv-\delta Uv = (\mu I-\delta)Uv$, comme demandé.

2- Su μ n'est pas une valeur propre de A, la matrice $\delta - \mu I$ est inversible, nous avons donc $Uv = (\mu I - \delta)^{-1}UBv$. Or la norme spectrale possède la propriété $||Uv||_2 = ||v||_2$ pour les matrices U unitaires, et elle est matricille, ainsi :

$$||v||_2 = ||Uv||_2 = ||(\mu I - \delta)^{-1}UBv||_2 \leq ||(\mu I - \delta)^{-1}||_2||UBv||_2 = ||(\mu I - \delta)^{-1}||_2||Bv||_2 \leq ||(\mu I - \delta)^{-1}||_2||B||_2||v||_2.$$

Puisque nous avons choissi v unitaire ceci reviens à $1 \leq ||(\mu I - \delta)^{-1}||_2||B||_2$.

3- La matrice $(\mu I - \delta)^{-1}$ est diagonale, les éléments de la diagonale sont les $\frac{1}{\mu-\lambda_i}$. Multipliée par son adjointe, c'est encore une matrice diagonale, dont les éléments diagonaux sont les $\frac{1}{|\mu-\lambda_i|^2}$. Ces éléments constituent son spectre, le rayon spectral, carré de la norme spectrale est le plus grand d'entre eux.

$$||(\mu I - \delta)^{-1}||_2^2 = \max_{1\leq i\leq n} \frac{1}{|\mu - \lambda_i|^2} = \frac{1}{\min_{1\leq i\leq n} |\mu - \lambda_i|^2} = \left(\frac{1}{|\mu - \lambda_{i_0}|}\right)^2$$

pour l'indice i_0 pour lequel le minimum est atteint, ainsi $||(\mu I - \delta)^{-1}||_2 = \frac{1}{|\mu-\lambda_{i_0}|}$ pour ce même indice.

En conclusion, si μ est l'une des valeurs propres λ_i, évidement $|\mu - \lambda_i| = 0 \leq ||B||_2$. Si μ n'est pas valeur propre de A, nous avons que $1 \leq \frac{1}{|\mu-\lambda_{i_0}|}||B||_2$; c'est-à-dire, $|\mu - \lambda_{i_0}| \leq ||B||_2$, pour cet indice.

4- En écrivant $A+\sigma_1 B = (A+\sigma B)+(\sigma_1-\sigma)B$, et puisque $A+\sigma B$ est supposée normale, le résultat ci-dessus montre que pour chaque valeur propre α de la matrice $A + \sigma_1 B$ il existe une valeur propre β de $A + \sigma B$ telle que $|\alpha - \beta| \leq ||(\sigma_1 - \sigma)B||_2 = |\sigma_1 - \sigma|||B||_2$.

Ainsi, en écartant la cas $B = 0$, trivial, si $\varepsilon > 0$ est donné, il suffit de prendre un $0 < \eta < \frac{\varepsilon}{||B||_2}$, pour avoir que si $|\sigma-\sigma_1| < \eta$, $|\alpha-\beta| < \varepsilon$. C'est le sens que l'on doit donner à la continuité de la dépendance des valeurs propres, en fonction de la matrice.

III - B (6 points)

Soient $A = (a_{ij}) \in M_n(\mathbb{C})$, D sa diagonale principale, $M(t) = tD + (1 - t)A$ pour $t \in [0, 1]$, et pour $i = 1, 2, \ldots, nD_i(t) = \{z \in \mathbb{C} \ / \ |z - a_{ii}| \leq (1 - t)\sum_{j=1,\ j\neq i}^n |a_{ij}|\}$. Où l'on remarque que $M(t)$ a les a_{ii} sur la diagonale, et les $(1 - t)a_{ij}$ hors de la diagonale, ainsi $D_i(t)$ est le disque de Gerschgorin-Hadamard de $M(t)$ correspondant à la ligne i.

Le fermé de $\mathbb{C}, W_i(t)$ est la réunion de tous ces disques sauf le i-ème.

1- Le théorème de Gerschgorin-Hadamard dit que les valeurs propres de chaque matrice sont contenues dans la réunion de ses disques. Or $D_1(t) \cup W_1(t)$ est la réunion de tous les disques de Gerschgorin-Hadamard de $M(t)$, il contient donc toutes les valeurs propres.

2- Puisque $M(1) = D = diag(a_{ii})$ ses valeurs propres sont les éléments de la diagonale principale de A.

3- Soit $t \in [0, 1]$, et i indice de ligne quelconque. Soit $z \in D_i(t)$ alors $|z - a_{ii}| \leq (1 - t)\sum_{j=1,\ j\neq i}^n |a_{ij}| \leq \sum_{j=1,\ j\neq i}^n |a_{ij}|$ car $t \geq 0$, ainsi $z \in D_i(0)$. Nous avons $D_i(t) \subset D_i(0)$.

4- Supposons que $D_1(0)$ ne contient aucune valeur propre de $M(t_1)$ pour $t_1 \in [0, 1]$ que l'on fixe. Toutes les valeurs propres de $M(t_1)$ sont donc à une distance de a_{ii} supérieure, strictement, à $\sum_{j=1,\ j\neq 1}^n |a_{1j}| =: a_1$. Soit $0 < \varepsilon < min\{|\beta_i - a_{11}| - a_1\}$ où le minimum porte sur les valeurs propres β_i de $M(t_1)$. Autrement dit, si β est une valeur propre de $M(t_1)$, $|\beta - a_{11}| - a_1 > \varepsilon$.

Pensons $M(t) = A + t(D - A) = (A + t_1(D - A)) + (t - t_1)(D - A)$ comme une perturbation par $(t - t_1)(D - A)$ de la matrice $A + t_1(D - A)$. Ainsi, par continuité, quelque soit $0 < \varepsilon$, il existe $0 < \eta$ tel que chaque valeur propre α de $M(t)$ est approchée à ε-près par une valeur propre de $M(t_1)$, β dès que $|t - t_1| < \eta$. C'est-à-dire $t \in [0, 1], |t - t_1| < \eta$, entraine que si α est une valeur propre de $M(t)$, il existe une valeur propre β de $M(t_1)$ qui vérifie $|\alpha - \beta| < \varepsilon$. On en déduit $|a_{11} - \alpha| \geq |a_{11} - \beta| - |\beta - \alpha| > a_1 + \varepsilon - |\beta - \alpha| > a_1$. Ce qui veut dire que α n'appartient pas au disque $D_1(0)$ comme il fallait démonter.

5- Soit $t_1 \in [0,1]$ tel que $D_1(0)$ contient au moins une valeur propre de $M(t_1)$, et que si $D_1(0) \cap W_1(0) = \emptyset$ la distance T, de a_{11} au compact $W-1(0)$ est strictement plus grande que a_1, le rayon de $D_1(0)$; posons $0 < \varepsilon < T - a_1$.

Par continuité, il existe $0 < \eta$, tel que si $t \in [0,1], |t - t_1| < \eta$ et si β une de ces valeurs propres de $M(t_1)$, il existe α valeur propre de $M(t)$ vérifiant $|\alpha - \beta| < \varepsilon$.

Maintenant si β une de ces valeurs propres de $M(t_1)$ qui est contenue dans $D_1(0)$, $|\beta - a_{11}| \leq a_1$. Proche de β, nous avons une valeur propre α comme ci-dessus. D'où $|a_{11} - \alpha| \leq |a_{11} - \beta| + |\beta - \alpha| < a_1 + \varepsilon < T$. La valeur propre α n'est pas dans $W_1(0)$ qui contient tous les autres disques $D_i(t), i > 1$. Puisqu'elle doit se trouver dans la réunion des disques $D_i(t)$, $\alpha \in D_1(t) \subset D_1(0)$.

6- Quitte à changer la numérotation, on peut supposer que $D_1(0)$, disque de Guerschgorin-Hadamard de A, ne rencontre pas la réunion des autres : $W_1(0)$. A l'aide des résultats ci-dessus, nous savons que l'ensemble $[0,1]$ est partagé en deux parties disjointes : l'ensemble des t pour lesquels l'une des valeurs propres de $M(t)$ est dans $D_1(0)$ et l'ensemble des t pour lesquels aucune des valeurs propres de $M(t)$ est élément de $D_1(0)$. L'exercice 4 montre que ce dernier ensemble est une partie ouverte de $[0,1]$. L'exercice 5 montre que le premier, son complémentaire, est aussi une partie ouverte. La connexité de $[0,1]$ entraine que l'une de ces parties est vide. Or en exercice 2, nous avons vu que l'élément 1 est dans l'ensemble des t pour lesquels l'une des valeurs propres de $M(1)$ est dans $D_1(0)$. Cette valeur propre est a_{11} le centre de $D_1(0)$. Alors, pour tout $t \in [0,1]$, $D_1(0)$ contient une valeur propre de $M(t)$ en particulier pour $t = 0$, $D_1(0)$ contient une valeur propre de A.

7- Supposons que la matrice $A \in M_n(\mathbb{C})$ a n disques de Gerschgorin-Hadamard deux-à-deux disjoints, en particulier, chaque disque est disjoint de la réunion des autres. Par 6- il contient une valuer propre de A. Les valeurs propres sont en nombre de n deux-à-deux distinctes. La matrice est diagonalisable.

La réciproque est fausse, par exemple la matrice unité, qui est bien diagonalisable, a tous ses disques de Guerschgorin-Hadamard confondus (et réduits à un point).

IV (5 points)

Soit $\mathcal{A} = \{A_1, A_2, \ldots, A_l\}$ un ensemble fini de matrices carrées de taille n à coéfficients réels. Fixons la suite infinie $(M_k)_k$ de matrices. vérifiant si $k \in \mathbb{N}$, il existe $1 \leq i \leq l$ tel que $M_k = A_i$. Notons Q_k le produit : $Q_k = M_k M_{k-1} \cdot \cdots \cdot M_1 M_0$ et $R_k = \frac{1}{||Q_k||_2} Q_k \in M_n(\mathbb{R})$. Qui est supposée converger, de limite R.

1- Par continuité de la norme, $1 = ||R_k||_2 \to ||R||_2$ Ce qui montre que $||R||_2 = 1$. En notant $\lambda_k(M) = ||MR_{k-1}||_2$, pour chaque $M \in M_n(\mathbb{R})$, clairement $0 \leq \lambda_k(M)$ puis $\lambda_k(M) = ||MR_{k-1}||_2 \leq ||M||_2 ||R_{k-1}||_2 = ||M||_2$. Les suites numériques $\lambda_k(M)$ sont toujours bornées.

2- Clairement, $M_k R_{k-1} = M_k \frac{1}{||Q_{k-1}||_2} Q_{k-1} = M_k \frac{M_{k-1} \cdots M_1 M_0}{||M_{k-1} \cdots M_1 M_0||_2} = \frac{M_k M_{k-1} \cdots M_1 M_0}{||M_k M_{k-1} \cdots M_1 M_0||_2} \frac{||M_k M_{k-1} \cdots M_1 M_0||_2}{||M_{k-1} \cdots M_1 M_0||_2} = ||M_k \frac{M_{k-1} \cdots M_1 M_0}{||M_{k-1} \cdots M_1 M_0||_2}||_2 R_k$ Ce qui donne $M_k R_{k-1} = ||M_k R_{k-1}||_2 R_k = \lambda_k(M_k) R_k$. Comme il fallait démontrer.

3- Si pour chaque i le nombre d'indices k pour lesquels la matrice M_k coincide avec la matrice A_i était fini, l'ensemble d'indices k pour les quels la matrice M_k coincide avec l'une des l matrices A_i serait fini, ce qui entraine que la suite matricielle serait finie ; Or nous considérons une suite infinie $(M_k)_k$. Soit donc i tel que la matrice A_i apparait pour un nombre infini d'indices $k \in \mathbb{N}$. Organisons ces indices en une suite extraite $(M_{k_p})_p$ de sorte, donc, que pour tout $p \in \mathbb{N}$, $M_{k_p} = A_i$ suite extraite constante. Attention, la suite R_{k_p} n'est pas pour autant constante ! La suite extraite correspondante $(\lambda_{k_p})_p$ est une suite réelle bornée, a donc une suite extraite convergente, notons $\lambda_{k_{p_l}} \to \lambda$ si $l \to +\infty$; le nombre λ est une valeur d'adhérence de la suite $(\lambda_k)_k$

L'identité prouvée en 2- s'écrit : $M_{k_{p_l}} R_{k_{p_l}-1} = \lambda_{k_{p_l}} R_{k_{p_l}}$. Or $M_{k_{p_l}} = A_i$, pour tout l, et tant $(R_{k_{p_l}-1})_l$ comme $(R_{k_{p_l}})_l$ sont des suites extraites de la suite $(R_k)_k$ qui est convergente, de limte R.

En passant donc, à la limite lorsque l tend vers l'infini, nous avons : $A_i R = \lambda R$.

4- La matrice R est de norme 1, elle a bien des colonnes non nulles. Soit X une colonne non nulle de R. Nous avons, produit par blocks, $A_i X = \lambda X$. Qui est bien un vecteur propre de A_i.

5- Pour $l = 1, n = 2, \mathcal{A} = \{A_1\}$ où $A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. On voit $M_k = A_1, Q_k = (A_1)^k = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}$ de norme de Frobenius $\sqrt{2 + k^2} \approx k$ qui n'est pas convergente.

La norme spectrale de $(A_1)^k$ est $\sqrt{1 + k^2 \frac{1 + \sqrt{1 + \frac{1}{k^2}}}{2}} \approx k$. La suite R_k converge et sa limite est $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. La valeur propre en cause, la seule est 1.

6- Pour que le produit infini, pour un choix $\mathbb{N} \to \{1, 2, \ldots, l\} : k \to i_k$,

$$\prod_{k=0}^{+\infty} A_{i_k}$$

soit convergent, il faut que la suite des produits partiels converge : $Q_l = \prod_{k=0}^{l} A_{i_k}$. Lorsque la limite est non nulle, ceci entraine la convergence de la suite $(R_l)_l$, normalisée comme ci-dessus. Or ceci entraine que l'on trouve une valeur propre réelle pour au moins l'une des matrices A_i ce qui a été exclu.

Un tel produit ne pet converger, sauf vers la matrice nulle éventuellement.