

Cours SIM 203 – Examen écrit, mardi 11 mai 2021

NOTE: Les matrices et les vecteurs considérés dans tous les exercices sont supposés à valeurs réelles.

Exercice 1. Calculer la décomposition en valeurs singulières (SVD) de la matrice

$$A = \begin{bmatrix} 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Exercice 2 (rotations de Givens). Les factorisations de la forme $A = QR$ interviennent (notamment) dans divers algorithmes de résolution de problèmes de moindres carrés ou de valeurs propres; le cours présente leur calcul par la méthode des réflecteurs de Householder, bien adaptée à la mise à zéro de toute une (portions de) colonne d'une matrice A .

Cet exercice aborde une autre technique également employée pour calculer une factorisation QR, appelée méthode des rotations de Givens; celle-ci est notamment utile quand il s'agit de mettre à zéro des termes "ciblés" de A . L'idée de base est la suivante: une fois choisis deux coefficients a, b situés dans une même colonne de $A \in \mathbb{R}^{m \times n}$, on applique une rotation plane au 2-vecteur de composantes a, b de sorte que la deuxième composante du 2-vecteur résultant soit nulle. On étend alors cette transformation de sorte qu'elle agisse sur deux lignes choisies de A et annule le terme de la deuxième ligne dans une colonne de A spécifiée.

- (a) On donne deux réels a, b . Montrer qu'il existe deux réels c, s vérifiant $c^2 + s^2 = 1$ tels que

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{Bmatrix} a \\ b \end{Bmatrix} = \begin{Bmatrix} r \\ 0 \end{Bmatrix}$$

(la valeur de r dans le vecteur résultat n'étant pas imposée *a priori*). Montrer que c, s peuvent être calculés, pour a, b donnés, au moyen de cinq opérations arithmétiques ($+, -, \times, /$) et un calcul de racine carrée. Proposer en particulier une méthode de calcul évitant le risque d'*overflow* qu'entraînerait une évaluation de a^2 ou b^2 quand l'un (au moins) de ces nombres est très grand.

- (b) Soit $A \in \mathbb{R}^{m \times n}$. On choisit deux lignes k, ℓ et une colonne p de A . Définir une matrice orthogonale $R(k, \ell, p)$ telle que $R(k, \ell, p)A$ laisse les lignes de A autres que k, ℓ inchangées et introduise un zéro à la position $a_{\ell p}$ de A . Quel est le nombre d'opérations arithmétiques total requis pour le calcul de $R(k, \ell, p)A$. La matrice $R(k, \ell, p)$ est connue sous le nom de matrice de rotation de Givens; interpréter $R(k, \ell, p)$ et expliquer cette appellation.

- (c) Soit $H \in \mathbb{R}^{n \times n}$ une matrice carrée de Hessenberg supérieure, c'est-à-dire de la forme

$$H = \begin{bmatrix} \times & \times & \times & \dots & \times \\ \times & \times & & & \times \\ 0 & \times & \times & & \times \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \times & \times \end{bmatrix}$$

(H vérifie donc $h_{ij} = 0$ pour $i - j > 1$). Décrire un algorithme fondé sur les rotations de Givens et réalisant une factorisation $H = QR$ de H , où $Q \in \mathbb{R}^{n \times n}$ est orthogonale et $R \in \mathbb{R}^{n \times n}$ est triangulaire supérieure.

L'algorithme évoqué en (c) est très utile pour le calcul de valeurs propres d'une matrice A par l'algorithme QR, après une factorisation préalable $A = PHP^T$ de A (P orthogonale, H Hessenberg supérieure).

Exercice 3 (itération d'Arnoldi interrompue). Soit $A \in \mathbb{R}^{n \times n}$ une matrice inversible. Les itérations d'Arnoldi permettent de calculer une matrice orthogonale $Q \in \mathbb{R}^{n \times n}$ et une matrice de Hessenberg $H \in \mathbb{R}^{n \times n}$ telles que $AQ = QH$ (section 4.5). On rappelle que la k -ème itération ($1 \leq k \leq n$) consiste à imposer que la k -ème colonne de l'égalité matricielle $AQ = QH$ soit vérifiée (l'égalité complète est donc vérifiée après la n -ème et dernière itération). On s'intéresse à la résolution d'un système linéaire $Ax = b$ pour un second membre $b \in \mathbb{R}^n$ donné, et démarre les itérations d'Arnoldi en posant $q_1 = b/\|b\|_2$. Pour tout entier k , on note $\mathcal{K}_k(A, b)$ le sous-espace de Krylov de dimension k associé à (A, b) . Cet exercice porte sur l'examen du cas particulier (non traité dans le cours) suivant:

- l'itération ℓ produit un coefficient sous-diagonal nul dans la ℓ -ème colonne de H : $h_{\ell+1, \ell} = 0$.

On suppose que ce numéro d'itération ℓ est le premier pour lequel cette situation apparaît. Dans tout cet exercice, ℓ désigne ce numéro particulier d'itération.

- Montrer que pour tout $k \leq \ell$ on a $\mathcal{K}_k(A, b) = \text{span}(q_1, \dots, q_k)$.
- Expliquer pourquoi l'itération ℓ ne permet pas de définir le vecteur $q_{\ell+1}$. Comment peut-on alors procéder pour effectuer l'itération $\ell+1$ et les itérations ultérieures?
- Montrer que le sous-espace de Krylov $\mathcal{K}_\ell(A, b)$ est stable par A : $A\mathcal{K}_\ell(A, b) \subseteq \mathcal{K}_\ell(A, b)$.
- Montrer que les sous-espaces de Krylov vérifient $\mathcal{K}_\ell(A, b) = \mathcal{K}_{\ell+1}(A, b) = \mathcal{K}_{\ell+2}(A, b) = \dots$ (stagnation à partir de l'itération ℓ).
- Montrer que la solution x de $Ax = b$ vérifie $x \in \mathcal{K}_\ell(A, b)$.
- Compte tenu des questions précédentes, quelles sont les conséquences du cas particulier considéré sur l'algorithme GMRES appliqué au système $Ax = b$?

Exercice 4 (preuve du théorème d'Eckart-Young-Mirsky). Soit $A \in \mathbb{R}^{m \times n}$ une matrice arbitraire, avec $m \geq n$. Soit $A = USV^T$ la SVD de A , avec $U = [u_1, \dots, u_m] \in \mathbb{R}^{m \times m}$, $V = [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$, les valeurs singulières de A étant ordonnées de sorte que $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$ (convention suivie dans le cours).

Pour un entier r tel que $1 \leq r < n$, on définit l'ensemble $\mathcal{M}_r \subset \mathbb{R}^{m \times n}$ des matrices $B \in \mathbb{R}^{m \times n}$ de rang r (donc, entre autres caractérisations, l'image de $B \in \mathcal{M}_r$ est un espace de dimension r). L'objet de cet exercice est de prouver le théorème d'Eckart-Young-Mirsky de meilleure approximation de A par une matrice de rang r pour la norme matricielle spectrale, qu'on peut énoncer ainsi:

$$(a) \min_{B \in \mathcal{M}_r} \|A - B\|_2 = \sigma_{r+1}, \quad (b) \hat{A}_r \in \arg \min_{B \in \mathcal{M}_r} \|A - B\|_2, \quad (\text{EYM})$$

la clause (b) signifiant donc que la SVD tronquée au rang r de A , notée \hat{A}_r (cf. Définition 6.5), minimise l'erreur en norme spectrale commise en approchant A par une matrice de rang $r < n$ et réalise donc la meilleure approximation de A par une matrice de rang r .

- Pourquoi le caractère réel de A implique-t'il que les matrices U, V contenant les vecteurs singuliers sont également réelles?
- Montrer que $\|A - \hat{A}_r\|_2 = \sigma_{r+1}$.
- Montrer que toute matrice $B \in \mathcal{M}_r$ peut se mettre sous la forme $B = XY^T$ avec $X \in \mathbb{R}^{m \times r}$ et $Y \in \mathbb{R}^{n \times r}$.
- Soit $B \in \mathcal{M}_r$ arbitrairement choisie. Montrer qu'il existe un vecteur non nul $w \in \text{span}(v_1, \dots, v_{r+1})$ tel que $Y^T w = 0$. Pour un tel w , montrer que $\|(A - B)w\|_2 \geq \sigma_{r+1} \|w\|_2$.
- Déduire des étapes précédentes la preuve de (EYM).

Cours SIM 203 – Corrigé de l'examen écrit du mardi 11 mai 2021

Exercice 1. On trouve sans difficulté

$$A^T A = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}, \quad A A^T = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

En classant les valeurs propres de $A^T A$ et $A A^T$ par ordre décroissant de module et avec les notations de la section 3.4 du cours, les éléments propres de ces deux matrices sont

$$\begin{aligned} \lambda_1 = 4, \lambda_2 = 0, \quad v_1 = e_2, v_2 = e_1 \quad (\text{pour } A^T A), \\ \lambda_1 = 4, \lambda_2 = \lambda_3 = 0, \quad u_1 = e_1, u_2 = e_2, u_3 = e_3 \quad (\text{pour } A A^T), \end{aligned}$$

On trouve alors immédiatement la SVD de A :

$$A = U S V^T, \quad U = [u_1, u_2, u_3] = I, \quad S = \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad V = [e_2, e_1] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Exercice 2.

- (a) Les conditions $c^2 + s^2 = 1$ et $sa + cb = 0$ sont par exemple vérifiées par

$$c = \frac{a}{\sqrt{a^2 + b^2}}, \quad s = -\frac{b}{\sqrt{a^2 + b^2}},$$

et ces formules sont évaluables en cinq opérations arithmétiques et une extraction de racine: a^2 , b^2 , $z := a^2 + b^2$, $z = \sqrt{z}$, $c = a/z$, $s = -b/z$.

Si a ou b est grand, l'évaluation intermédiaire de $a^2 + b^2$ présente un risque d'*overflow*, qui peut être évité par une normalisation judicieuse. Si $a \leq b$, on procède ainsi:

$$z = -a/b, \quad c = 1/\sqrt{1+z^2}, \quad s = z/\sqrt{1+z^2},$$

le cas $b \leq a$ étant traité de façon similaire. On vérifie aisément que la méthode (préférable) ci-dessus demande le même nombre d'opérations.

- (b) On définit $R = R(k, \ell, p) \in \mathbb{R}^{m \times m}$ comme la matrice identité en-dehors des lignes et colonnes k, ℓ , dont les intersections forment une matrice 2×2 de rotation:

$$R(k, \ell, p) = \begin{bmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & c & \dots & -s & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & s & \dots & c & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{bmatrix}$$

De plus, les coefficients c, s sont calculés comme en (a) pour $a = a_{kp}$ et $b = a_{\ell p}$. Le calcul de $R(k, \ell, p)A$ ne modifie que les lignes k, ℓ de A , et se résume aux évaluations de

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{Bmatrix} a_{kj} \\ a_{\ell j} \end{Bmatrix} \quad 1 \leq j \leq n$$

Chacun de ces produits matrice-vecteur demande 6 opérations arithmétiques (et la deuxième ligne s'annule par construction de $R(k, \ell, p)$ si $j = p$), donc le coût total est $6n - 3$ opérations.

(c) La factorisation QR de la matrice H demande uniquement l'annulation des $n - 1$ termes de la première sous-diagonale:

- Multiplier H à gauche par $R(1, 2, 1)H$ (avec c, s définis judicieusement) introduit un zéro en h_{21} , et demande $6n - 3$ opérations. La matrice $R(1, 2, 1)H$ obtenue est stockée en place (les coefficients initiaux de H étant remplacés par leurs valeurs modifiées par le calcul de $R(1, 2, 1)H$).
- Multiplier le résultat à gauche par $R(2, 3, 2)H$ (avec c, s définis judicieusement) introduit un zéro en h_{32} , et demande $6n - 9$ opérations (les zéros de la première colonne pouvant être ignorés). On vérifie sans difficulté que les zéros situés dans la première colonne de H sont préservés.
- [...] Multiplier le résultat à gauche par $R(k, k + 1, k)H$ défini judicieusement introduit un zéro en $h_{(k+1),k}$, et demande $6n - 6k + 3$ opérations (les zéros des $k - 1$ premières colonnes pouvant être ignorés).
- [...] Finalement, multiplier le résultat à gauche par $R(n - 1, n, n - 1)H$ introduit un zéro en $h_{n,(n-1)}$, et ne demande que 9 opérations

La suite d'opérations ci-dessus produit une matrice dont tous les termes de la première sous-diagonale sont maintenant nuls, préserve les zéros sous-diagonaux déjà présents dans la matrice initiale H , et conduit à

$$[R(n - 1, n, n - 1) \dots R(2, 3, 2) R(1, 2, 1)]H = R$$

où R est triangulaire supérieure. La matrice $R(n - 1, n, n - 1) \dots R(2, 3, 2) R(1, 2, 1)$ est orthogonale, et son inverse est $R(1, 2, 1)^T R(2, 3, 2)^T \dots R(n - 1, n, n - 1)^T =: Q$. Avec cette définition de Q , on a $H = QR$, et la factorisation souhaitée est obtenue. Le nombre total d'opérations arithmétiques nécessaire est

$$\sum_{k=1}^{n-1} (6n - 6k + 3) = 3(n^2 - 1)$$

Exercice 3.

(a) Le fait qu'on ait $\mathcal{K}_k(A, b) = \text{span}(q_1, \dots, q_k)$ pour tout $k \leq \ell$ se montre par récurrence sur k .

Les itérations d'Arnoldi étant initialisées par le choix $q_1 = b/\|b\|_2$, on a bien $\mathcal{K}_1(A, b) = \text{span}(b) = \text{span}(q_1)$. Supposons que $\mathcal{K}_k(A, b) = \text{span}(q_1, \dots, q_k)$ pour un certain entier k . L'itération k repose sur la vérification de la k -ème colonne de l'égalité matricielle $AQ = QH$, H étant une matrice de Hessenberg, ce qui se traduit (équation (4.15) du cours) par

$$Aq_k = h_{1k}q_1 + h_{2k}q_2 \dots + h_{(k+1),k}q_{k+1}. \quad (\text{A})$$

Pour $k < \ell$, $h_{(k+1),k} \neq 0$ par hypothèse, et on a donc

$$q_{k+1} = \frac{1}{h_{(k+1),k}} Aq_k - \frac{1}{h_{(k+1),k}} (h_{1k}q_1 + h_{2k}q_2 \dots + h_{kk}q_k)$$

Le vecteur entre parenthèses appartient à $\mathcal{K}_k(A, b)$ (hypothèse de récurrence), tandis que $Aq_k \in \mathcal{K}_{k+1}(A, b)$ (par définition de $\mathcal{K}_{k+1}(A, b)$ et puisque $q_k \in \mathcal{K}_k(A, b)$ par hypothèse de récurrence). Puisque les $\mathcal{K}_k(A, b)$ sont des espaces vectoriels emboîtés, l'égalité ci-dessus implique que $q_{k+1} \in \mathcal{K}_{k+1}(A, b)$, puis finalement que $\mathcal{K}_{k+1}(A, b) = \text{span}(q_1, \dots, q_{k+1})$.

(b) Le raisonnement ci-dessus ne fonctionne pas pour $k = \ell$ car $h_{(\ell+1),\ell} = 0$ par hypothèse: l'égalité (A) devient

$$Aq_\ell = h_{1\ell}q_1 + h_{2\ell}q_2 \dots + h_{\ell\ell}q_\ell. \quad (\text{A}_\ell)$$

et ne permet plus de définir le prochain vecteur $q_{\ell+1}$. On peut poursuivre en choisissant $q_{\ell+1}$ comme un vecteur arbitraire orthogonal à (q_1, \dots, q_ℓ) et normé, et poursuivre en appliquant (A) pour $k = \ell + 1, \ell + 2, \dots$ tant qu'on ne rencontre pas à nouveau une valeur nulle de $h_{(\ell'+1),\ell'}$ pour une itération $k = \ell'$ ultérieure.

(c) Pour $k = \ell$, (A_ℓ) entraîne que $Aq_\ell \in \mathcal{K}_\ell(A, b)$; d'autre part les vecteurs $Aq_1, \dots, Aq_{\ell-1}$ appartiennent à $A\mathcal{K}_{\ell-1}(A, b) \subset \mathcal{K}_\ell(A, b)$ par définition de $\mathcal{K}_\ell(A, b)$. Par conséquent, tout vecteur $x \in \mathcal{K}_\ell(A, b) = \text{span}(q_1, \dots, q_\ell)$ vérifie $Ax \in \mathcal{K}_\ell(A, b)$.

(d) Par définition de $\mathcal{K}_{\ell+1}(A, b)$, on a

$$\begin{aligned}\mathcal{K}_{\ell+1}(A, b) &= \text{span}(b, Ab, A^2b, \dots, A^\ell b) \\ &= \text{span}(b, A(b), A(Ab), \dots, A(A^{\ell-1}b)) \\ &= \text{span}(b) + A\mathcal{K}_\ell(A, b).\end{aligned}$$

Comme $\text{span}(b) \subset \mathcal{K}_\ell(A, b)$ et (par la question (c)) $A\mathcal{K}_\ell(A, b) \subset \mathcal{K}_\ell(A, b)$, on en déduit que $\mathcal{K}_{\ell+1}(A, b) \subset \mathcal{K}_\ell(A, b)$. Puisque par ailleurs les sous-espaces de Krylov sont emboîtés ($\mathcal{K}_k(A, b) \subset \mathcal{K}_{k+1}(A, b)$ pour tout entier k), on a l'égalité $\mathcal{K}_\ell(A, b) = \mathcal{K}_{\ell+1}(A, b)$. On peut alors poursuivre ce raisonnement par récurrence pour montrer que $\mathcal{K}_\ell(A, b) = \mathcal{K}_{\ell+1}(A, b) = \mathcal{K}_{\ell+2}(A, b) = \dots$

(e) On pose $x = Qy$, ce qui revient à développer x sur la base q_1, \dots, q_n : $x = y_1q_1 + \dots + y_nq_n$. On peut alors appliquer à A la réduction de Hessenberg $AQ = QH$, et le système $Ax = b$ prend la forme

$$Ax = b \implies AQy = b \implies QHy = b \implies Hy = Q^T b = \|b\|_2 q_1, \quad (\text{H})$$

la dernière égalité résultant du choix d'initialisation des itérations d'Arnoldi. Par ailleurs, le fait que $h_{(\ell+1),\ell} = 0$ entraîne que H est triangulaire par blocs:

$$H = \begin{bmatrix} H_{1:\ell, 1:\ell} & H_{1:\ell, (\ell+1):n} \\ 0 & H_{(\ell+1):n, (\ell+1):n} \end{bmatrix}.$$

Puisque de plus $b = \|b\|_2 q_1$, le système final (H) se résout par une méthode de remontée par blocs:

$$\begin{bmatrix} H_{1:\ell, 1:\ell} & H_{1:\ell, (\ell+1):n} \\ 0 & H_{(\ell+1):n, (\ell+1):n} \end{bmatrix} \begin{Bmatrix} y_{1:\ell} \\ y_{(\ell+1):n} \end{Bmatrix} = \begin{Bmatrix} \|b\|_2 q_1 \\ 0 \end{Bmatrix},$$

qui donne $y_{(\ell+1):n} = 0$ puis $H_{1:\ell, 1:\ell} y_{1:\ell} = \|b\|_2 q_1$. La solution x est donc dans $\text{span}(q_1, \dots, q_\ell) = \mathcal{K}_\ell(A, b)$.

(f) L'itération ℓ de l'algorithme GMRES résout le problème de moindres carrés $\|Ax - b\|_2^2 \rightarrow \min$ pour $x \in \mathcal{K}_\ell(A, b)$, et cette itération doit donc trouver la solution du système linéaire dans la situation considérée compte tenu des résultats précédents. GMRES a donc convergé et s'arrête à l'itération ℓ (poursuivre les itérations d'Arnoldi etc. au-delà est donc sans objet dans ce cadre).

Exercice 4.

(a) Les colonnes de U et V sont les vecteurs propres de AA^T et $A^T A$, respectivement. Les matrices AA^T et $A^T A$ étant SPD réelles, leurs vecteurs propres sont également réels.

(b) On a

$$A - \widehat{A}_r = \sum_{i=r+1}^n \sigma_i u_i v_i^T,$$

et cette formule constitue la SVD réduite de $A - \widehat{A}_r$ (les vecteurs manquants de U et V étant associés à la valeur singulière nulle de $A - \widehat{A}_r$). Par conséquent, la plus grande valeur singulière de $A - \widehat{A}_r$ est σ_{r+1} , et donc $\|A - \widehat{A}_r\|_2 = \sigma_{r+1}$ par le théorème 3.4.

- (c) Puisque $\text{Dim}(\text{Im}(B)) = r$ pour tout $B \in \mathcal{M}_r$, il existe r vecteurs x_1, \dots, x_r linéairement indépendants et nr coefficients y_{jk} ($1 \leq j \leq n, 1 \leq k \leq r$) tels que

$$b_j = \sum_{k=1}^r y_{jk} x_k, \quad 1 \leq j \leq n$$

c'est-à-dire $B = XY^T$ avec $X = [x_1, \dots, x_r]$ et $Y = [y_{jk}]$.

- (d) Posons $w = w_1 v_1 + \dots + w_{r+1} v_{r+1}$. L'équation $Y^T w = 0$ prend la forme du système matriciel homogène

$$\begin{bmatrix} y_1^T v_1 & \dots & y_1^T v_{r+1} \\ \vdots & & \vdots \\ y_r^T v_1 & \dots & y_r^T v_{r+1} \end{bmatrix} \begin{Bmatrix} w_1 \\ \vdots \\ w_{r+1} \end{Bmatrix} = \begin{Bmatrix} 0 \\ \vdots \\ 0 \end{Bmatrix}$$

de format $r \times (r+1)$, qui est sous-déterminé et a donc nécessairement une solution w non nulle.

Pour un tel w , on a $(A-B)w = Aw - XY^T w = Aw = \sum_{i=1}^{r+1} \sigma_i w_i u_i$, et donc (les vecteurs v_i étant orthonormaux)

$$\|(A-B)w\|_2^2 = \sum_{i=1}^{r+1} \sigma_i^2 w_i^2 \geq \sigma_{r+1}^2 \sum_{i=1}^{r+1} w_i^2 = \sigma_{r+1}^2 \|w\|_2^2$$

- (e) L'inégalité ci-dessus s'écrit $\|(A-B)w\|_2 / \|w\| \geq \sigma_{r+1}$, et entraîne $\|A-B\|_2 \geq \sigma_{r+1}$ pour tout $B \in \mathcal{M}_r$. Par ailleurs $\|A-B\|_2 = \sigma_{r+1}$ pour $B = \hat{A}_r \in \mathcal{M}_r$ (question (b)). Ces deux éléments combinés prouvent (EYM).