MATH-115(a) Algèbre linéaire avancée II
Printemps 2023
Question ouvertes de l'examen

July 6, 2023

Question 1

Considérons la forme bilinéaire $\langle \cdot, \cdot \rangle$ sur $\mathbb{R}^{n\times n}$ par $\langle A, B\rangle := \mathrm{Tr}(A^T B)$. Posons également le groupe orthogonal $\mathcal{O}(n) := \{S \in \mathbb{R}^{n\times n} : S^T S = I\}$.

1. (3 points) Montrer que $\langle \cdot, \cdot \rangle$ définit un produit scalaire de $\mathbb{R}^{n\times n}$ qui induit la norme de Frobenius $\|\cdot\|_F$.

2. (2 points) En déduire que la norme de Frobenius est invariante par action à gauche et à droite du groupe orthogonal. C'est-à-dire, montrer que

$$\|SA\|_F = \|AS\|_F = \|A\|_F \qquad \forall S \in \mathcal{O}(n).$$

   En particulier, on a $\|S\|_F^2 = n$ pour tout $S \in \mathcal{O}(n)$.

3. (1 point) En déduire que, pour toute matrice $A \in \mathbb{R}^{n\times n}$, minimiser $\|A - S\|_F^2$ sur $S \in \mathcal{O}(n)$ est équivalent à maximiser $\langle A, S\rangle$ sur $S \in \mathcal{O}(n)$.

4. (2 points) Montrer que, si $\Sigma \in \mathbb{R}_+^{n\times n}$ est diagonale à coefficients positifs ou nuls, alors l'identité $I \in \mathcal{O}(n)$ maximise la quantité $\langle \Sigma, S\rangle$, $S \in \mathcal{O}(n)$.

   Indice : montrer d'abord que $|S_{ii}| \leq 1$ pour $i = 1, \ldots, n$.

5. (2 points) Soit $A = U\Sigma V$ la décomposition en valeurs singulières d'une matrice $A \in \mathbb{R}^{n\times n}$ quelconque. Décrire les propriétés de U, Σ et V données par le théorème de décomposition.

   Conclure, à l'aide des questions précédentes, que UV est une des matrices orthogonales les plus proches de A. C'est-à-dire,

$$UV \in \arg\min_{S\in\mathcal{O}(n)} \|A - S\|_F\,.$$

Question 1 - solution

a)
- Le produit bilinéaire est symétrique : $\langle A, B\rangle = \mathrm{Tr}\,(A^T B) = \mathrm{Tr}\,((A^T B)^T) = \mathrm{Tr}\,(B^T A) = \langle B, A\rangle$.
- Le produit bilinéaire est défini positif : $\langle A, A\rangle = \mathrm{Tr}\,(A^T A) = \sum_{ij} A_{ij}^2 \geq 0$ avec égalité si et seulement si $A = 0$.
- Le produit bilinéaire induit la norme de Frobenius : $\langle A, A\rangle = \mathrm{Tr}\,(A^T A) = \sum_{ij} A_{ij}^2 = \|A\|_F^2$.

b)
- Pour tout $S \in \mathcal{O}(n)$, $\|SA\|_F^2 = \langle SA, SA\rangle = \mathrm{Tr}\,((SA)^T SA) = \mathrm{Tr}\,(A^T A) = \|A\|_F^2$, car $S^T S = 1$.
- Similairement, $\|AS\|_F^2 = \langle AS, AS\rangle = \mathrm{Tr}\,((AS)^T AS) = \mathrm{Tr}\,(A^T ASS^T) = \|A\|_F^2$, puisque $\mathrm{Tr}\,(AB) = \mathrm{Tr}\,(BA)$ et $SS^T = 1$.
- $\|S\|_F^2 = \|I\|_F^2 = n$.

c)
- On a $\|A - S\|_F^2 = \|A\|_F^2 + \|S\|_F^2 - 2\langle A, S\rangle$. Comme $\|S\|_F^2 = n$ est constant, minimiser à gauche est équivalent à maximiser $\langle A, S\rangle$.

d)
- Pour tout $S \in \mathcal{O}(n)$, $S^T S = 1$ implique que les colonnes de S sont orthonormals. En particulier, les colonnes sont de norme 1, et donc chaque composante de S est inférieurs à 1.
- $\langle \Sigma, S\rangle = \mathrm{Tr}\,(\Sigma^T S) = \sum_{ij} \Sigma_{ji} S_{ji} = \sum_i \Sigma_{ii} S_{ii} \leq \sum_i \Sigma_{ii} = \langle \Sigma, I_n\rangle$ d'après $S_{ii} \leq 1$.

e)
- U, V sont **réelles, orthogonales** dans $\mathbb{R}^{n\times n}$. $\Sigma \in \mathbb{R}^{n\times n}$ est **diagonale** avec diagonale **positive ou nulle**.
- $\|A - S\|_F^2 = \|U\Sigma V - S\|_F^2 = \left\|\Sigma - U^T S V^T\right\|_F^2$ par la question b).
- Ceci est minimisé lorsque $U^T S V^T = 1$, c'est-à-dire $S = UV$, par la question d).

Question 2

Soit K un corps, $A \in K^{n \times n}$ et $f(x, y) := x^T A y$ une forme bilinéaire symétrique non dégénérée de $V = K^n$.

1. (1 point) Montrer que $\ker(A) = \{0\}$, et donc que A est de rang plein.

Considérons désormais un sous-espace quelconque W de V de dimension m. Posons $w_1, \ldots, w_m$ une base de W. Définissons l'application $\phi : V \to K^m$ par $\phi(v)_i := f(v, w_i)$ pour tout $i = 1, \ldots, m$.

2. (3 points) Montrer que $\phi(v) = QAv$, pour une matrice $Q \in K^{m \times n}$ de rang m.

3. (2 points) Montrer que $\ker(\phi) = W^\perp := \{v \in V \mid f(v, w) = 0 \ \forall w \in W\}$, et que $\operatorname{rang} QA = m$.

4. (1 point) En déduire que $\dim(W) + \dim(W^\perp) = \dim(V)$.

5. (2 point) Montrer que $W \cap W^\perp = \{0\}$ est équivalent au fait que $f_{|W \times W}$ soit non dégénérée.

6. (1 point) Conclure que $V = W \oplus W^\perp$ si et seulement si $f_{|W \times W}$ est non dégénérée.

Question 2 - solution

a)
- $Ax = 0$ implique que $f(y, x) = 0$ pour tout y. Donc $x = 0$ car f est non-dégénérée.

b)
- $\phi(v) = \begin{pmatrix} w_1^T A v \\ \vdots \\ w_m^T A v \end{pmatrix} = \begin{pmatrix} w_1^T A \\ \vdots \\ w_m^T A \end{pmatrix} v = QAv$, où $Q = \begin{pmatrix} w_1^T \\ \vdots \\ w_m^T \end{pmatrix} \in \mathbb{R}^{m \times n}$.
- Comme les lignes de Q sont linéairement indépendantes, $\operatorname{rank}(Q) = m$.

c)
- Si $\phi(x) = 0$, alors $f(v, w_i) = 0$ pour tout i, et donc $f(v, w) = 0$ pour chaque $w \in W$ par linéarité.
- Si $f(v, w) = 0$ pour tout $w \in W$, alors, en particulier, $\phi(v) = 0$.
- Comme A est de rang plein, on a $\operatorname{rank}(QA) = \operatorname{rank}(Q) = m$.

d)
- Le théorème du rang donne $\dim(V) = \dim \ker \phi + \operatorname{rank}(QA) = \dim(W^\perp) + \dim(W)$, par b) et c).

e)
- Si $W \cap W^\perp = 0$ et $f(w, w') = 0$ pour certains $w \in W$ et chaque $w' \in W$, alors $w \in W^\perp$ et $w = 0$. Cela montre que la restriction de f est non dégénérée.
- Si la restriction de f est non dégénérée et $w \in W \cap W^\perp$, alors $f(w, w') = 0$ pour chaque $w' \in W$. Par conséquent, $w = 0$ et $W \cap W^\perp = 0$.

f)

- $W \cap W^\perp = 0$ et $\dim(V) = \dim(W^\perp) + \dim(W)$ implique que $V = W \oplus W^\perp$, puisque la concaténation des bases de W et $W^\perp$ donne une base de V (elle est linéairement indépendante et de taille n).

Question 3

Soit $A \in \mathbb{R}^{m \times n}$ une matrice non nulle. Posons $\sigma_1 \geq \cdots \geq \sigma_r > 0$ les valeurs singulières non nulles de A, et $\|\cdot\|_F$ la norme de Frobenius.

1. (1 point) Montrer que $r \geq 1$ (A admet au moins une valeur singulière non nulle).
2. (3 points) Montrer que, pour tout $k \in \{1, \ldots, r\}$, on a
$$\sqrt{k} \cdot \sigma_k \leq \|A\|_F \leq \sqrt{r} \cdot \sigma_1.$$
3. (2 points) En déduire qu'il existe une matrice $B \in \mathbb{R}^{m \times n}$ de rang inférieur ou égal à k telle que la plus grande valeur singulière de $(A - B)$ est bornée supérieurement par $\|A\|_F / \sqrt{k+1}$.

Question 3 - solution

a)
- Si aucune valeur propre non nulle n'existe, alors la décomposition en valeurs singulières $A = U\Sigma V$, avec $\Sigma = 0$, signifie que $A = 0$, ce qui est supposé ne pas être le cas.

b)
- Selon le cours, $\|A\|_F^2 = \sum_{i=1}^r \sigma_i^2$, et $\sigma_1 \geq \cdots \geq \sigma_r$.
- $\|A\|_F^2 \geq \sum_{i=1}^k \sigma_k^2 + \sum_{i=k+1}^r 0 = k\sigma_k^2$.
- $\|A\|_F^2 \leq \sum_{i=1}^r \sigma_1^2 = r\sigma_1^2$.

c)
- Soit $A = U\Sigma V$ la SVD de A. Nous prenons les k premières valeurs singulières pour construire B. $B = U \begin{pmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_k & \\ & & & 0 \end{pmatrix} V$, ce qui implique que $A - B$ a pour plus grande valeur singulière $\sigma_{k+1} \leq \|A\|_F / \sqrt{k+1}$, selon b).

Question 4

Soit $A \in \mathbb{R}^{n\times n}$ une matrice symétrique et soit

$$A = U \cdot \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} U^T$$

une factorisation de A telle que $U = (u_1, \ldots, u_n) \in \mathbb{R}^{n\times n}$ est orthogonale et $\lambda_1 \geq \cdots \geq \lambda_n$.

1. (2 points) Montrer que l'ensemble $\{u_1, \ldots, u_n\}$ forme une base orthonormale de vecteurs propres de A associés aux valeurs propres $\lambda_1, \ldots, \lambda_n$.
2. (4 points) Montrer que, pour tout $1 \leq \ell < n$, on a

$$\max_{\substack{x\in S^{n-1}\\ x\perp u_1,\ldots,x\perp u_\ell}} x^T Ax = \lambda_{\ell+1},$$

   et que $u_{\ell+1}$ est une solution optimale.

Question 4 - solution

a)
- On a $AU = DU$, ce qui donne $Au_i = \lambda_i u_i$ pour chaque i. De manière équivalente, $Au_i = UDU^{-1}u_i = UDe_i = \lambda_i Ue_i = \lambda_i u_i$.
- U est orthogonal, donc ses colonnes forment une base orthonormale. En particulier, $u_i^T u_j = \delta_{ij}$.

b)
- $x = \sum_{i=1}^n = \alpha_i u_i$, donc $Ax = \sum_{i=1}^n = \alpha_i Au_i = \sum_{i=1}^n = \alpha_i \lambda_i u_i$.
- Les colonnes de U sont orthonormales, donc $x^T Ax = \sum_i \alpha_i^2 \lambda_i$.
- Si x est orthogonal à $u_1, \ldots, u_\ell$, alors $\alpha_1 = \cdots = \alpha_\ell = 0$. Si $x \in S^{n-1}$, alors $\sum_i \alpha_i^2 = 1$. Donc $x^T Ax = \sum_i \alpha_i^2 \lambda_i \leq \lambda_{\ell+1} \sum_{i\geq \ell+1} \alpha_i^2 = \lambda_{\ell+1}$.

c)
- U est orthogonal, donc $u_{\ell+1}$ appartient à la sphère et est orthogonal à $u_1, \ldots, u_\ell$. $u_{\ell+1}^T Au_{\ell+1} = \lambda_{\ell+1}$.