Polytech Marseille, Département Informatique

TD d'analyse de données. 1

1. Soient A_1, A_2, B_1, B_2 quatre matrices carrées d'ordre n, et $A = [A_1, A_2]$ et $B' = [B_1', B_2']$ les matrices blocs associées. Quelles sont les dimensions des matrices A et B ? Peut-on définir leurs produits à droite et à gauche ? Peut-on donner une expression pour le produit AB ?

2. Si A est une matrice carrée d'ordre n telle que $A' = (a_1', \dots, a_n')$, où $a_1, \dots, a_n$ sont les lignes de A, comment s'écrivent les matrices $A'A$ et AA' ? Que peut-on dire du rang et des valeurs propres de ces deux matrices ?

3. Soit A une matrice rectangulaire (n, p) de rang plein. Montrez que si $p < n$, $A'A$ est une matrice carrée symétrique inversible, définie et positive. Que peut-on dire de ses valeurs propres ? Qu'appelle-t-on valeurs singulières de A ? Quelle est sa factorisation SVD ?

   Application : Soit $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$; Déterminez ses valeurs singulières et sa SVD.

   On peut définir la pseudo inverse B de A. Dans quel cas aura-t-on $B = (A'A)^{-1}A'$? Quelles sont les propriétés de la pseudo inverse ?

4. La matrice X contient des échantillons de m caractères d'individus. Pour chaque colonne, on a ainsi un échantillon d'ordre n indépendant. On note X_i la i-ème ligne de X et $X^{(j)}$ sa j-ème colonne. On note $\mu = \mathrm{E}(X_i)$, et $\hat{\mu} = (\widehat{\mu_1}, \dots, \widehat{\mu_m})$, avec $\widehat{\mu_j} = \frac{1}{n} \sum_{i=1}^{n} X_{i,j} = \frac{1}{n} 1_n X^{(j)}$

   Que représentent $\widehat{\mu_j}$ et $\hat{\mu}$?

   On note $1_{r,s}$ une matrice (r, s) formée de 1, et $1_s = 1_{1,s}$. Montrez que $\hat{\mu} = \frac{1}{n} 1_n X$

   Soit $Y = X - \frac{1}{n}\, {}'1_n \hat{\mu}$. On peut alors écrire $Y = AX$, avec $A = I_n - \frac{1}{n}\, {}'1_n 1_n$. Que représente Y ?

   Comment s'écrit la matrice A ? Vérifiez que $V = \frac{1}{n}\, {}'Y\, Y$ vérifie : $V = \frac{1}{n}\, {}'X\, X - {}'\hat{\mu}\, \hat{\mu}$

5. Rappelez la définition d'une distance. Qu'est-ce qu'une distance euclidienne ? Proposez des exemples de produits scalaires, et les distances euclidiennes associées. Proposez différentes distances pour des points de R^n.

6. Rappelez les propriétés d'un projecteur, d'un projecteur orthogonal, et suivant une matrice de projection M.

7. Comment s'écrit la fonction caractéristique d'une loi exponentielle ? En déduire son espérance, sa variance et tous ses moments.

8. On veut tester les valeurs d'un générateur aléatoire à l'aide d'un test de Kolmogorov-Smirnov. Comment s'écrit le critère du test ? quel est le seuil au risque de 5% (qui donne $c = 1.36$) pour une population de 10000 valeurs ?

9. Peut-on estimer les paramètres d'une loi uniforme avec le maximum de vraisemblance ?

10. Si A est une matrice carrée symétrique, et B une matrice rectangulaire (p, n), on définit pour un vecteur colonne X d'ordre n, $f(X) = X'AX$ et $g(X) = BX$. Quel est le gradient de f ? Pour minimiser f sous la contrainte $g(X) = b$, quelle(s) méthode(s) pouvez-vous proposer ?

Analyse de données : TD1

Exo 1

Matrice bloc → matrice de matrice

matrice

$$\begin{bmatrix} A \\ B \end{bmatrix}^T \cdot \begin{bmatrix} A' & B' \end{bmatrix}^T = \underbrace{AA' + BB'}_{\text{matrice}}$$

$A = [A_1, A_2]$ $A : n \times 2n$

$B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$ $B : 2n \times n$ } Donc AB et BA existent

Les transposé vanish toute ensemble

$$AB = A_1 B_1 + A_2 B_2$$

$$BA = \begin{bmatrix} B_1 A_1 & B_1 A_2 \\ B_2 A_1 & B_2 A_2 \end{bmatrix}$$

Exo 2

matrice carré d'ordre n scalaire

• $A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ $A^tA = \sum_{i=1}^{n} a_i^t a_i$ $AA^t = \begin{bmatrix} a_1 a_1^t & a_1 a_2^t & \cdots & a_1 a_n^t \\ a_2 a_1^t & & & \vdots \\ \vdots & & & \\ a_n a_1^t & \cdots & & a_n a_n^t \end{bmatrix}$

$A^t = [a_1^t, a_2^t, \dots, a_n^t]$

• $\dim(\mathrm{Im}) + \dim(\ker) = \dim(\mathbb{E})$ pour un endomorphisme

• $\mathrm{rang}(A) = \mathrm{rang}(A^t)$, $\mathrm{rang}(AA^t) \stackrel{?}{=} \mathrm{rang}(A^tA)$

$X \in \ker(AA^t)$ si $AA^tX = 0 \Leftrightarrow (A^TA)(A^TX) = 0$

A^TX equation $Y\,(A^TX) = 0$ $\to A^TX \in \ker(Y)$

$$A^TX \in \ker(A^TA)$$

* AA^T et A^TA ont les m̂ valeurs propres ≥ 0
* $[\mathrm{rang}(AA^t) = \mathrm{rang}(A^TA)] \Rightarrow \dim(\ker(AA^T)) = \dim(\ker(A^TA))$

Si λ est vp de AA^T :

$\exists X \neq 0$ tq $(AA^T)X = \lambda X$ — scalaire donc on déplace où veut le produit matriciel

$\Rightarrow \underbrace{(A^TA)}_{\text{nouveau}}(A^TX) = \lambda(A^TX) \Rightarrow$ bon si $A^TX \neq 0$

- si $A^TX \neq 0$: λ vp
- si $A^TX = 0$, $A\underbrace{A^TX}_{\text{nul}} = \lambda X = 0 \Rightarrow \lambda = 0$

    ↳ 0 est vp de A^TA : oui

Donc elles ont les m̂ vp

Pourquoi ≥ 0 ?

λ vp de AA^T :

$$\left.\begin{array}{l} \bullet\ AA^TX = \lambda X \Rightarrow X^TAA^TX = \lambda \underbrace{X^TX}_{>0} \\[2ex] \text{Si } \left.\begin{array}{l} Y = A^TX \\ Y^T = X^TA \end{array}\right\} X^TAA^TX = Y^TY \geq 0 \end{array}\right\} \lambda \geq 0$$

μ est valeur singulière de A

$\Updownarrow$

$\lambda = \mu^2$ est vp de AA^T (A^TA) respectivement, $\mu \geq 0$

Factorisation Matrice

Matrice $\rightarrow A = L \cdot U \leftarrow$ Matrice triang sup

(L : Matrice triang inf)

- $\rightarrow$ Si A symétrique : $L = U^T$ (cholesch)
- $\rightarrow$ Si A est diagonalisable :

$$A = QDQ^{-1}$$ (Q $\rightarrow$ Matrice de passage vers la base)

$$D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \qquad A^{-1} = QD^{-1}Q^{-1}$$

Diagonalisable $\Leftarrow$ polynome carac a des racines simples

→ Si A $n > p$ (resp $n < p$)

u et v unitaire

SVD ⟹ $A = \underbrace{U^t \Lambda V}_{\text{matrice de passage vers la base triangulaire}}$ $U : n \times n$, $V : p \times p$

$$A = \begin{bmatrix} \mu_1 & & & \\ & \mu_2 & 0 & \\ & & \ddots & \\ & & & \mu_n \end{bmatrix} : n \times p$$

Appli

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \qquad AA^T = \begin{bmatrix} 14 & 11 \\ 11 & 14 \end{bmatrix} \qquad A^TA = \begin{bmatrix} 5 & 8 & 5 \\ 8 & 13 & 9 \\ 5 & 9 & 10 \end{bmatrix}$$

Valeurs singulières de A : (Vp & $Vs(A)$ = Vp & $Vs(AA^T)$)

$$\begin{cases} (14-\lambda)x + 11y = 0 \\ 11x + (14-\lambda)y = 0 \end{cases}$$

ou

$$\begin{vmatrix} 14-x & 11 \\ 11 & 14-x \end{vmatrix} = (14-x)^2 - 11^2 = (14-x+11)(14-x-11)$$
$$= (25-x)(3-x)$$

Vp : 25 & 3

Vs : 5 et $\sqrt{3}$

$$\Lambda = \begin{bmatrix} 5 & 0 & 0 \\ 0 & \sqrt{3} & 0 \end{bmatrix} \begin{matrix} \to 25 \text{ 1ère colonne après} \\ \to 3 \text{ 2ème} \end{matrix}$$

Vecteurs propres de AA^T :

$AA^T \begin{pmatrix} x \\ y \end{pmatrix} = 25 \begin{pmatrix} x \\ y \end{pmatrix} \iff 14x + 11y = 25x \iff x = y$

$AA^T \begin{pmatrix} x \\ y \end{pmatrix} = 3 \begin{pmatrix} x \\ y \end{pmatrix} \iff 14x + 11y = 3x \iff -x = y$

$$U = \frac{1}{\sqrt{2}} \begin{array}{cc} x=y & x=-y \\ \left[\begin{array}{cc} 1 & 1 \\ 1 & -1 \end{array}\right] & \\ \to Vp \text{ du } 25 & \to Vp \text{ du } 3 \end{array}$$

$A = U^t \Lambda V = U \Lambda V^t$

$V_1^t = U^t \Lambda_1^{-1} A$ $\qquad \Lambda_1 = \begin{bmatrix} 5 & 0 \\ 0 & \sqrt{5} \end{bmatrix}$

le petit 1
matrice carré

$V = \begin{bmatrix} V_1 \\ \_\_ \end{bmatrix}$

↳ V_1 à laquelle on rajoute 1 ligne de norme 1 et orthogonale aux 2 précédantes

$$V_1^t = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1/5 & 0 \\ 0 & \frac{1}{\sqrt{3}} \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

$$= \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{5} + \frac{2}{\sqrt{3}} & \frac{2}{5} + \frac{3}{\sqrt{3}} & \frac{3}{5} + \frac{1}{\sqrt{3}} \\ \frac{1}{5} - \frac{2}{3} & \frac{2}{5} - \frac{3}{\sqrt{3}} & \frac{3}{5} - \frac{1}{\sqrt{3}} \end{bmatrix}$$

Exo 5

Distance dans E :

$d : E \times E \to \mathbb{R}^+$
$(x, y) \to d(x, y)$

- $d(x,y) = d(y,x)$
- $d(x,z) \leqslant d(x,y) + d(y,z)$
- $d(x,y) = 0 \Leftrightarrow x = y$

$\langle x, y \rangle$

$\|x\|_2 = \sqrt{\langle x, x \rangle}$

$d(x,y) = \|x - y\|$

* $\|\lambda x\| = |\lambda| \cdot \|x\|$
* $\|x + y\| \leqslant \|x\| + \|y\|$
* $\|x\| = 0 \Leftrightarrow x = 0$

Si $E = \mathbb{F}_2^n$

$\langle u, v \rangle = \sum_{k=1}^{n} u_k v_k \quad \in \{0, 1, \dots, n\}$

$\langle u, u \rangle = \sum_{k=1}^{n} u_k^2 = \sum_{k=1}^{n} u_k \quad (u_k \in \{0,1\})$ → nbre bit = 1

$d(u,v) = \|u - v\| = \sqrt{\text{nbre de bit à 1 dans } (u-v)} = \sqrt{d_h(u,v)}$

↳ = composante ≠ entre u et v

car 0 − 0 = 0
1 − 1 = 0

racine distance Hamming

Analyse de données : TD 1

Exo 7

- $f_X(x) = \lambda e^{-\lambda x} \quad (x \geq 0)$ $\qquad \varphi_X : \mathbb{R} \to \mathbb{C}$, $\quad u \mapsto E(e^{iuX})$

$$\varphi_X(u) = \int^{+\infty} \lambda e^{-\lambda x} e^{iux} dx$$

$$= \lambda \int_0^{+\infty} e^{(iu-\lambda)x} dx = \frac{\lambda}{iu-\lambda} \left[e^{\overbrace{(iu-\lambda)}^{<0}x} \right]_0^{+\infty} = -\frac{\lambda}{iu-\lambda}$$

- Moments de X d'ordre m

$$\int x^m f_X(x) dx$$

$$\varphi_X(u) = 1 + iuE(X) + \frac{(iu)^2}{2} E(X^2) + \dots + \frac{(iu)^m}{m!} E(X^m) + o(u^m)$$

DL en 0

$$\frac{\lambda}{\lambda - iu} = \frac{1}{1 - \frac{iu}{\lambda}} = \frac{1}{1-v} = 1 + v + v^2 + \dots + v^m$$

$$= 1 + \frac{iu}{\lambda} + \left(\frac{iu}{\lambda}\right)^2 + \dots + \left(\frac{iu}{\lambda}\right)^m + (u^m)$$

$$E(X) = \frac{1}{\lambda}, \quad E(X^2) = \frac{2}{\lambda^2}, \quad E(X^3) = \frac{3!}{\lambda^3}, \quad E(X^n) = \frac{n!}{\lambda^n}$$

$= E(X^2) - E(X)^2$

$$\Rightarrow V(X) = \frac{1}{\lambda^2}, \quad \sigma_X = \frac{1}{\lambda}$$

Exo 9

Estimation : Modèle paramétrique

$\theta = \lambda$: $\exp(\lambda)$

$\theta = (\mu, \sigma)$: $N(\mu, \sigma)$

$\theta = (a, b)$: $\text{unif}(a, b)$

On cherche à estimer la valeur de θ à partir d'un échantillon.

Modèle standard

$\boxed{F_\theta} \to x_1, \dots, x_n$

Maximum de vraisemblence

$$\underset{\theta}{\mathrm{Max}} \, (\underbrace{f_\theta(x_1) \times f_\theta(x_2) \times \dots \times f_\theta(x_n)}_{L(\theta, x)})$$

$L(\theta, x)$ Fonction de vraisemblence

$\ell(\theta, x) = \underset{\ln}{\log}(L)$ - log-vraisemblence

Modèle exponentiel

$\theta = \lambda \quad f_\theta(x) = \lambda e^{-\lambda x}, \; x \geq 0$

$$L(\lambda, x) = (\lambda e^{-\lambda x_1}) \dots (\lambda e^{-\lambda x_n})$$
$$= \lambda^n e^{-\lambda(x_1 + \dots + x_n)}$$

$\ell(\lambda, x) = n \ln(\lambda) - \lambda \left(\sum_1^n x_i \right)$ recherche min/max

$\ell' = \frac{n}{\lambda} - \sum_1^n x_i = 0$

$$\Rightarrow \lambda = \frac{n}{\sum_1^n x_i}$$

loi unif (a, b)

$f_\theta(x) = \frac{1}{b-a}, \; x \in [a, b]$

$L(\theta, x) = \frac{1}{(b-a)^n}$

$x_k \in [a, b]$

$a \leq x_1 \leq \dots \leq x_n \leq b$

comment maximiser $L(\theta, x)$?

$\begin{cases} a = x_1 \\ b = x_n \end{cases}$

Exo 10

A matrice $n \times n$ symétrique

$f(x) = X^t A X \qquad X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

ensemble de colonnes de A

$A = [C_1 C_2 \dots C_n] = \begin{bmatrix} l_1 \\ l_2 \\ \vdots \\ l_n \end{bmatrix} \qquad \frac{\partial f}{\partial x_1}(X) =$

$A = (a_{ij})$

ensemble des lignes de A

$f(x) = X^t Y \qquad Y = AX$

$Y = \begin{bmatrix} l_1 x \\ l_2 x \\ l_n x \end{bmatrix}$ $\to l_1$ de $A \times X$

$= x_1 y_1 + \dots + x_n y_n$

$y_k = l_k X = \sum_1^n a_{kj} x_j$

$$f(x) = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} x_j$$

$$= \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{k=1}^{n} \left[x_k \sum_{j \neq k} a_{kj} x_j \right]$$ $x_1 \sum \ldots + x_n \sum \ldots$

$\sum$ l'en reste qui ont a changé

$$\frac{\partial f}{\partial x_i} = 2 \quad a_{ii} x_i + \sum_{j \neq i} a_{ij} x_j + \sum_{k \neq i} x_k a_{kj}$$

i : indice ligne

$$= 2 \sum_{j=1}^{n} a_{ij} x_j$$ La dérivée par rapport à x_i est deux fois le produit de la ligne i par le vecteur colonne

$$\begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = 2 \begin{bmatrix} l_1 x \\ l_n x \end{bmatrix} = 2AX$$

Problème

$$\min_X (X^t A X)$$

cond nécessaire: $AX = 0$ $\qquad X \in \ker(A)$

cond suffisante: $f(x) \geq 0$

$X \in \ker(A) \Rightarrow f(x) \geq 0$

$A = CC^T \qquad f(x) = x^t C \, C^t x = (xC)^2$ ← passer

X est min global de f

$\min f(x) \quad \overbrace{\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n}^{g(x)} = c$

Si il y a plusieurs contraintes

Consigne $\begin{cases} g(x) = BX \\ g(x) = b \end{cases}$

$BX = b$ $\qquad$ avec p contraintes on a $B\ (p \times n)$

$$f(x) + \lambda [g(x) - c] = h_\lambda(x)$$

↳ multiplicateur de lagrange

On cherche à minimiser h_λ, où $\lambda \in \mathbb{R}$ est un paramètre

Si $\exists \lambda \in \mathbb{R}$ tq X est minimum de h_λ et $g(x) = c$

alors pour $g(y) = c$ $\qquad h_\lambda(x) \leq h_\lambda(y)$

$f(x) \leq f(y)$

$$\frac{\partial}{\partial x_i}\left(\lambda\left[g(x)-c\right]\right) = \lambda \alpha_i$$

$$\lambda \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \lambda \alpha$$

$$\begin{cases} 2AX + \lambda\alpha = 0 \Rightarrow 2X^tAX + \lambda C = 0 \\ \alpha^t X = C = X^t\alpha \quad \middle| \quad \lambda = -2\dfrac{X^tAX}{C} \end{cases}$$

Min f_*

$X^tX = 1$

$h_\lambda(X) = X^tAX + \lambda(X^tX - 1)$

$\text{grad}\, h_\lambda(X) = 2AX + \lambda 2X = 0$ ← perdu — a priori on définit comme on veut

$\Leftrightarrow AX = \lambda X \Rightarrow \underline{\lambda \text{ vp}}$ starfoullah

<u>CN</u> : λ vp de A

et X vecteur propre associé

(d'où $X^tAX = \lambda X^tX = \lambda$)

$\Rightarrow \lambda$ + petite vp de A