FONDEMENTS DU MACHINE LEARNING

2024/09/03

Programme :

- Introduction
- Notations, rappels d'algèbre linéaire

LOGISTIQUE

https://www.lamsade.dauphine.fr/~croyer/coursFML.html

clement.royer@lamsade.dauphine.fr

Emploi du temps

Mardi $13^H 45 - 15^H 15$

Du 03/09 au 15/10

Du 29/10 au 12/11

Notation : $N = 0{,}3 \underbrace{CC}_{\text{contrôle continu}} + 0{,}7 \underbrace{E}_{\text{Exam en janvier}}$

Sujet du cours : Utilisation de modèles linéaires pour des tâches d'apprentissage

→ Approche de base en sciences des données

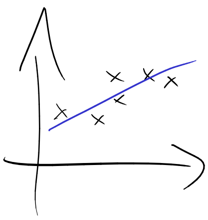
Plan du cours

① Réduction de dimension

≃ Apprentissage non supervisé

Outils : SVD, ACP, valeurs propres

Principe : Représenter une matrice de données en comprimant l'information



② Régression linéaire

≃ Apprentissage supervisé

Outils : Moindres carrés linéaires

Systèmes linéaires

NOTATIONS ET RAPPELS D'ALGÈBRE LINÉAIRE

<u>But</u> : Fixer les notations

$\mathbb{R}$: ensemble des nombres réels $(a, b, c, \alpha, \beta, \gamma, \dots)$

$\mathbb{R}^n$: ensemble des vecteurs à $n \geq 1$ composantes/coordonnées réelles

$$\mathbb{R}^1 \equiv \mathbb{R}$$

$$\forall v \in \mathbb{R}^n, \quad v = [v_i]_{i=1..n} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

$$v_i \in \mathbb{R} \ \forall i = 1..n$$

<u>Opérations sur les vecteurs</u> :

définition

$$\forall (u, v) \in (\mathbb{R}^n)^2, \quad u + v := [u_i + v_i]_{i=1..n}$$

$$\forall \alpha \in \mathbb{R}, \quad \alpha v := [\alpha v_i]_{i=1..n}$$

α scalaire

Produit scalaire :

$$\forall (u, v) \in (\mathbb{R}^n)^2, \quad u^T v := \sum_{i=1}^{n} u_i v_i$$

"u transpose v"

(Autres notations :
$<u, v>$
$u \cdot v$)

Norme (euclidienne) :

$$\forall u \in \mathbb{R}^n, \quad \|u\| := \sqrt{u^T u} = \sqrt{\sum_{i=1}^{n} u_i^2}$$

Matrices

$\mathbb{R}^{m\times n}$: ensemble des matrices à coefficients réels avec m lignes et n colonnes
$m \geq 1$ $\quad$ $n \geq 1$

$$A \in \mathbb{R}^{m\times n} \qquad A = [a_{ij}]_{\substack{i=1..m\\ j=1..n}}$$

$m = n$: matrice carrée

"vecteur ligne" $\quad$ "vecteur colonne"

$$\mathbb{R}^{1\times 1} \equiv \mathbb{R}$$

$$\mathbb{R}^{m\times 1} \equiv \mathbb{R}^m \qquad (\triangle\!\!\!! \ \mathbb{R}^{1\times n} \text{ n'est pas } \mathbb{R}^n)$$

$$A \in \mathbb{R}^{1\times n} \qquad A = [a_1 \ \ldots \ a_n]$$

$$a = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

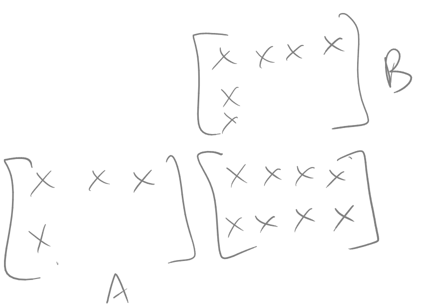
Opérations matricielles

- $\forall (A,B) \in (\mathbb{R}^{m\times n})^2, \qquad A + B := [a_{ij} + b_{ij}]_{ij}$

  $\forall \alpha \in \mathbb{R}, \qquad \alpha A := [\alpha a_{ij}]_{\substack{1\leq i\leq m\\ 1\leq j\leq n}}$

- $\forall A \in \mathbb{R}^{m\times n}, \ \forall B \in \mathbb{R}^{n\times p},$

$$AB := [c_{ik}]_{\substack{1\leq i\leq m\\ 1\leq k\leq p}} \in \mathbb{R}^{m\times p}$$



$$c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk}$$

) produit scalaire du vecteur représentant la i-ème ligne de A et de celui représentant la k-ième colonne de B

• Transposée d'une matrice :

$\forall A \in \mathbb{R}^{m\times n}$, $A = [a_{ij}]_{\substack{1\le i\le m \\ 1\le j\le n}}$, la transposée de A

est $\underbrace{A^T}_{\text{"A transpose"}} := [\tilde{a}_{ji}]_{\substack{1\le j\le n \\ 1\le i\le m}} \in \mathbb{R}^{n\times m}$, $\tilde{a}_{ji} = a_{ij} \ \forall\, i,j$

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 4 & 7 \end{bmatrix} \qquad A^T = \begin{bmatrix} 1 & 0 \\ 3 & 4 \\ 2 & 7 \end{bmatrix}$$

$\forall u \in \mathbb{R}^n$, $u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$, $\quad u^T = [u_1 \ \dots \ u_n] \in \mathbb{R}^{1\times n}$

$\forall (u,v) \in (\mathbb{R}^n)^2$, $\quad \underbrace{u^T v = \underbrace{(u^T)}_{\in \mathbb{R}^{1\times n}} \overbrace{v}^{\in \mathbb{R}^n \equiv \mathbb{R}^{n\times 1}}}_{\in \mathbb{R}^{1\times 1} \equiv \mathbb{R}}$

$$\underbrace{v}_{\in \mathbb{R}^{n\times 1}} \overbrace{u^T}^{\in \mathbb{R}^{1\times n}} \in \mathbb{R}^{n\times n}$$

$$u^T v \qquad [u_1 \ — \ u_n] \begin{bmatrix} v_1 \\ | \\ v_n \end{bmatrix} = \left[\sum u_i v_i\right] \qquad \begin{bmatrix} v_1 \\ | \\ v_n \end{bmatrix} [u_1 \ — \ u_n] = \begin{bmatrix} u_1 v_1 & \dots & u_n v_1 \\ u_1 v_n & \dots & u_n v_n \end{bmatrix}$$

Produit matrice-vecteur : $\forall A \in \mathbb{R}^{m\times n}$, $\forall v \in \mathbb{R}^n$,

$$Av := \left[\sum_{j=1}^{n} a_{ij} v_j\right]_{1\le i\le m} \in \mathbb{R}^m$$

$\forall B \in \mathbb{R}^{n\times n}$
$\forall v \in \mathbb{R}^n$

$$v^T B v = v^T (Bv) = (v^T B) v = (v^T) B (v)$$

$$:= \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} v_i v_j$$

$\forall A \in \mathbb{R}^{m\times n}$, $\forall u \in \mathbb{R}^m$, $\forall v \in \mathbb{R}^n$, $\quad u^T A v = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} u_i v_j$

Propriétés et classes de matrices

- $A \in \mathbb{R}^{n\times n}$ est symétrique si $A^T = A$

  Ex) $I_n = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}$

- $A \in \mathbb{R}^{n\times n}$ est semi-définie positive si
  1) $A^T = A$
  2) $\forall v \in \mathbb{R}^n$, $v^T A v \geq 0$

  Ex) I_n, $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$

  $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

  Notation : $A \succeq 0$

- $A \in \mathbb{R}^{n\times n}$ est définie positive si
  1) $A \succeq 0$
  2) $\forall v \in \mathbb{R}^n$, $\|v\| \neq 0$, $v^T A v > 0$

  $v \neq 0_{\mathbb{R}^n} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$

  Notation : $A \succ 0$

  Ex) $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, I_n

- $A \in \mathbb{R}^{n\times n}$ est inversible si $\exists B \in \mathbb{R}^{n\times n}$ telle que $AB = BA = I_n$

  B s'appelle l'inverse de A, et on la note A^{-1}

- Un réel $\lambda \in \mathbb{R}$ est une valeur propre de $A \in \mathbb{R}^{n\times n}$ si $\exists v \in \mathbb{R}^n$, $v \neq 0$, $Av = \lambda v$

  On dit alors que v est un vecteur propre associé à la valeur propre λ

- $A \in \mathbb{R}^{n \times n}$ est diagonalisable si il existe $(\lambda_1, \dots, \lambda_n) \in (\mathbb{R})^n$ et $P \in \mathbb{R}^{n \times n}$ inversible tels que

matrice diagonale

$$\underbrace{A = P \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} P^{-1}}_{\text{factorisation de la matrice } A \;/\; \text{décomposition spectrale / en valeurs propres}}$$

↳ L'existence d'une décomposition spectrale et des valeurs propres n'est pas garantie pour une matrice A quelconque

↳ En revanche, si $A \in \mathbb{R}^{n \times n}$ est symétrique, alors elle admet une décomposition spectrale de la forme

$$P \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} P^{-1} \quad \text{avec} \quad P^{-1} = P^T$$

(On dit que P est une matrice orthogonale)

Si en plus $A \succeq 0$ (resp. $A \succ 0$), alors

$\lambda_i \geq 0$ (resp. $\lambda_i > 0$). $\forall\, i = 1..n$

- Valeurs propres / Matrices semi-définies positives
  $\Rightarrow$ ACP (Analyse en Composantes Principales)

- Généralisation : valeurs singulières pour $A \in \mathbb{R}^{m \times n}$ $\forall m, n$
  $\Rightarrow$ SVD