

Rappels VIP : Algèbre linéaire et Analyse

Afshine AMIDI et Shervine AMIDI

6 octobre 2018

Notations générales

□ **Vecteur** – On note $x \in \mathbb{R}^n$ un vecteur à n entrées, où $x_i \in \mathbb{R}$ est la $i^{\text{ème}}$ entrée :

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

□ **Matrice** – On note $A \in \mathbb{R}^{m \times n}$ une matrice à m lignes et n colonnes, où $A_{i,j} \in \mathbb{R}$ est l'entrée située à la $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne :

$$A = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{pmatrix} \in \mathbb{R}^{m \times n}$$

Remarque : le vecteur x défini ci-dessus peut être vu comme une matrice $n \times 1$ et est aussi appelé vecteur colonne.

□ **Matrice identité** – La matrice identité $I \in \mathbb{R}^{n \times n}$ est une matrice carrée avec des 1 sur sa diagonale et des 0 partout ailleurs :

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Remarque : pour toute matrice $A \in \mathbb{R}^{n \times n}$, on a $A \times I = I \times A = A$.

□ **Matrice diagonale** – Une matrice diagonale $D \in \mathbb{R}^{n \times n}$ est une matrice carrée avec des valeurs non nulles sur sa diagonale et des zéros partout ailleurs.

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}$$

Remarque : on note aussi $D = \text{diag}(d_1, \dots, d_n)$.

Opérations matricielles

□ **Vecteur-vecteur** – Il y a deux types de multiplication vecteur-vecteur :

— Produit scalaire : pour $x, y \in \mathbb{R}^n$, on a :

$$x^T y = \sum_{i=1}^n x_i y_i \in \mathbb{R}$$

— Produit dyadique : pour $x \in \mathbb{R}^m, y \in \mathbb{R}^n$, on a :

$$xy^T = \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_n \\ \vdots & & \vdots \\ x_m y_1 & \cdots & x_m y_n \end{pmatrix} \in \mathbb{R}^{m \times n}$$

□ **Matrice-vecteur** – Le produit de la matrice $A \in \mathbb{R}^{m \times n}$ et du vecteur $x \in \mathbb{R}^n$ est un vecteur de taille \mathbb{R}^m , tel que :

$$Ax = \begin{pmatrix} a_{r,1}^T x \\ \vdots \\ a_{r,m}^T x \end{pmatrix} = \sum_{i=1}^n a_{c,i} x_i \in \mathbb{R}^m$$

où $a_{r,i}^T$ sont les vecteurs-ligne et $a_{c,j}$ sont les vecteurs-colonne de A et x_i sont les entrées de x .

□ **Matrice-matrice** – Le produit des matrices $A \in \mathbb{R}^{m \times n}$ et $B \in \mathbb{R}^{n \times p}$ est une matrice de taille $\mathbb{R}^{m \times p}$, tel que :

$$AB = \begin{pmatrix} a_{r,1}^T b_{c,1} & \cdots & a_{r,1}^T b_{c,p} \\ \vdots & & \vdots \\ a_{r,m}^T b_{c,1} & \cdots & a_{r,m}^T b_{c,p} \end{pmatrix} = \sum_{i=1}^n a_{c,i} b_{r,i}^T \in \mathbb{R}^{m \times p}$$

où $a_{r,i}^T, b_{r,i}^T$ sont des vecteurs-ligne et $a_{c,j}, b_{c,j}$ sont des vecteurs-colonne de A et B respectivement.

□ **Transposée** – La transposée est une matrice $A \in \mathbb{R}^{m \times n}$, notée A^T , qui est telle que ses entrées sont renversées.

$$\forall i, j, \quad A_{i,j}^T = A_{j,i}$$

Remarque : pour des matrices A, B , on a $(AB)^T = B^T A^T$.

□ **Inverse** – L'inverse d'une matrice carrée inversible A est notée A^{-1} et est l'unique matrice telle que :

$$AA^{-1} = A^{-1}A = I$$

Remarque : toutes les matrices carrées ne sont pas inversibles. Aussi, pour des matrices A, B , on a $(AB)^{-1} = B^{-1}A^{-1}$.

□ **Trace** – La trace d'une matrice carrée A , notée $\text{tr}(A)$, est la somme de ses entrées diagonales :

$$\text{tr}(A) = \sum_{i=1}^n A_{i,i}$$

Remarque : pour toutes matrices A, B , on a $\text{tr}(A^T) = \text{tr}(A)$ et $\text{tr}(AB) = \text{tr}(BA)$.

□ **Déterminant** – Le déterminant d'une matrice carrée $A \in \mathbb{R}^{n \times n}$ notée $|A|$ ou $\det(A)$ est exprimée récursivement en termes de $A_{\setminus i, \setminus j}$, qui est la matrice A sans sa $i^{\text{ème}}$ ligne et $j^{\text{ème}}$ colonne, de la manière suivante :

$$\det(A) = |A| = \sum_{j=1}^n (-1)^{i+j} A_{i,j} |A_{\setminus i, \setminus j}|$$

Remarque : A est inversible si et seulement si $|A| \neq 0$. Aussi, $|AB| = |A||B|$ et $|A^T| = |A|$.

Propriétés matricielles

□ **Décomposition symétrique** – Une matrice donnée A peut être exprimée en termes de ses parties symétrique et antisymétrique de la manière suivante :

$$A = \underbrace{\frac{A + A^T}{2}}_{\text{Symétrique}} + \underbrace{\frac{A - A^T}{2}}_{\text{Antisymétrique}}$$

□ **Norme** – Une norme est une fonction $N : V \rightarrow [0, +\infty[$ où V est un espace vectoriel, et tel que pour tous $x, y \in V$, on a :

- $N(x + y) \leq N(x) + N(y)$
- $N(ax) = |a|N(x)$ pour a scalaire
- si $N(x) = 0$, alors $x = 0$

Pour $x \in V$, les normes les plus utilisées sont récapitulées dans le tableau ci-dessous :

Norme	Notation	Définition	Cas
Manhattan, L^1	$\ x\ _1$	$\sum_{i=1}^n x_i $	LASSO
Euclidien, L^2	$\ x\ _2$	$\sqrt{\sum_{i=1}^n x_i^2}$	Ridge
p -norme, L^p	$\ x\ _p$	$\left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}}$	Inégalité de Hölder
Infini, L^∞	$\ x\ _\infty$	$\max_i x_i $	Convergence uniforme

□ **Dépendance linéaire** – Un ensemble de vecteurs est considéré comme étant linéairement dépendant si un des vecteurs de cet ensemble peut être défini comme une combinaison des autres.

Remarque : si aucun vecteur ne peut être noté de cette manière, alors les vecteurs sont dits linéairement indépendants.

□ **Rang d'une matrice** – Le rang d'une matrice donnée A est notée $\text{rang}(A)$ et est la dimension de l'espace vectoriel généré par ses colonnes. Ceci est équivalent au nombre maximum de colonnes indépendantes de A .

□ **Matrice semi-définie positive** – Une matrice $A \in \mathbb{R}^{n \times n}$ est semi-définie positive et est notée $A \succeq 0$ si l'on a :

$$A = A^T \quad \text{et} \quad \forall x \in \mathbb{R}^n, \quad x^T A x \geq 0$$

Remarque : de manière similaire, une matrice A est dite définie positive et est notée $A \succ 0$ si elle est semi-définie positive et que pour tout vector x non-nul, on a $x^T A x > 0$.

□ **Valeur propre, vecteur propre** – Étant donné une matrice $A \in \mathbb{R}^{n \times n}$, λ est une valeur propre de A s'il existe un vecteur $z \in \mathbb{R}^n \setminus \{0\}$, appelé vecteur propre, tel que :

$$Az = \lambda z$$

□ **Théorème spectral** – Soit $A \in \mathbb{R}^{n \times n}$. Si A est symétrique, alors A est diagonalisable par une matrice orthogonale réelle $U \in \mathbb{R}^{n \times n}$. En notant $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, on a :

$$\exists \Lambda \text{ diagonal, } A = U \Lambda U^T$$

□ **Décomposition en valeurs singulières** – Pour une matrice A de dimensions $m \times n$, la décomposition en valeurs singulières est une technique de factorisation qui garantit l'existence d'une matrice unitaire U $m \times m$, d'une matrice diagonale Σ $m \times n$ et d'une matrice unitaire V $n \times n$, tel que :

$$A = U \Sigma V^T$$

Analyse matricielle

□ **Gradient** – Soit $f : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ une fonction et $A \in \mathbb{R}^{m \times n}$ une matrice. Le gradient de f par rapport à A est une matrice de taille $m \times n$, notée $\nabla_A f(A)$, telle que :

$$\left(\nabla_A f(A)\right)_{i,j} = \frac{\partial f(A)}{\partial A_{i,j}}$$

Remarque : le gradient de f est seulement défini lorsque f est une fonction donnant un scalaire.

□ **Hessienne** – Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction et $x \in \mathbb{R}^n$ un vecteur. La hessienne de f par rapport à x est une matrice symétrique $n \times n$, notée $\nabla_x^2 f(x)$, telle que :

$$\left(\nabla_x^2 f(x)\right)_{i,j} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$$

Remarque : la hessienne de f est seulement définie lorsque f est une fonction qui donne un scalaire.

□ **Opérations de gradient** – Pour des matrices A, B, C , les propriétés de gradient suivants sont bons à savoir :

$$\nabla_A \text{tr}(AB) = B^T$$

$$\nabla_{A^T} f(A) = (\nabla_A f(A))^T$$

$$\nabla_A \text{tr}(ABA^T C) = CAB + C^T A B^T$$

$$\nabla_A |A| = |A| (A^{-1})^T$$