# Notes de cours : Mathématiques pour l'Intelligence Artificielle

Ton Nom

15 septembre 2025

## Table des matières


**1 LoRA : Low-Rank Adaptation** **1**

**2 Décomposition en valeurs singulières (SVD)** **1**
2.1 Définition . . . . . 1
2.2 Interprétation géométrique . . . . . 2

**3 Espaces fondamentaux d'une matrice** **2**

**4 Théorème spectral** **2**

**5 Caractérisations des matrices définies positives** **2**
5.1 Définie positive (PD) . . . . . 3
5.2 Définie semi-positive (PSD) . . . . . 3

**6 Déterminant et trace d'une matrice** **3**
6.1 Définition formelle du déterminant . . . . . 3
6.2 Trace . . . . . 4


## 1 LoRA : Low-Rank Adaptation

LoRA (Low-Rank Adaptation) est une méthode permettant d'adapter des grands modèles de langage en ne modifiant qu'un petit nombre de paramètres.

— L'idée est de décomposer une matrice de poids $W \in \mathbb{R}^{d\times d}$ en une perturbation de bas rang :

$$W' = W + \Delta W, \quad \Delta W = AB, \quad A \in \mathbb{R}^{d\times r},\ B \in \mathbb{R}^{r\times d},\ r \ll d.$$

— Seules les matrices $A$ et $B$ sont apprises lors de l'adaptation.

## 2 Décomposition en valeurs singulières (SVD)

### 2.1 Définition

Toute matrice $M \in \mathbb{R}^{m\times n}$ admet une décomposition en valeurs singulières :

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

avec :

— $U \in \mathbb{R}^{m\times m}$ matrice orthogonale,
— $V \in \mathbb{R}^{n\times n}$ matrice orthogonale,
— $\Sigma \in \mathbb{R}^{m\times n}$ matrice diagonale contenant les valeurs singulières $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$.

### 2.2 Interprétation géométrique

— La SVD décrit une matrice comme une suite de transformations : rotation, dilatation (par les valeurs singulières), puis rotation.
— Les valeurs singulières indiquent l'étirement maximal que la matrice applique à un vecteur.

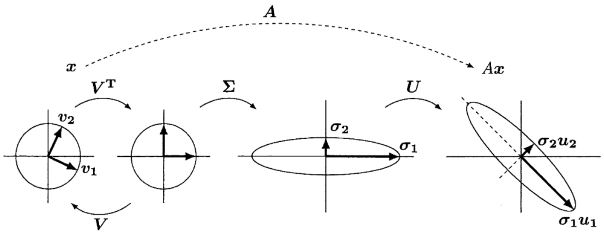


FIGURE 1 – Caption

## 3 Espaces fondamentaux d'une matrice

Soit $A \in \mathbb{R}^{m\times n}$. On définit les espaces suivants :

**Définition 3.1** (Noyau et image).

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n \mid Ax = 0\},$$
$$\mathcal{R}(A) = \{y \in \mathbb{R}^m \mid y = Ax,\ x \in \mathbb{R}^n\}.$$

**Définition 3.2** (Co-noyau et co-image).

$$\mathcal{N}(A^T) = \{y \in \mathbb{R}^m \mid A^T y = 0\},$$
$$\mathcal{R}(A^T) = \{x \in \mathbb{R}^n \mid x = A^T y,\ y \in \mathbb{R}^m\}.$$

**Propriété 3.1** (Orthogonalité). *Les quatre espaces sont reliés par les relations suivantes :*

$$\mathcal{R}(A) \perp \mathcal{N}(A^T), \qquad \mathcal{R}(A^T) \perp \mathcal{N}(A).$$

**Propriété 3.2** (Lien avec la SVD). *Soit $A = U\Sigma V^T$ une décomposition en valeurs singulières.*
— *Les colonnes de $U$ associées aux valeurs singulières non nulles forment une base de $\mathcal{R}(A)$.*
— *Les colonnes de $U$ associées aux valeurs singulières nulles forment une base de $\mathcal{N}(A^T)$ (le co-noyau).*
— *Les colonnes de $V$ associées aux valeurs singulières non nulles forment une base de $\mathcal{R}(A^T)$ (la co-image).*
— *Les colonnes de $V$ associées aux valeurs singulières nulles forment une base de $\mathcal{N}(A)$.*

## 4 Théorème spectral

**Théorème 4.1** (Théorème spectral). *Toute matrice symétrique réelle $A \in \mathbb{R}^{n\times n}$ est diagonalisable dans une base orthonormée :*

$$A = Q\Lambda Q^T,$$

*où $Q$ est orthogonale et $\Lambda$ diagonale contenant les valeurs propres réelles de $A$.*

## 5 Caractérisations des matrices définies positives

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

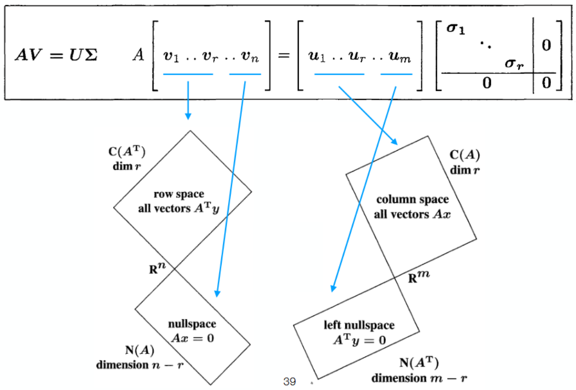


FIGURE 2 – Caption

### 5.1 Définie positive (PD)

**Propriété 5.1.** *Les conditions suivantes sont équivalentes :*

1. *Forme quadratique strictement positive :*

$$x^T Ax > 0, \quad \forall x \in \mathbb{R}^n \setminus \{0\}.$$

2. *Toutes les valeurs propres de $A$ sont strictement positives : $\lambda_i > 0,\ \forall i$.*
3. *$A$ admet une factorisation de Cholesky $A = LL^T$ avec $L$ triangulaire inférieure à diagonale strictement positive.*
4. *Critère de Sylvester : tous les mineurs principaux sont strictement positifs :*

$$\det(A_{1:k,1:k}) > 0, \quad \forall k = 1, \dots, n.$$

### 5.2 Définie semi-positive (PSD)

**Propriété 5.2.** *Les conditions suivantes sont équivalentes :*

1. *Forme quadratique non négative :*

$$x^T Ax \geq 0, \quad \forall x \in \mathbb{R}^n.$$

2. *Toutes les valeurs propres de $A$ sont non négatives : $\lambda_i \geq 0,\ \forall i$.*
3. *$A$ admet une factorisation (généralisée) $A = LL^T$ avec $L$ triangulaire inférieure (diagonales éventuellement nulles).*
4. *Critère de Sylvester généralisé : tous les mineurs principaux sont non négatifs :*

$$\det(A_{1:k,1:k}) \geq 0, \quad \forall k = 1, \dots, n.$$

## 6 Déterminant et trace d'une matrice

### 6.1 Définition formelle du déterminant

**Définition 6.1** (Déterminant)**.** *Soit $A = (a_{ij}) \in \mathbb{R}^{n \times n}$. Le **déterminant** de $A$ est défini par :*

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)},$$

*où :*

— $S_n$ *désigne le groupe symétrique, i.e. l'ensemble des permutations de* $\{1, 2, \dots, n\}$,
— $\mathrm{sgn}(\sigma)$ *est la signature de la permutation* $\sigma$, *qui vaut* $+1$ *si* $\sigma$ *est paire, et* $-1$ *si* $\sigma$ *est impaire.*

**Propriété 6.1.** *Quelques propriétés fondamentales du déterminant :*
— $\det(AB) = \det(A)\det(B)$,
— $\det(A^T) = \det(A)$,
— $\det(I_n) = 1$,
— $A$ *est inversible* $\iff \det(A) \neq 0$,
— *si* $A$ *est triangulaire (supérieure ou inférieure), alors* $\det(A)$ *est le produit des éléments diagonaux.*

**Exemple 6.1.** *Pour* $n = 2$ *:*

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \mathrm{sgn}(\mathrm{id})(ad) + \mathrm{sgn}((12))(bc) = ad - bc.$$

—

## 6.2 Trace

**Définition 6.2** (Trace)**.** *La* ***trace*** *d'une matrice carrée* $A \in \mathbb{R}^{n \times n}$ *est définie comme la somme de ses éléments diagonaux :*

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

**Propriété 6.2.** *Propriétés fondamentales de la trace :*
— $\mathrm{Tr}(A + B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$,
— $\mathrm{Tr}(\alpha A) = \alpha\,\mathrm{Tr}(A)$, *pour tout scalaire* $\alpha$,
— $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$, *si* $A$ *et* $B$ *sont de tailles compatibles,*
— *si* $A$ *est symétrique et diagonalisable* $A = Q\Lambda Q^T$, *alors*

$$\mathrm{Tr}(A) = \sum_{i=1}^{n} \lambda_i,$$

*c'est-à-dire la somme des valeurs propres.*

**Exemple 6.2.**

$$\mathrm{Tr} \begin{pmatrix} 4 & 2 \\ 3 & 4 \end{pmatrix} = 4 + 4 = 8.$$