MIA Séance 2 : Algèbre linéaire

EVEN Lwena

https://thomas.schatz.cogserver.net/pdf/teaching/M2_IAAA/MIA/2022-2023/CM2.pdf

2 Matrices

Rappels :

— Produit matrice - vecteur : soient $M = (m_{ij})_{1\leq i\leq n, 1\leq j\leq m}$ et $v = (v_i)_{1\leq i\leq m}$.

$$Mv = \begin{bmatrix} \sum\limits_{j=1}^{m} m_{1j} v_j \\ \vdots \\ \sum\limits_{j=1}^{m} m_{nj} v_j \end{bmatrix}$$

— Produit matrice - matrice : soient $A \in \mathbb{R}^{m\times n}$ et $B \in \mathbb{R}^{n\times l}$.

$$C = AB = (c_{ij})_{1\leq i\leq m, 1\leq j\leq l} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

2.1 Matrice d'une application linéaire

La matrice d'une application linéaire permet de représenter cette application en dimensions finies d'une manière unifiée, permissible aux outils de calcul matriciel.

Méthode pour calculer la matrice d'une application linéaire :

E, F des espaces vectoriels de dimension finie.
$V = (v_1, ..., v_p)$ base de E
$W = (w_1, ..., w_n)$ base de F
$f : E \to F$, fonction linéaire

1. Calculer l'image des vecteurs de la base de l'espace de départ par la fonction d'intérêt $f(v_j)$
2. Effectuer la décomposition linéaire de ces images à l'aide des vecteurs de la base de l'espace d'arrivée $f(v_j) = \lambda_1 w_1 + ... + \lambda_n w_n$
3. Les coefficients λ_i de la décomposition des $f(v_j)$ par rapport aux vecteurs de la base de l'espace d'arrivée forment les colonnes de la matrice de l'application linéaire

L'espace vectoriel $\mathbb{R}_d[X]$ définit formellement les polynômes de degré d, indépendamment du type d'objet. Y sont définis l'addition ainsi que la multiplication par un scalaire.
Par exemple : $1 + X = (1, 1, 0, ...)$ ou encore $X + 3X^2 = (0, 1, 3, 0, ...)$

Exemple. Construire la matrice associée à l'application linéaire suivante

$$\begin{array}{cccl} f & : & \mathbb{R}_2[X] & \to & \mathbb{R}_3[X] \\ & & P = a_0 + a_1 X + a_2 X^2 & \mapsto & f(P) = (a_2 + a_0)X^3 + a_1 X^2 + a_0 \end{array}$$

1. $f(1) = f((1, 0, 0)) = X^3 + 1$
   $f(X) = f((0, 1, 0)) = X^2$
   $f(X^2) = f((0, 0, 1)) = X^3$
2. $f(1) = 1 \times w_1 + 1 \times w_4$
   $f(X) = 1 \times w_3$
   $f(X^2) = 1 \times w_4$
3. $M(f, V, W) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

Note : La multiplication matricielle est imposée par les propriétés des applications linéaires. En effet, en prenant A et B deux matrices associées à deux applications linéaires et C leur composition, alors définir la multiplication telle que $C = \text{mult}(A, B)$ mène nécessairement à la formule connue du produit matriciel.

2.2 Changement de base

$$P(V, W) = P(V \to W) = M(Id_E, V, W)$$

Changer de base revient à changer de système de coordonnées. La matrice de passage décrite ci-dessus transforme un vecteur v exprimé en terme de ses coordonnées dans V, en le même vecteur en terme de ses coordonnées dans W.

Exemple. *Calculer la matrice de passage de $B = ((1,0),(0,1))$ la base canonique dans l'espace euclidien $\mathbb{R}^2$ vers $C = ((1,1),(-1,1))$.*
Exprimons les vecteurs de la base B en fonction des vecteurs de la base C, ce qui revient à trouver λ_1, λ_2 tels que :
$(1,0) = \lambda_1(1,1) + \lambda_2(-1,1) \iff \lambda_1 = \frac{1}{2}, \lambda_2 = -\frac{1}{2}$
$(0,1) = \lambda_1(1,1) + \lambda_2(-1,1) \iff \lambda_1 = \frac{1}{2}, \lambda_2 = \frac{1}{2}$
$P(B \to C) = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$

2.3 Propriétés et opérations matricielles

Quelques propriétés :

- $A \times (B \times C) = (A \times B) \times C$
- $A \times (B + C) = A \times B + A \times C$
- $(A + B) \times C = A \times C + B \times C$
- $c(A \times B) = (cA) \times B = A \times (cB)$
- ⚠ $AB \neq BA$ en général
- $(A + B)^T = A^T + B^T$
- $(\alpha A)^T = \alpha A^T$
- ⚠ $(AB)^T = B^T A^T$

Matrices par bloc.
Soient deux matrices partitionnées $P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$ et $Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$. On peut alors utiliser les opérations usuelles :

- Addition : $P + Q = \begin{bmatrix} P_{11} + Q_{11} & P_{12} + Q_{12} \\ P_{21} + Q_{21} & P_{22} + Q_{22} \end{bmatrix}$
- Multiplication par un scalaire : $\alpha P = \begin{bmatrix} \alpha P_{11} & \alpha P_{12} \\ \alpha P_{21} & \alpha P_{22} \end{bmatrix}$
- Multiplication de matrices : $P \times Q = \begin{bmatrix} P_{11}Q_{11} + P_{12}Q_{21} & P_{11}Q_{12} + P_{12}Q_{22} \\ P_{21}Q_{11} + P_{22}Q_{21} & P_{21}Q_{12} + P_{22}Q_{22} \end{bmatrix}$
  ⚠ Attention aux dimensions des blocs qui doivent permettre tous les produits.

Interprétation des opérations matricielles.

- Produit matrice-vecteur : $y = Ax = x_1 a_1 + ... + x_n a_n$ (avec $a_1, .., a_n$ les colonnes de A)
- Produit matrice-matrice : $c_{ij} = \tilde{a}_i^T b_j = \langle \tilde{a}_i, b_j \rangle = \sum_{k=1}^{n} a_{ik} b_{kj}$

  ou $C = \sum_i a_i \tilde{b}_i^T = \begin{bmatrix} \\ a_1 \\ \\ \end{bmatrix} \begin{bmatrix} & \tilde{b}_1^T & \end{bmatrix} + \begin{bmatrix} \\ a_2 \\ \\ \end{bmatrix} \begin{bmatrix} & \tilde{b}_2^T & \end{bmatrix} + ...$ qui représente une somme de N matrices de rang au plus 1.

3 Angles et orthogonalité

Produit scalaire.

$$\begin{array}{rccl} (\,|\,) & : & E \times E & \to & \mathbb{R} \\ & & (v, w) & \mapsto & (v|w) \end{array}$$

Propriétés de la forme :

1. Bilinéaire : linéaire en v pour w fixé et en w pour v fixé
2. Symétrique : $\forall v, w \in E^2, (v|w) = (w|v)$
3. Définie positive : $\forall v \in E \setminus \{0_E\}, (v|v) > 0$

Lorsqu'il est associé à la norme Euclidienne : $(v|w) = v^T w = \sum_{i=1}^{n} v_i w_i$

Rappels. $\cos\theta = \frac{(v|w)}{\sqrt{(v|v)(w|w)}}$. Deux vecteurs sont **orthogonaux** si et seulement si $(v|w) = 0$.

Algorithme de Gram-Schmidt. Il permet de construire une base orthonormale.

Gram-Schmidt procedure : étant donné k vecteurs indépendants $(a_1, a_2, \ldots, a_k) \in E^k$, trouver des vecteurs orthonormaux $(q_1, q_2, \ldots, a_k)$, tels que pour tout $r \leq k$, $(q_1, q_2, \ldots, q_r)$ soit une base orthonormale de $\mathrm{Vect}(a_1, a_2, \ldots, a_r)$

- step 1a. $\tilde{q}_1 := a_1$
- step 1b. $q_1 := \tilde{q}_1 / \|\tilde{q}_1\|$ (normalize)
- step 2a. $\tilde{q}_2 := a_2 - (q_1^T a_2) q_1$ (remove q_1 component from a_2)
- step 2b. $q_2 := \tilde{q}_2 / \|\tilde{q}_2\|$ (normalize)
- step 3a. $\tilde{q}_3 := a_3 - (q_1^T a_3) q_1 - (q_2^T a_3) q_2$ (remove q_1, q_2 components)
- step 3b. $q_3 := \tilde{q}_3 / \|\tilde{q}_3\|$ (normalize)
- etc.

Factorisation QR. Q est constituée des vecteurs de base orthonormale, et R est une matrice triangulaire supérieure dont les coefficients sont calculés durant Gram-Schmidt. A doit être une matrice à coefficients réels de rang k (à colonnes indépendantes).

$$\begin{bmatrix} a_1 & a_2 & \ldots & a_k \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & \ldots & q_k \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \ldots & r_{1k} \\ 0 & r_{22} & \ldots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & r_{kk} \end{bmatrix}$$

Matrice orthogonale. Une matrice est orthogonale si ses colonnes forment une base orthonormale. Dans le cas de matrices rectangulaires, on parle de matrices semi-orthogonales. Soit Q une matrice orthogonale de taille $m \times k$, alors $Q^T Q = I_k$ car $q_i^T q_j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{sinon.} \end{cases}$.

L'intérêt d'une base orthonormée est qu'elle permet de calculer la décomposition de n'importe quel vecteur comme $v = \sum_{i=1}^{n} (v|v_i) v_i$. Dans le cas général, pour inverser une matrice, on passe alors de $\mathcal{O}(n^3)$ à $\mathcal{O}(n)$ car pour une matrice orthogonale, $Q^{-1} = Q^T$ car $Q^T Q = I = Q Q^T$

4 Structure des applications linéaires et matrices

Comment comprendre une application linéaire grâce à sa représentation matricielle?

4.1 Décomposition en valeurs singulières

Soient $A \in \mathbb{R}^{m \times n}$
$U \in \mathbb{R}^{m \times m}$ orthogonale telle que $U^T U = U U^T = I_m$
$V \in \mathbb{R}_{n \times n}$ orthogonale telle que $V^T V = V V^T = I_n$
$\Sigma \in \mathbb{R}^{m \times n}$ diagonale avec $\sigma_1, \ldots, \sigma_r \in (\mathbb{R}^{+*})^r$ ordonnés tels que $\sigma_1 \geq \ldots \geq \sigma_r > 0$ où r représente le rang de A, c'est-à-dire le nombre de colonnes linéairement indépendantes de A.

Alors, pour A quelconque, $\exists AV = U\Sigma \iff A = U\Sigma V^T$.

$$\begin{bmatrix} Av_1 & \dots & Av_n \end{bmatrix} = \begin{bmatrix} U \begin{bmatrix} \sigma_1 \\ 0 \\ \vdots \end{bmatrix} & \dots & U \begin{bmatrix} \vdots \\ \sigma_r \\ \vdots \end{bmatrix} & U \begin{bmatrix} 0 \\ 0 \\ \vdots \end{bmatrix} \dots \end{bmatrix} = \begin{bmatrix} \sigma_1 u_1 & \dots & \sigma_r u_r & 0 & \dots \end{bmatrix}$$

On a donc : $Av_i = \sigma_i u_i$. Si A était carrée, on aurait $Av = \lambda v$ pour les valeurs propres. Ce n'est pas possible pour les matrices rectangulaires car elles représentent des applications linéaires pour lesquelles l'espace d'entrée de A est différent de l'espace de sortie.

$$A = U(\Sigma V^T) = \sum_{i=1}^{m} u_i (\widetilde{\Sigma V^T})_j^T = \sum_{i=1}^{r} \sigma_i u_i{}_v^T$$

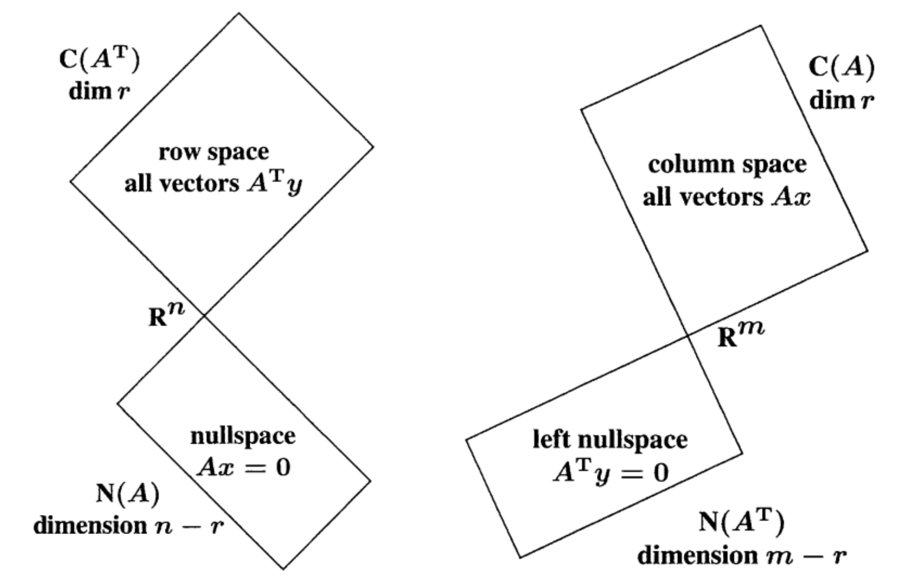
Appliquons la décomposition de A à un vecteur v quelconque pour décomposer les étapes d'une application linéaire quelconque associée à A : $Av = (U(\Sigma(V^T v)_{(1)})_{(2)})_{(3)}$

1. V^T : changement de repère orthonormé
2. Σ : mise à l'échelle par σ_i le long des nouveaux axes
3. U : à nouveau changement de repère orthonormé, transformation orthogonale

4.2 Espaces associés à une matrice

Le **noyau** d'une application linéaire $f : E \leftarrow V$ est définit comme suit : $\mathrm{Ker}(f) = \{v | v \in E \text{ et } f(v) = 0_F\}$. C'est l'ensemble des vecteurs de l'espace de départ dont l'image par f vaut 0_F

L'**image** d'une même application linéaire est définie comme suit : $\{v | v \in F \text{ et } \exists w \in E, v = f(w)\}$. C'est l'ensemble des vecteurs de l'espace d'arrivée qui ont un antécédent par f dans l'espace de départ. On l'appelle aussi *column space* $C(A)$.



$C(A^T)$ et $N(A)$ sont orthogonaux. De la même manière, $C(A)$ et $N(A^T)$ sont aussi orthogonaux. Leur somme est directe : $C(A^T) \oplus N(A) = \mathbb{R}^m \iff C(A^T) \cap N(A) = \{0\}$, et de la même manière $C(A) \oplus N(A^T) = \mathbb{R}^n \iff C(A) \cap N(A^T) = \{0\}$

Si $\dim(C(A)) = r$, alors $\dim(N(A^T)) = m - r$, il s'agit des dimensions "non atteignables" de l'espace d'arrivée, qui n'ont pas d'antécédent par f. $N(A^T)$ est le *co-noyau* car il s'agit du noyau de A^T

De la même manière, si $\dim(N(A)) = n - r$, alors $\dim(C(A^T)) = r$, c'est la *co-image*.

Exemple. Trouver la décomposition en valeurs singulières de A^{-1}.
Pour montrer que Σ est inversible : $\Sigma = U^T AV$ or A est inversible. Posons $S = V^T A^{-1} U$. On a donc $\Sigma S = I = S\Sigma$ donc $S = \Sigma^{-1}$, avec $S = \begin{bmatrix} \frac{1}{\sigma_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sigma_n} \end{bmatrix}$.
$A^{-1} = V\Sigma^{-1}U^T$. Il faut ensuite réordonner les colonnes de V, U et les σ pour obtenir un ordre décroissant.