Université Aix-Marseille

MIA

Scribe Séance 2

Aguedo Alexandre, Durand Romain
Promo 2023-2024

Tuteur : M.SCHATZ

Septembre 2023

Table des matières

0.1 Produit scalaire

Définition: dans un espace vectoriel réel (i.e. K = **R**), un *produit scalaire* est une forme bilinéaire symétrique définie positive

Forme : $(\mid) : \left| \begin{array}{ccc} E \times E & \to & \mathbf{R} \\ (v, w) & \mapsto & (v \mid w) \end{array} \right.$

Bilinéaire : linéaire en v pour w fixé et en w pour v fixé

Symétrique : Pour tout $v, w \in E^2,\ (v \mid w) = (w \mid v)$

Définie positive : Pour tout $v \in E \setminus \{0_E\}, (v \mid v) > 0$

Exemple :
Le produit scalaire canonique de $\mathbb{R}^n$, si $x = (x_1, ..., x_n)$ et $y = (y_1, ..., y_n)$ sont dans $\mathbb{R}^n$, on pose :

$$(x|y) = x_1 y_1 + \cdots + x_n y_n = X^\top Y$$

où X et Y sont les matrices colonnes des coordonnées de x et y.

0.2 Normes

Norme associée au produit scalaire :

$$\|v\| = \sqrt{(v|v)}$$

Exemple pour la norme Euclidienne (norme L2) dans $\mathbb{R}^n$, avec $v = (v_1, ..., v_n)$:

$$\|\mathbf{v}\|_2 = \sqrt{\sum_{i=1}^n v_i^2}$$

Distance d induite par la norme :

$$\forall (x, y) \in E \times E \quad d(x, y) = \|y - x\|$$

0.3 Orthogonalité

Vecteurs orthogonaux: si et seulement si $(v \mid v) = 0$

Famille orthogonale:

$(v_i)_{i \in I}$, telle que pour tout $(i,j) \in I^2$, $i \neq j$, $(v_i \mid v_j) = 0$

Famille orthonormale: famille orthogonale telle que

$$\text{pour tout } i \in I, \ \|v_i\| = 1$$

Théorème : Une famille de vecteurs orthogonaux est libre.

Comment construire des familles orthonormales ?

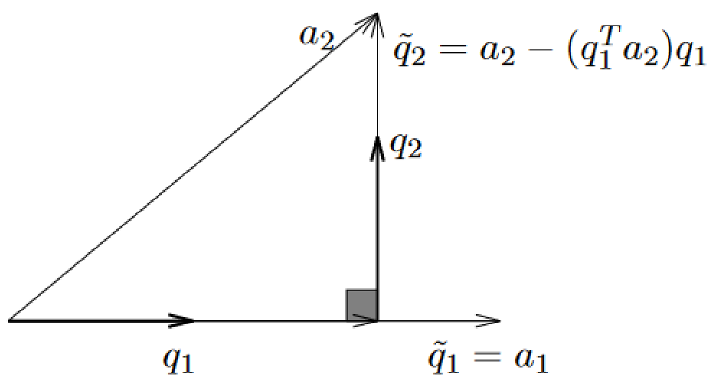
On utilise la **Méthode de Gram-Schmidt** :

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.

Note : Pour trouver, $q2$ on soustrait à $a2$ la composante de $a2$ aligner selon $q1$, soit le projeté orthogonal de $a2$ sur $q1$, ou encore $(q_1^T \cdot a_1)q_1$.

On peut le voir sur cette figure :



0.3.1 Matrice Orthogonal

Définition : Une matrice est dite semi-orthogonale si ces colonnes sont orthonormales.
Une matrice semi-orthogonale carrée est dite orthogonale. **Quelques propriétés d'une matrice orthogonale :**

- Une matrice réelle A est orthogonale si et seulement si elle est inversible et son inverse est égale à sa transposée : $A^{-1} = A^t$. On a donc $AA^t = AA^t = In$.
- Une matrice orthogonale représente une base orthonormée.
- Également, une matrice carrée est orthogonale si et seulement si sa transposée l'est (c.-à-d. $AA^t = In$), donc si et seulement si ses vecteurs lignes sont orthogonaux deux à deux et de norme 1.

Note : A^tA a pour coefficient $c_{ij} = \widetilde{a}_i^T a_j = (\widetilde{a}_i|a_j)$ qui, si A est orthogonal, est égale à 1 si $i = j$ et 0 sinon, car les colonnes de A sont orthogonales deux à deux et de norme 1. On a bien $A^tA = In$.

0.3.2

Transformation QR :

$$\underbrace{\begin{bmatrix} a_1 & a_2 & \cdots & a_k \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} q_1 & q_2 & \cdots & q_k \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1k} \\ 0 & r_{22} & \cdots & r_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{kk} \end{bmatrix}}_{R}$$

k : rang de la matrice A

$$r_{ij} = q_i^T a_j$$ Calculé durant Gram-Schmidt

A : Matrice de dimensions nxk
Q : Matrice orthogonale de dimension nxk
R : Matrice triangulaire supérieur de dimension kxk

Intérêt ?

Soit E un espace vectoriel de dimension finie n muni d'un produit scalaire $(\,.\, | \,.\,)$ et $V = (v_1, \ldots, v_n)$ une base orthonormale de E. Alors, pour tout $v \in E$,

$$v = \sum_{i=1}^{n} (v \mid v_i) v_i.$$

On peut calculer la décomposition d'un vecteur dans une base orthonormé simplement avec des produits scalaire.

La transformation QR est, par exemple, utilisée pour le calcul de solutions de systèmes linéaires. Effectivement, $AX = Y$ peut alors s'écrire : $QRX = Y$ ou $RX = QTY$.

On a alors une résolution rapide du système sans avoir à calculer la matrice inverse de A.

0.3.3 Décomposition en valeur singulière dans $\mathbb{R}$:

Etant donné une matrice rectangulaire quelconque

(à coefficients réels)

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix}$$

Quelle transformation linéaire représente-t-elle ?

Par exemple : $M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$

Sous cette forme, il est très difficile de savoir la transformation linéaire que cette matrice représente, mais on peut se faire une idée en observant sa décomposition en valeur singulière.

Théorème : Soit M une matrice $m \times n$ dont les coefficients appartiennent à $\mathbb{R}$. Alors il existe une factorisation de la forme :

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

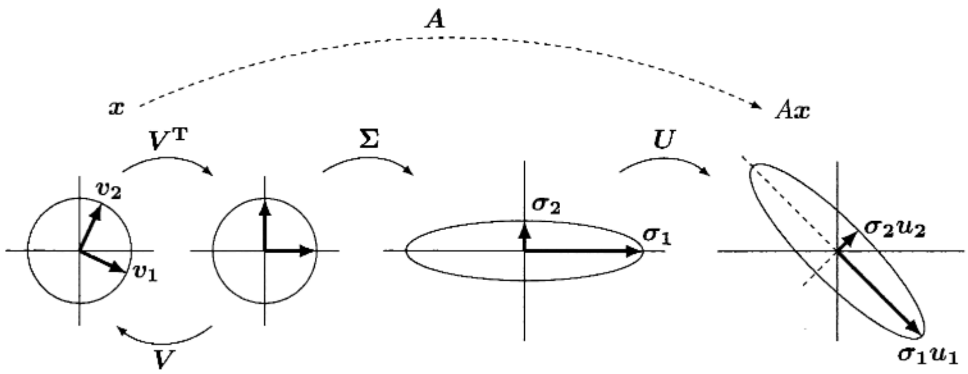
$$\boldsymbol{AV = U\Sigma} \qquad A \begin{bmatrix} \boldsymbol{v}_1 \,..\, \boldsymbol{v}_r \,..\, \boldsymbol{v}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{u}_1 \,..\, \boldsymbol{u}_r \,..\, \boldsymbol{u}_m \end{bmatrix} \left[\begin{array}{ccc|c} \boldsymbol{\sigma}_1 & & & \\ & \ddots & & \mathbf{0} \\ & & \boldsymbol{\sigma}_r & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right]$$

A : Matrice à coefficient réelle de dimension $m \times n$
U : Matrice orthogonal de dimension $m \times m$, base de $\mathbb{R}^m$
Σ : Matrice diagonale de dimension $m \times n$ dont les éléments diagonaux, appelés valeurs singulières de A, sont triés par ordre décroissant et positifs.
V : Matrice orthogonal de dimension $n \times n$, base de $\mathbb{R}^n$

Remarque :

— Nous avons prouvé dans l'exercice 2.6 que $A = \sum_i \sigma_i u_i v_i^T$.
Les σ_i sont les valeurs singulières de A, les u_i forme la matrice U et les v_i la matrice V.
Cette décomposition est intéressante car elle nous donne des informations sur A et les transformations qu'elle applique sur un vecteur d'entrée x en effectuant le produit Ax.

Ce résultat est une projection de x dans une certaine direction (multiplication de x par le vecteur $\sigma_i u_i v_i^T$), avec des facteurs d'échelle donnés par les valeurs singulières σ_i.



Par exemple : $M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$

$$U \approx \begin{pmatrix} -.21 & -.89 & .41 \\ -.52 & -.25 & -.82 \\ -.83 & .39 & .41 \end{pmatrix} \quad V \approx \begin{pmatrix} -.48 & -.57 & -.66 \\ .78 & .08 & -.62 \\ .41 & -.82 & .41 \end{pmatrix}$$

$$\begin{pmatrix} \sigma_1 & \sigma_2 & \sigma_3 \end{pmatrix} \approx \begin{pmatrix} 16.85 & 1.07 & 0 \end{pmatrix}$$

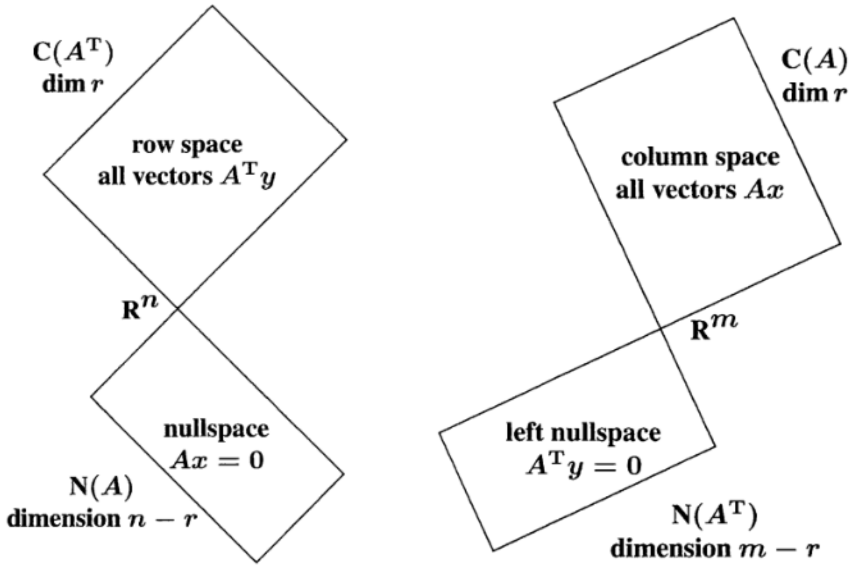
Cette décomposition matricielle nous indique que la matrice M appliqué sur x va grandement "élargir" selon la première composante (facteur 17), n'aura pas d'incidence sur la deuxième et va annuler la troisième.

— Dans l'exercice 15 nous avons démontré que $\det(Q^T AQ) = \det(A)$.Cette propriété signifie que le déterminant d'une matrice est invariant sous des transformations similaires.

0.3.4 Espace associé à une matrice

Le **noyau** d'une application linéaire $f : E \leftarrow V$ est définit comme suit : $\text{Ker}(f) = \{v | v \in E \text{ et } f(v) = 0_F\}$. 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*.

Cette partie est tirée du scribe 2.1 de EVEN Lwena.

0.4 Théorème spectral

Si S est une matrice symétrique, réelle de taille $m \times m$, alors il existe une matrice orthogonale réelle Q de taille $m \times m$ et une matrice diagonale réelle Λ de taille $m \times m$ telles que $S = Q\Lambda Q^T$.

- Matrice symétrique : une matrice est dite symétrique si elle est égale à sa transposée, donc par rapport a la diagonale. $A = A^T$

- Matrice orthogonale : une matrice carrée est dite orthogonale si le produit de la matrice et sa transposée donne la matrice identité. $Q^TQ = QQ^T = I$

- Matrice diagonale : Une matrice est dite diagonale si tous ses éléments hors de la diagonale principale sont nuls.

La matrice Q contient les vecteurs propres de S et la matrice λ contient les valeurs propres de S sur sa diagonale. Le théorème spectrale est également équivalent à $A = U\Sigma U^T$. U doit être une matrice orthogonale et Σ est une matrice diagonale.

Remarque : À l'aide du résultat de l'exercice 2.9 et de la série entière $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, nous obtenons la formule $e^A = QDiag(e^{\lambda^1}...e^{\lambda^n})Q^t$ avec les λ_i les valeurs propres de A. Cette formule à par exemple des applications dans les systèmes d'équations différentielles.

0.5 Positivité

- Matrice définie positive : Une matrice symétrique réelle est dite définie positive, noté $S \succ 0$, ssi pour toute matrice colonne u, $u^T Su > 0$. Le produit scalaire de u avec Su est toujours positif pour n'importe quel vecteur colonne u non nul.
- Matrice semi-définie positive : Une matrice symétrique réelle est dite semi-définie positive, noté $S \succeq 0$, ssi pour toute matrice colonne u, $u^T Su \geq 0$. Le produit scalaire de u avec Su est toujours non négatif.

Relation d'ordre sur les matrices symétriques réelles :
- $S_1 \prec S_2$ ssi $S_2 - S_1 \succ 0$. Cette formule signifie que que la différence entre S_1 et S_2 est définie positive.
- $S_1 \preceq S_2$ ssi $S_2 - S_1 \succeq 0$. Cette formule signifie que que la différence entre S_1 et S_2 est définie semi-définie positive.

La notion de positivité est importante pour l'optimisation, car une matrice hessienne positive indique un minimum local d'une fonction. Elle est également utilisé en probabilités.

0.6 Diagonalisation

Soit A une matrice à coefficients réels de taille $m \times m$. $\lambda \in C$ est une valeur propre de A ssi il existe $v \in C^m$, non-nul, tel que $Av = \lambda v$, i.e. l'image de v par A est dans la même direction que v. Un tel v est appelé un vecteur propre de A associé à la valeur propre λ.
Voici comment on effectue une diagonalisation :

Diagonalisation

suppose $v_1, \ldots, v_n$ is a *linearly independent* set of eigenvectors of $A \in \mathbf{R}^{n \times n}$:

$$Av_i = \lambda_i v_i, \quad i = 1, \ldots, n$$

express as

$$A \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$$

define $T = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ and $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$, so

$$AT = T\Lambda$$

and finally

$$T^{-1}AT = \Lambda$$

Si $v_1, ..., v_n$ est un ensemble linéairement indépendant de vecteurs propres de A dans $R^{n \times n}$ tels que $Av_i = \lambda_i v_i$ pour $i = 1, ..., n$, alors on peut exprimer A en termes de ses vecteurs propres et valeurs propres.
On définit T comme la matrice dont les colonnes sont les vecteurs propres $v_1, ..., v_n$ et Λ la matrice

diagonale de valeurs propres $\lambda_1, ..., \lambda_n$. Donc $AT = T\Lambda$ et finalement $T^{-1}AT = \Lambda$
La diagonalisation permet de simplifier l'analyse et la manipulation des matrices, en particulier de résoudre des systèmes d'équations différentielles et analyser les mouvements des systèmes linéaires.
Les matrices nilpotentes, c'est a dire qui ont la propriété qu'une certaine puissance de la matrice est la matrice zéro ($A^k = 0$), sont généralement non-diagonisables, il est donc important de vérifier ce point avant d'essayer pour rien.

0.7 Forme canonique de Jordan

La forme canonique de Jordan est une représentation d'une matrice sous forme quasi-diagonale. Utile pour les matrices non-diagonisables (comme les matrices nilpotentes).
Toute matrice carrée peut être transformée en une forme canonique de Jordan via une transformation de similarité, c'est a dire qu'il existe une matrice inversible P telle que $J = P^{-1}AP$, où J est la forme de Jordan de la matrice A.
Une matrice est dite être en forme de Jordan si elle est composée de blocs de Jordan. un bloc de Jordan est une matrice carrée associée à une valeur propre spécifique. Sur la diagonale de chaque bloc, on retrouve la valeur propre, et juste au-dessus de la diagonale, on a des 1, avec tous les autres éléments étant 0.
Voici l'explication en image de ce que je vient d'écrire :

$$T^{-1}AT = J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_q \end{bmatrix}$$

where

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \in \mathbf{C}^{n_i \times n_i}$$

is called a *Jordan block* of size n_i with eigenvalue λ_i (so $n = \sum_{i=1}^{q} n_i$)

Comme pour la diagonalisation d'une matrice, la forme canonique de Jordan permet comme la diagonalisation de résoudre des systèmes d'équations différentielles. Toutes les matrices ont une forme de Jordan, diagonisables ou non.

0.8 Déterminant et trace

Pour une matrice A carrée $n \times n$:
$\det(A) = \sum_{\sigma \in S_n} (\mathrm{sgn}(\sigma) \Pi_{i=1}^{n} a_{i,\sigma i})$
$\det(AB) = \det(A)\det(B)$
$\mathrm{Tr}(A) = \sum_{i=1}^{n} a_{i,i}$
$\mathrm{Tr(AB)} = \mathrm{Tr(BA)}$
$\mathrm{Tr}(A_1 A_2 ... A_k) = \mathrm{Tr}(A_2 A_3 ... A_k A_1) = \mathrm{Tr}(A_k A_1 A_2 ... A_{k-1})$

— Le déterminant est une valeur scalaire qui peut être calculée à partir des éléments d'une matrice carrée. Il donne des informations sur la matrice, notamment si elle est inversible ou non. Si le déterminant est non nul, la matrice est inversible.

— La trace d'une matrice est la somme de ses valeurs diagonales. Elle est importante car elle est invariante sous changement de base, cela signifie qu'elle reste la même quelle que soit la base choisie pour représenter la matrice.

0.9 Polynôme caractéristique et théorème de Cayley-Hamilton

Le théorème dit que pour toute matrice $A \in R^{n \times n}$, on a $\chi(A) = 0$. Ce qui signifie que la matrice satisfait son propre polynôme caractéristique. Le polynôme caractéristique d'une matrice A est donné par :
$\chi(S) = \det(sI - A)$
Si $p(s) = a_0 + a_1 s + ... + a_k s^k$ est un polynôme et A est une matrice appartenant à $R^{n \times n}$, alors :
$p(A) = a_0 I + a_1 A + ... + a_k A^k$
La corollaire dit que si A est une matrice de $R^{n \times n}$ et si $p(s)$ est un polynôme tel que $p(A) = 0$, alors tous les termes de degré n ou plus dans $p(s)$ peuvent être exprimés comme une combinaison linéaire des termes $I, A, A^2, ..., A^{n-1}$
Cela signifie que si un polynôme évalué à une matrice donne la matrice nulle, alors les termes de degré élevé de ce polynôme peuvent être réduits en utilisant les puissances inférieures de la matrice jusqu'à A^{n-1}, ce qui permet de simplifier les calculs lors de puissance élevée de matrice, résoudre des équations différentielles linéaires et est généralement utilisé pour prouver d'autres résultats en algèbre linéaires.