Mathématiques pour l’intelligence artificielle

UE MIA, M2 IAAA, AMU, 2022-2023

Thomas Schatz

Vendredi 9 septembre 2022

Scribes

Plan du cours (prévisionnel)

9 séances de 3h

Partie 1 (DM1) : algèbre linéaire et probabilités

1. Notions de bases sur les preuves (+ Algèbre linéaire?)
2. Algèbre linéaire (+ Probabilités?)
3. Probabilités

Partie 2 (DM2): statistique et optimisation

4. Statistiques
5. Optimisation

Partie 3 (DM3):

6. Optimisation sous contraintes
7. Optimisation stochastique
8. Théorie de l'apprentissage

9. Putting it all together

Algèbre linéaire

1. Espaces vectoriels et fonctions linéaires
2. Matrices
3. Angles et orthogonalité
4. Structure des applications linéaires et matrices
5. Espaces de matrices et normes
6. Algèbre linéaire numérique

Matrice

- Objet central de l'algèbre linéaire (en pratique)
- Idée générale: simplement un tableau de nombres en deux dimensions

$$M = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{np} \end{pmatrix}$$

Matrice d’une fonction linéaire

Multiplication matrice/vecteur colonne: règle

Multiplication matrice-matrice: règle

Matrice

- Objet central de l'algèbre linéaire (en pratique)
- Idée générale: simplement un tableau de nombres en deux dimensions
- Pourquoi est-ce utile ?

$$\mathrm{M} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{np} \end{pmatrix}$$

Matrice

- Objet central de l’algèbre linéaire (en pratique)
- Idée générale: simplement un tableau de nombres en deux dimensions
- Pourquoi est-ce utile ?
  - Permet de représenter les fonctions linéaires en dimensions finie de manière simple à appréhender et pratique pour les calculs

$$\mathrm{M} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{np} \end{pmatrix}$$

Matrice d'une fonction linéaire

E, F 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
$M = M(f, V, W)$ matrice de f de V vers W.

$$\mathrm{M} = \begin{array}{cl}
\begin{pmatrix}
a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\
a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\
\vdots & \vdots & & \vdots & & \vdots \\
a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\
\vdots & \vdots & & \vdots & & \vdots \\
a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{np}
\end{pmatrix} &
\begin{array}{l}
\leftarrow w_1 \\
\leftarrow w_2 \\
\\
\leftarrow w_i \\
\\
\leftarrow w_n
\end{array} \\
\begin{array}{cccccc}
\uparrow & \uparrow & & \uparrow & & \uparrow \\
f(v_1) & f(v_2) & & f(v_j) & & f(v_p)
\end{array} &
\end{array}$$

An example vector space

$$\mathbb{R}_2[X] = \mathrm{Vect}(1, X, X^2)$$

$$P \in \mathbb{R}_2[X] \Leftrightarrow \big(\text{there exists } (a_0, a_1, a_2) \in \mathbb{R}^3,\ P = a_0 + a_1 X + a_2 X^2\big)$$

Formellement:

$$1 = (1, 0, 0, \dots)$$
$$X = (0, 1, 0, \dots)$$
$$X^2 = (0, 0, 1, \dots)$$

Soit $x \in \mathbb{R}$

$$e_x : \left| \begin{array}{ccc} \mathbb{R}_2[X] & \to & \mathbb{R} \\ P = a_0 + a_1 X + a_2 X^2 & \mapsto & P(x) = a_0 + a_1 x + a_2 x^2 \end{array} \right.$$

Soit M une matrice à coefficients réels de taille n pas n

$$e_M : \left| \begin{array}{ccc} \mathbb{R}_2[X] & \to & \mathcal{M}_n(\mathbb{R}) \\ P = a_0 + a_1 X + a_2 X^2 & \mapsto & P(M) = a_0 + a_1 M + a_2 M^2 \end{array} \right.$$

Matric d'une fonction linéaire, exemple

$$V = (1, X, X^2) \text{ base de } \mathbb{R}_2[X]$$

$$W = (1, X, X^2, X^3) \text{ base de } \mathbb{R}_3[X]$$

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

$$M(f, V, W)?$$

$$\mathrm{M} = \begin{array}{cc} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1j} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2j} & \dots & a_{2p} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{i1} & a_{i2} & \dots & a_{ij} & \dots & a_{ip} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nj} & \dots & a_{np} \end{pmatrix} & \begin{array}{l} \leftarrow w_1 \\ \leftarrow w_2 \\ \\ \leftarrow w_i \\ \\ \leftarrow w_n \end{array} \\ \begin{array}{cccc} \uparrow & \uparrow & \uparrow & \uparrow \\ f(v_1) & f(v_2) & f(v_j) & f(v_p) \end{array} & \end{array}$$

Matrice d’une fonction linéaire

Multiplication matrice/vecteur colonne: interprétation

Multiplication matrice-matrice: interprétation

Matrice d’une fonction linéaire

L’évaluation en un point d’une fonction linéaire et la composition d’application linéaire se réduit (en dimension finie) à l’application “mécanique” de règles de calcul matriciel

Pas trop “mécanique” quand même:

Crimes contre les matrices
http://ee263.stanford.edu/notes/matrix_crimes.pdf

Changement de base

Toute base peut servir de système de coordonnées pour E

Changement de base == changement de système de coordonnées

Matrice de passage de V à W: $P(V, W) = M(Id_E, V, W)$

Définition: transforme un vecteur v exprimé en terme de ses coordonnées dans V, en le même vecteur en terme de ses coordonnées dans W

Pour la trouver: les colonnes de P(V,W) sont les coordonnées des éléments de V (dans un ordre fixé) dans W

Propriétés du produit matriciel

$$A \times (B \times C) = (A \times B) \times C$$

$$A \times (B + C) = A \times B + A \times C$$

$$(A + B) \times C = A \times C + B \times C$$

$$c(A \times B) = (cA) \times B = A \times (cB)$$

Transposée

Transposée de A : obtenue en échangeant les lignes et les colonnes de A

$$(A+B)^{\mathsf{T}} = A^{\mathsf{T}} + B^{\mathsf{T}}, \qquad (\alpha A)^{\mathsf{T}} = \alpha A^{\mathsf{T}}$$

$$(AB)^{\mathsf{T}} = B^{\mathsf{T}} A^{\mathsf{T}}$$

Matrice par blocs

La matrice

$$\mathbf{P} = \begin{bmatrix} 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 2 & 2 & 2 \\ 3 & 3 & 4 & 4 & 4 \\ 3 & 3 & 4 & 4 & 4 \\ 3 & 3 & 4 & 4 & 4 \end{bmatrix}$$

peut être partitionnée en quatre blocs

$$\mathbf{P}_{11} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \mathbf{P}_{12} = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}, \mathbf{P}_{21} = \begin{bmatrix} 3 & 3 \\ 3 & 3 \\ 3 & 3 \end{bmatrix}, \mathbf{P}_{22} = \begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \end{bmatrix}.$$

On peut alors écrire la matrice par bloc comme :

$$\mathbf{P}_{\text{partitionnee}} = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix}.$$

Interpretation des opérations matricielles

write $A \in \mathbf{R}^{m\times n}$ in terms of its columns:

$$A = \left[\begin{array}{cccc} a_1 & a_2 & \cdots & a_n \end{array}\right]$$

then $y = Ax$ can be written as

$$y = x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$$

write A in terms of its rows:

$$A = \left[\begin{array}{c} \tilde{a}_1^T \\ \tilde{a}_2^T \\ \vdots \\ \tilde{a}_n^T \end{array}\right]$$

then $y = Ax$ can be written as

$$y = \left[\begin{array}{c} \tilde{a}_1^T x \\ \tilde{a}_2^T x \\ \vdots \\ \tilde{a}_m^T x \end{array}\right]$$

Interpretation des opérations matricielles

$$c_{ij} = \tilde{a}_i^T b_j = \langle \tilde{a}_i, b_j \rangle$$

$$C = \left[\begin{array}{c} c_1 \cdots c_p \end{array} \right] = AB = \left[\begin{array}{c} Ab_1 \cdots Ab_p \end{array} \right]$$

$$C = \left[\begin{array}{c} \tilde{c}_1^T \\ \vdots \\ \tilde{c}_m^T \end{array} \right] = AB = \left[\begin{array}{c} \tilde{a}_1^T B \\ \vdots \\ \tilde{a}_m^T B \end{array} \right]$$

$$C = \sum_i a_i \tilde{b}_i^T$$

Algèbre linéaire

1. Espaces vectoriels et fonctions linéaires
2. Matrices
3. Angles et orthogonalité
4. Structure des applications linéaires et matrices
5. Espaces de matrices et normes
6. Algèbre linéaire numérique

Angles et orthogonalité

Définition: dans un espace vectoriel réel, 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 : $$\text{Pour tout } v, w \in E^2,\ (v \mid w) = (w \mid v)$$

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

Angles et orthogonalité

Norme associée au produit scalaire : $\|v\| = \sqrt{(v \mid v)}$

Exemple: produit scalaire et norme Euclidienne sur $\mathbf{R}^n$

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

Famille orthogonale:

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

Famille orthonormale: famille orthogonale telle que

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

Angles et orthogonalité

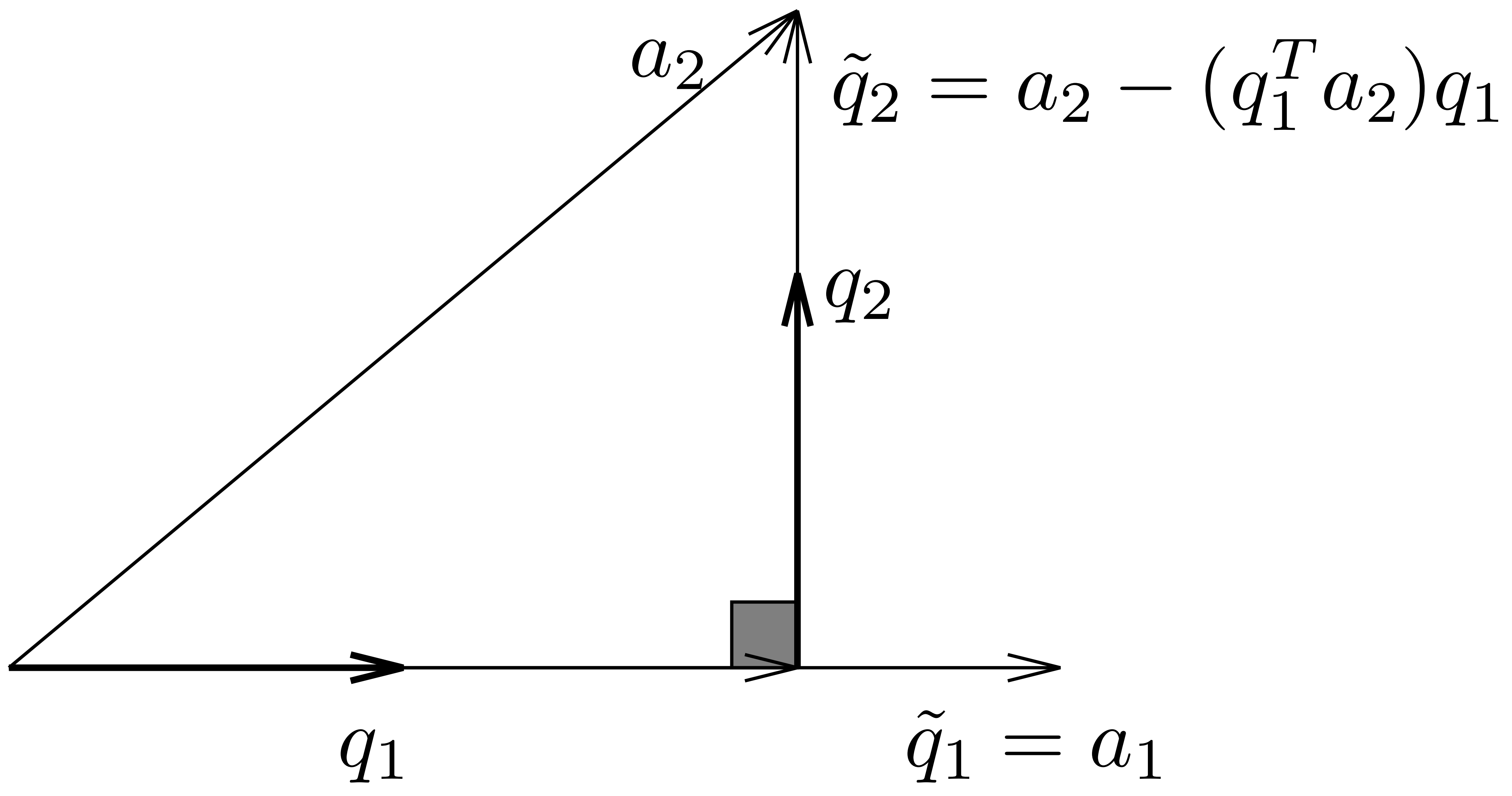
Théorème : une famille de vecteurs orthogonaux est libre

Comment construire des familles orthonormales?

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.

Orthogonalité



Gram-Schmidt et factorisation 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

Une matrice est dite semi-orthogonale si ces colonnes sont orthonormales.

Si Q est une matrice semi-orthogonale de taille $m \times k$, alors $Q^T Q = I_k$

Une matrice semi-orthogonale carrée est dite orthogonale.

Angles et orthogonalité

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 $v \in E$,

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

Pause

Algèbre linéaire

1. Espaces vectoriels et fonctions linéaires
2. Matrices
3. Angles et orthogonalité
4. Structure des applications linéaires et matrices
5. Espaces de matrices et normes
6. Algèbre linéaire numérique

Décomposition en valeurs singulières

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 ?

Décomposition en valeurs singulières

$$\boldsymbol{AV = U\Sigma} \qquad A\begin{bmatrix} \\ \boldsymbol{v}_1 \, .. \, \boldsymbol{v}_r \, .. \, \boldsymbol{v}_n \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \boldsymbol{u}_1 \, .. \, \boldsymbol{u}_r \, .. \, \boldsymbol{u}_m \\ \\ \end{bmatrix} \left[\begin{array}{ccc|c} \boldsymbol{\sigma_1} & & & \\ & \ddots & & \mathbf{0} \\ & & \boldsymbol{\sigma_r} & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right]$$

Interpretation des opérations matricielles

$$c_{ij} = \tilde{a}_i^T b_j = \langle \tilde{a}_i, b_j \rangle$$

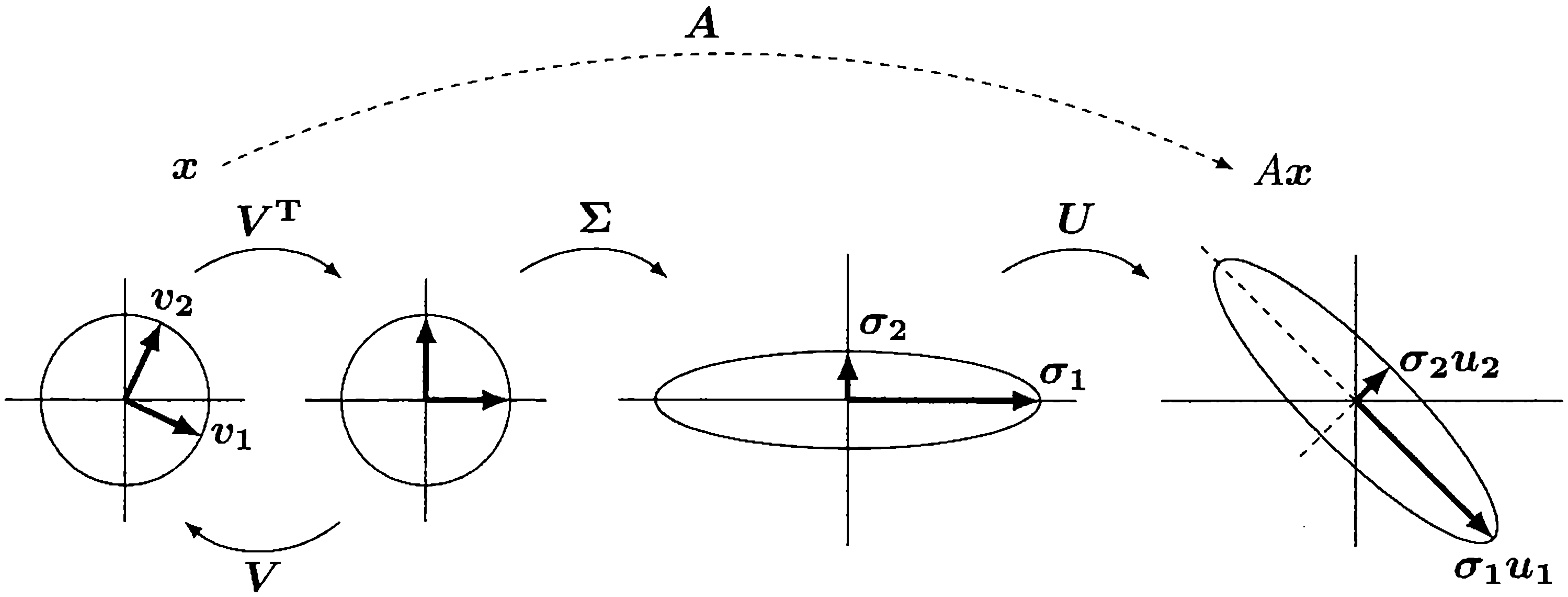
$$C = \left[\begin{array}{c} c_1 \cdots c_p \end{array}\right] = AB = \left[\begin{array}{c} Ab_1 \cdots Ab_p \end{array}\right]$$

$$C = \left[\begin{array}{c} \tilde{c}_1^T \\ \vdots \\ \tilde{c}_m^T \end{array}\right] = AB = \left[\begin{array}{c} \tilde{a}_1^T B \\ \vdots \\ \tilde{a}_m^T B \end{array}\right]$$

$$C = \sum_i a_i \tilde{b}_i^T$$

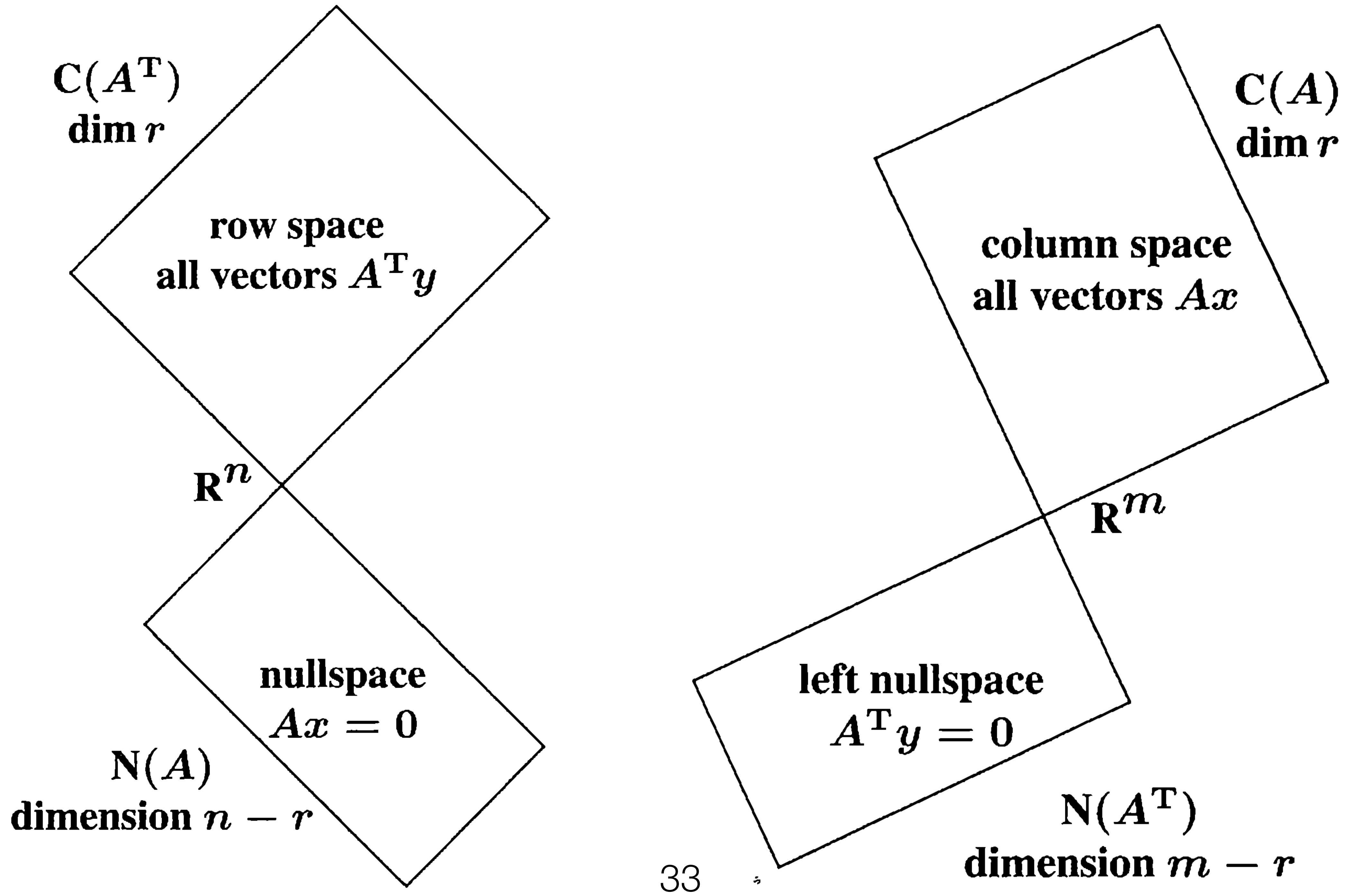
Décomposition en valeurs singulières

$$AV = U\Sigma \qquad A\begin{bmatrix} v_1 \,..\, v_r \,..\, v_n \end{bmatrix} = \begin{bmatrix} u_1 \,..\, u_r \,..\, u_m \end{bmatrix}\left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right]$$



Espaces associés à une matrice

$$AV = U\Sigma \qquad A\begin{bmatrix} & & \\ v_1 \,..\, v_r \,..\, v_n \\ & & \end{bmatrix} = \begin{bmatrix} & & \\ u_1 \,..\, u_r \,..\, u_m \\ & & \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right]$$



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 $S = Q\Lambda Q^T$.

Une matrice symétrique réelle est dite définie positive, noté $S \succ 0$ ssi pour toute matrice colonne u, $u^T Su > 0$.

Une matrice symétrique réelle est dite semi-définie positive, $S \succeq 0$, ssi pour toute matrice colonne u, $u^T Su \geq 0$.

Relation d'ordre sur les matrices symétriques réelles:

$S_1 \prec S_2$ ssi $S_2 - S_1 \succ 0$

et

$S_1 \preceq S_2$ ssi $S_2 - S_1 \succeq 0$

Diagonalisation

Soit A une matrice à coefficients réels de taille $m \times m$. $\lambda \in \mathbf{C}$ est une valeur propre de A ssi il existe $v \in \mathbf{C}^m$, 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.

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

$$A \left[\begin{array}{ccc} v_1 & \cdots & v_n \end{array} \right] = \left[\begin{array}{ccc} v_1 & \cdots & v_n \end{array} \right] \left[\begin{array}{ccc} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{array} \right]$$

define $T = \left[\begin{array}{ccc} v_1 & \cdots & v_n \end{array} \right]$ and $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$, so

$$AT = T\Lambda$$

and finally

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

Forme canonique de Jordan

any matrix $A \in \mathbf{R}^{n\times n}$ can be put in *Jordan canonical form* by a similarity transformation, $i.e.$

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

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 $n = \sum_{i=1}^{q} n_i$)

Déterminant et trace

A matrice carrée $n \times n$

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

$\operatorname{sgn}(\sigma)$ est la parité du nombre d'élement dans une décomposition de σ en une séquence de transpositions (échange de deux éléments).

$$\det(AB) = \det(A)\det(B)$$

$$\operatorname{Tr}(A) = \sum_{i=1}^{n} a_{i,i} \quad \operatorname{Tr}(\mathrm{AB}) = \operatorname{Tr}(\mathrm{BA})$$

$$\operatorname{Tr}(A_1 A_2 \ldots A_k) = \operatorname{Tr}(A_2 A_3 \ldots A_k A_1) = \cdots = \operatorname{Tr}(A_k A_1 A_2 \ldots A_{k-1})$$

Théorème de Cayley-Hamilton

Cayley-Hamilton theorem: for any $A \in \mathbf{R}^{n\times n}$ we have $\mathcal{X}(A) = 0$, where $\mathcal{X}(s) = \det(sI - A)$

Corollaire : pour tout entier naturel p, $A^p \in \mathrm{Vect}(I, A, A^2, \dots, A^{n-1})$

Algèbre linéaire

1. Espaces vectoriels et fonctions linéaires
2. Matrices
3. Angles et orthogonalité
4. Structure des applications linéaires et matrices
5. Espaces de matrices et normes
6. Algèbre linéaire numérique

Espaces de fonctions linéaires

Espace de matrice, muni de l'addition et la multiplication matricielle et de la multiplication par un scalaire:

$$\mathcal{M}_{m,n}(\mathbf{R})$$

Espace de fonctions linéaires muni de l'addition et de la composition de fonctions linéaires et de la multiplication par un scalaire:

$$\mathcal{L}(E, F)$$

Normes de vecteurs

A function $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ is called a *vector norm* if it has the following properties:

1. $\|\mathbf{x}\| \geq 0$ for any vector $\mathbf{x} \in \mathbb{R}^n$, and $\|\mathbf{x}\| = 0$ if and only if $\mathbf{x} = \mathbf{0}$
2. $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$ for any vector $\mathbf{x} \in \mathbb{R}^n$ and any scalar $\alpha \in \mathbb{R}$
3. $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ for any vectors $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$.

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}.$$

$$|\mathbf{x}^T\mathbf{y}| \leq \|\mathbf{x}\|_2\|\mathbf{y}\|_2$$

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| + \cdots + |x_n|$$

$$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\mathbf{x}^T\mathbf{x}}$$

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

Si Q est une matrice orthogonale,
$\|Qx\|_2 = \|x\|_2$

Normes de matrices

A matrix norm is a function $\|\cdot\| : \mathbb{R}^{m\times n} \to \mathbb{R}$ that has the following properties:

- $\|A\| \geq 0$ for any $A \in \mathbb{R}^{m\times n}$, and $\|A\| = 0$ if and only if $A = 0$
- $\|\alpha A\| = |\alpha|\|A\|$ for any $m \times n$ matrix A and scalar α
- $\|A + B\| \leq \|A\| + \|B\|$ for any $m \times n$ matrices A and B

$$\|A\| = \sup_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$$

$$\|A\|_2 = \sigma_1$$

$$\|A\|_F = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}.$$

$$\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$$

Application de la notion de norme: lien valeurs propres, valeurs singulières

$$\sigma_r \leq |\lambda| \leq \sigma_1$$

Application de la notion de norme: lien valeurs propres, valeurs singulières

$$\sigma_r \leq |\lambda| \leq \sigma_1$$

Éléments propres de $A^T A$ et AA^T ?

Algèbre linéaire

1. Espaces vectoriels et fonctions linéaires
2. Matrices
3. Angles et orthogonalité
4. Structure des applications linéaires et matrices
5. Espaces de matrices et normes
6. Algèbre linéaire numérique

Algèbre linéaire numérique

Exemple : produit matrice-matrice $A\colon\ m \times n,\ B\colon\ n \times p.$

Produit matriciel $C = AB$:

$$C = 0$$

Boucle $i = 1..m,\ j = 1..p,\ k = 1..n$:

$$c_{ij} = c_{ij} + a_{ik} b_{kj}$$

Ordre des trois boucles ?

i en premier $C = \begin{bmatrix} \tilde{c}_1^T \\ \vdots \\ \tilde{c}_m^T \end{bmatrix} = AB = \begin{bmatrix} \tilde{a}_1^T B \\ \vdots \\ \tilde{a}_m^T B \end{bmatrix}$

j en premier $C = \begin{bmatrix} c_1 \cdots c_p \end{bmatrix} = AB = \begin{bmatrix} Ab_1 \cdots Ab_p \end{bmatrix}$

k en premier $C = AB = \sum_{k=1}^{n} a_k \tilde{b}_i^T$

Algèbre linéaire numérique

Stabilité numérique

Algèbre linéaire numérique

Stabilité numérique

Supposons qu'un processus nous donne la solution estimée $\hat{x}$ d'un système linéaire $Ax = b$ en effectuant toutes les opérations matricielles de manière exacte mais qu'il y a des erreurs d'arrondi dans le stockage en nombre flottants de A et de b dans la mémoire (en pratique les opérations matricielles effectuées pour trouver x introduisent davantage d'erreurs d'arrondis, ce modèle donne donc une borne supérieure sur la qualité possible d'un algorithme numérique pour résoudre des systémes linéaires).

Alors,

$$\text{si } \mathbf{u}\kappa_\infty(A) \leq .5 \qquad \frac{\|x - \hat{x}\|_\infty}{\|x\|_\infty} \leq 4\mathbf{u}\kappa_\infty(A)$$

$\mathbf{u}$ est le unité d'arrondi, égale à la moitié de lécart entre 1 et le plus petit nombre flottant strictement supérieur à 1. Pour les nombres flottants IEEE single précision, $\mathbf{u}$ est d'environ 10^{-7}. Il est d'environ 10^{-16} pour les nombres flottants IEEE double précision.

$\kappa_\infty(A) := \|A\|_\infty \|A^{-1}\|_\infty$ est le conditionnement de A pour la norme ∞

Summary

Notions les plus importantes

- Matrice d'une application linéaire
- Différentes vues sur les opérations matricielles
- Produit scalaire, orthogonalité, matrices orthogonales et bases orthonormales
- Décomposition en valeurs singulières
- Normes de vecteurs et de matrices
- Stabilité numérique

Crédits

Certaines figures de

- Strang, Gilbert. *Linear algebra and learning from data*. Cambridge: Wellesley-Cambridge Press, 2019
- Boyd, Stephen EE263 - Introduction to Linear Dynamical Systems (https://see.stanford.edu/Course/EE263)
- Wikipedia