Fondamentaux d'algèbre linéaire

Réductions de matrices

Anne Sabourin

Semaine 1

- Espaces vectoriels réels
- Applications linéaires
- Matrices

Semaine 2

- Produit scalaire, projections, interprétations géométriques
- Réductions de matrices

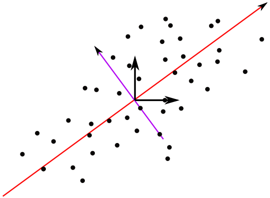
Anne Sabourin

Réduction de matrices

Principe : Matrice quelconque → produit de matrices simples, **ex** : matrices orthogonales, dilatations . . .

Applications

- Inversion de matrice, résolution de système linéaire, solution moindres carrés.
- Analyse en composantes principales (ACP) : directions de « plus grande variance » d'un jeu de données.

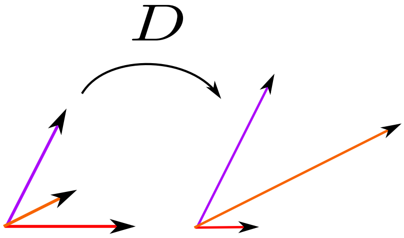


Matrice diagonale, vecteurs propres, valeurs propres

- Une matrice carrée D est diagonale si tous ses coefficients sont nuls sauf les coefficients diagonaux

$$D = \begin{pmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_n \end{pmatrix} \Leftrightarrow f(\mathbf{u}_i) = \sigma_i \mathbf{u}_i \text{ pour } f \leftrightarrow_{\mathcal{U}} M.$$

- On dit que $\mathbf{u}_i$ est **vecteur propre de** f associé à la **valeur propre** σ_i

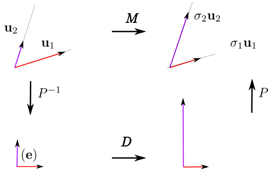


Anne Sabourin

Matrice/ application diagonalisable

- $f : \mathbb{R}^n \to \mathbb{R}^n$ est **diagonalisable** si $\exists \mathcal{U} : f(\mathbf{u}_i) = \sigma_i \mathbf{u}_i$
- M est **diagonalisable** si $f \leftrightarrow_e M$ l'est.
  $\Leftrightarrow \exists P$, matrice de passage, telle que
  $$D = P^{-1} M P \text{ diagonale.}$$
- Colonnes de P : vecteurs propres de M et f.

M (ou f) est diagonalisable s'il existe une base de vecteurs propres pour M (ou f).



Diagonalisation des matrices symétriques

- Une matrice (carrée) M est **symétrique** si $M^\top = M$.

Théorème

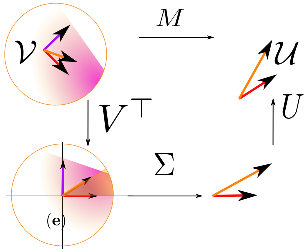
Toute matrice symétrique réelle est diagonalisable en base orthornormée

- « **en base orthornormée** » : avec une matrice de passage orthogonale.
- Autrement dit, si M est symétrique et $[f] = M$, il existe une base orthonormée $\mathcal{U}$, telle que $[f]_\mathcal{U}^\mathcal{U}$ soit diagonale.
- ou encore : si M est symétrique, il existe une matrice orthogonale P et une matrice diagonale D telles que

$$M = PDP^\top .$$

Décomposition en valeurs singulières (SVD)

- But : avoir un résultat similaire pour une matrice M ni symétrique, ni carrée.
- Principe : Ecrire la matrice comme le produit de
  - Une matrice orthogonale
  - Une matrice « diagonale »
  - Une deuxième matrice orthogonale.



SVD : existence

Soit $M \in \mathbb{R}^{n \times p}$. On peut trouver

- $U \in \mathbb{R}^{n \times n}$ et $V \in \mathbb{R}^{p \times p}$, orthogonales ($UU^T = 0, VV^T = 0$)
- une matrice 'diagonale' $\Sigma \in \mathbb{R}^{n \times p}$ de coefficients diagonaux positifs ou nuls,

telles que

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

Cette décomposition est la SVD de M.

remarques :

- matrice Σ 'diagonale' :

$$\Sigma = \left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_n & \end{array} \right) \text{ ou } \Sigma = \left(\begin{array}{ccc} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_p \\ \hline & 0 & \end{array} \right)$$

- U et V ne sont pas uniques, mais les valeurs propres σ_i le sont (à l'ordre près)

SVD en pratique : avec Python

```
import numpy as np
from numpy import linalg
M = np.matrix([[1,2,3],[4,5,6]])
U, s, W =linalg.svd(M)
```

Complexité : $O(p\,n^2)$ opérations, si $p \geq n$.

- `s` est une `array` unidimensionelle (la diagonale de Σ).
- `W` correspond à $V^\top$ dans nos notations : on a

$$M = U \Sigma W$$

- pour reconstruire M

```
    S=np.zeros( (2,3),dtype='float')
    S[:2 , :2] = np.diag(s)
    S=np.matrix(S)
    M_recons= U * S * W
```

SVD : interprétation « application linéaire »

- Soit f linéaire, $f \sim_{(\mathbf{e})_p,(\mathbf{e})_n} M$.
- Existence de la SVD $\Leftrightarrow$ existence de $\mathcal{U}, \mathcal{V}$, orthonormales, telles que

$$f \sim_{\mathcal{U},\mathcal{V}} \Sigma.$$

$\mathcal{U}$: colonnes de U ; $\mathcal{V}$: colonnes de V ;

Vecteurs, valeurs singulières

$$f(\mathbf{v}_i) = \sigma_i \, \mathbf{u}_i, \quad i \leq \min(n, p)$$

$\mathbf{u}_i$: vecteurs singuliers à gauche,
$\mathbf{v}_i$: vecteurs singuliers à gauche,
σ_i : valeurs singulières.

- si $p > n$, $f(v_i) = 0$ pour $i > n$;
- $\mathrm{Im}\, f \subset \mathrm{Vect}(u_1, \ldots, u_p)$.

SVD : interpretation « matrices »

$$M = U\Sigma V^\top \Leftrightarrow MV = \Sigma U$$
$$\Leftrightarrow MV_i = \sigma_i U_i \ , \ i \leq \min(n,p)$$

(V_i, U_i : colonnes de V et U.)

- V_i « vecteurs singuliers à droite »de M
- U_i « vecteurs singuliers à gauche »de M
- σ_i : valeurs singulières.

$$M = \left(\begin{array}{c|c|c} \mathbf{u}_1 & \dots & \mathbf{u}_n \end{array}\right) \left(\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_n & \end{array}\right) \left(\begin{array}{c} \mathbf{v}_1^\top \\ \hline \vdots \\ \hline \mathbf{v}_p^\top \end{array}\right)$$

Fondamentaux d'algèbre linéaire

Réductions de matrices

Anne Sabourin
INSTITUT Mines-Télécom

Application : solution de moindres carrés

Problème de stat : trouver $X^* = \arg\min_{X \in \mathbb{R}^p} \|MX - Y\|^2$.

- Supposons $M \in \mathbb{R}^{n \times p}$, $n < p$.
- X^* vérifie

$$M^\top M X^* = M^\top Y \qquad \text{cf :cours de stat}$$

Solution si $M^\top M$ non inversible ? (rang $r \leq n$)

Ne pas calculer directement $M^\top M$ puis inverser (`NaN`) :
utiliser la SVD de M !

Anne Sabourin

INSTITUT Mines-Télécom

Solution : pseudo-inverse et SVD

déf : pseudo-inverse de M : La matrice $M^\dagger$ telle que, $\forall\, Y$,

le vecteur $\hat{X} = M^\dagger Y$ soit

- solution de $\min_X \|MX - Y\|^2$
- de plus faible norme euclidienne parmi les solutions.

Existence, unicité : admis.

$$M^\dagger \textbf{ se calcule avec la SVD } : M^\dagger = V\Sigma^\dagger U^\top$$

où

$$\Sigma^\dagger = \left(\begin{array}{cccc|c} \sigma_1^{-1} & & & & \\ & \ddots & & & \\ & & \sigma_r^{-1} & & 0 \\ & & & 0 & \end{array}\right)$$

Moindres carrés : conclusion

$$\text{Solution choisie : } X^* = \mathbf{V}\,\mathbf{\Sigma}^\dagger\,\mathbf{U}\,Y \qquad (= \hat{X})$$

- Existe même si $M^\top M$ non inversible !
- Stable si $M^\top M$ inversible mais avec très petites valeurs propres (colonnes 'presque' colinéaires).
- Coût numérique : la décomposition $U,\ \Sigma,\ V^\top$.

Conclusion du cours

Après ce cours - et du travail personnel, vous devriez pouvoir aborder sereinement :

- les statistiques dans le modèle linéaire
- Les techniques de classifications basées sur la séparation par des hyper-plans (perceptron, SVM)
- la réduction de dimension (PCA)