

Analyse en composantes principales

Séance 6 – Projection, reconstruction et inertie

Christophe Ambroise 

UEVE, UMR CNRS 8071

19 octobre 2026

1

Un modèle latent

2

Modèle latent gaussien

Une structure linéaire de dimension q , perturbée par un bruit isotrope.

$$\mathbf{x} = \mu + W\mathbf{z} + \varepsilon,$$

$$\mathbf{z} \sim \mathcal{N}(\mathbf{0}, I_q),$$

$$\mathbf{x} \mid \mathbf{z} \sim \mathcal{N}(\mu + W\mathbf{z}, \sigma^2 I_p),$$

$$\mathbf{x} \sim \mathcal{N}(\mu, WW^\top + \sigma^2 I_p).$$

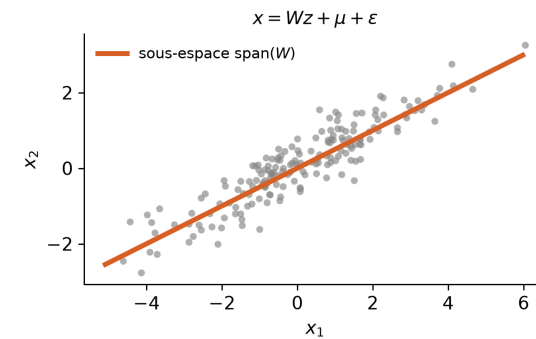


Figure 1

ACP probabiliste : Tipping et Bishop (1999).

3

Algorithme EM

E – moments conditionnels des coordonnées latentes.

M – paramètres (μ, W, σ^2) par vraisemblance complète espérée.

$$Q(\theta \mid \theta^{(t)}) = \mathbb{E}_{Z \mid X, \theta^{(t)}} [\log p(X, Z \mid \theta)],$$
$$\theta^{(t+1)} = \arg \max_{\theta} Q(\theta \mid \theta^{(t)}).$$

$\sigma^2 \rightarrow 0$: sous-espace estimé = sous-espace principal classique.

4

Réduire la dimension

5

Le problème

n individus, p variables quantitatives, souvent redondantes.

$q < p$ coordonnées synthétiques.

Deux exigences : variance conservée, reconstruction fidèle.

Une ligne par individu, une colonne par variable.

$$X = \begin{pmatrix} x_1^\top \\ \vdots \\ x_n^\top \end{pmatrix} \in \mathbb{R}^{n \times p}.$$

6

Le centre du nuage

Le barycentre des individus.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad X_c = X - \mathbf{1}_n \bar{x}^\top.$$

ACP = géométrie du nuage **centré**.

7

Deux géométries

Variables centrées

ACP de covariance

Unités conservées

Variables centrées-réduites

ACP de corrélation

Variances égales à 1

La standardisation change la question géométrique.

8

L'inertie totale

La dispersion quadratique autour du centre.

$$I_T = \sum_{i=1}^n \|x_i - \bar{x}\|^2 = \|X_c\|_F^2.$$

9

Un sous-espace

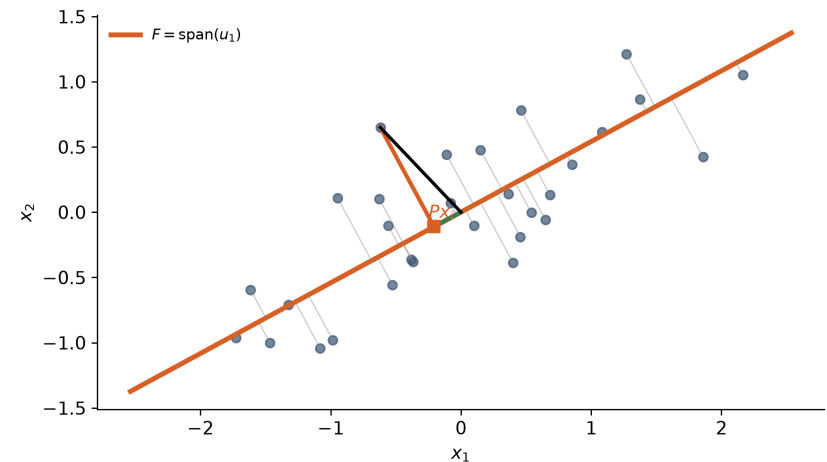
$F \subset \mathbb{R}^p$, dimension q ; P_F son projecteur orthogonal.

$$\hat{x}_i = \bar{x} + P_F(x_i - \bar{x}), \quad r_i = (I_p - P_F)(x_i - \bar{x}).$$

10

Le triangle rectangle

Pour chaque individu : inertie = projection + résidu.



Pythagore, pour le nuage

La somme des identités individuelles.

$$\underbrace{I_T}_{\text{inertie totale}} = \underbrace{\sum_i \|P_F(x_i - \bar{x})\|^2}_{I_{\text{proj}}(F)} + \underbrace{\sum_i \|(I_p - P_F)(x_i - \bar{x})\|^2}_{I_{\text{res}}(F)}.$$

$I_T = I_{\text{proj}} + I_{\text{res}}$ – égalité exacte.

Deux critères, un optimum

À dimension q fixée, même sous-espace optimal.

$$\arg \max_{\dim(F)=q} I_{\text{proj}}(F) = \arg \min_{\dim(F)=q} I_{\text{res}}(F).$$

La solution spectrale

Matrice de covariance

Convention population, cohérente avec l'inertie.

$$S = \frac{1}{n} X_c^\top X_c, \quad S = S^\top \succeq 0.$$

15

Une direction unitaire

Coordonnées projetées : $c = X_c u$.

Variance portée par l'axe u .

$$\frac{1}{n} \|c\|^2 = \frac{1}{n} \|X_c u\|^2 = u^\top S u, \quad \|u\| = 1.$$

16

Premier axe principal

Tout maximiseur est un vecteur propre associé à λ_1 .

$$\max_{\|u\|=1} u^\top S u = \lambda_1, \quad S u_1 = \lambda_1 u_1, \quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq$$

17

Premier axe : démonstration

Théorème spectral : $S = \sum_j \lambda_j u_j u_j^\top$, base (u_j) orthonormée.

$$u = \sum_j a_j u_j, \quad \sum_j a_j^2 = 1, \\ u^\top S u = \sum_j \lambda_j a_j^2 \leq \lambda_1 \sum_j a_j^2 = \lambda_1.$$

- Égalité en $u = u_1$: maximum atteint.
- Égalité $\iff a_j = 0$ dès que $\lambda_j < \lambda_1$: u dans l'espace propre de λ_1 , donc $Su = \lambda_1 u$.
- Positivité : $\lambda_j = u_j^\top S u_j = \frac{1}{n} \|X_c u_j\|^2 \geq 0$.

18

Sous-espace principal

Les q premiers axes engendrent le meilleur sous-espace.

$$F_q = \text{span}(u_1, \dots, u_q), \quad P_q = U_q U_q^\top.$$

19

Sous-espace principal : démonstration

P projecteur sur F , $h_j = \|Pu_j\|^2 \in [0, 1] : \sum_j h_j = \text{tr } P = q$
et $I_{\text{proj}}(F) = n \text{tr}(PS) = n \sum_j \lambda_j h_j$.

$$\begin{aligned} \frac{1}{n} I_{\text{proj}}(F) - \sum_{j \leq q} \lambda_j &= \sum_{j \leq q} \lambda_j (h_j - 1) + \sum_{j > q} \lambda_j h_j \\ &\leq \lambda_q \left(\sum_j h_j - q \right) = 0. \end{aligned}$$

- Égalité pour F_q ($h_j = 1$ si $j \leq q$, 0 sinon) ; unique si $\lambda_q > \lambda_{q+1}$.
- $q = 1 : h_j = a_j^2$, preuve du premier axe.

20

Composantes principales

Scores des individus dans la nouvelle base ; covariances nulles.

$$C = X_c U, \quad \frac{1}{n} C^\top C = \Lambda.$$

21

Formule de reconstitution

Reconstruction de rang q .

$$\hat{X}_q = C_q U_q^\top = X_c U_q U_q^\top = X_c P_q.$$

22

Théorème d'Eckart–Young

$\hat{X}_q$ est la meilleure approximation de rang au plus q .

$$\hat{X}_q \in \arg \min_{\text{rang}(A) \leq q} \|X_c - A\|_F^2, \quad \|X_c - \hat{X}_q\|_F^2 = n \sum_{j>q} \lambda_j.$$

23

Inerties spectrales

Chaque axe porte $n\lambda_j$ unités d'inertie.

$$I_T = n \sum_{j=1}^p \lambda_j = n \text{tr}(S),$$

$$I_{\text{proj}}(F_q) = n \sum_{j=1}^q \lambda_j,$$

$$I_{\text{res}}(F_q) = n \sum_{j=q+1}^p \lambda_j.$$

24

Dimension q

Taux d'inertie expliquée.

$$\rho_q = \frac{I_{\text{proj}}(F_q)}{I_T} = \frac{\sum_{j=1}^q \lambda_j}{\sum_{j=1}^p \lambda_j}.$$

Seuil, coude, interprétabilité — trois critères, aucun automatisme.

25

Iris sans étiquette

26

Standardisation

150 fleurs, quatre mesures en centimètres, quatre variances unitaires.

species retirée avant tout calcul.

$$z_{ij} = \frac{x_{ij} - \bar{x}_j}{s_j}, \quad \text{tr}(S) = 4, \quad I_T = np = 600.$$

27

Le spectre

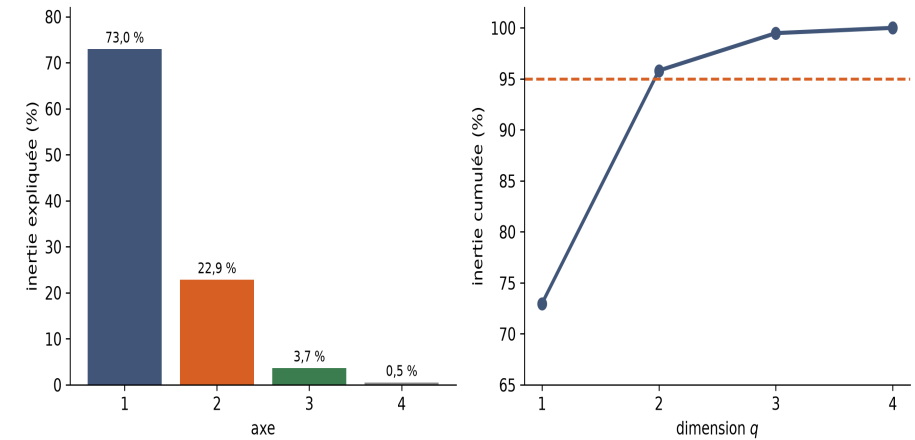


Figure 3

Deux axes : **95,8 %** de l'inertie.

28

Le plan principal

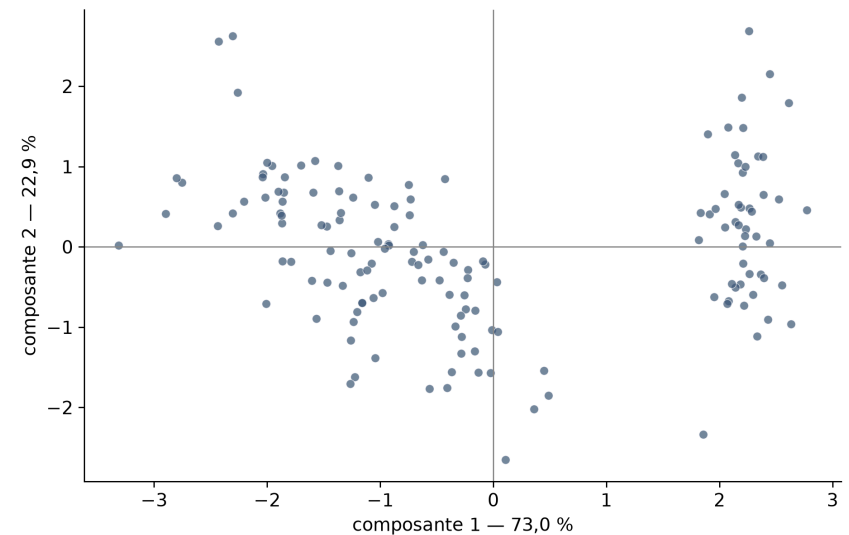


Figure 4

Aucune classe dans la construction.

29

Lecture *a posteriori*

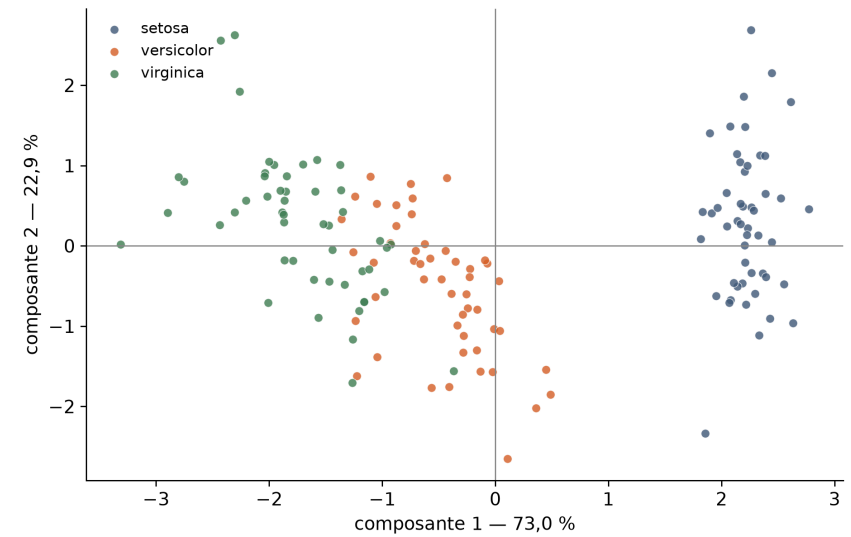


Figure 5

Les espèces colorent un résultat déjà calculé.

30

Corrélation variable-composante

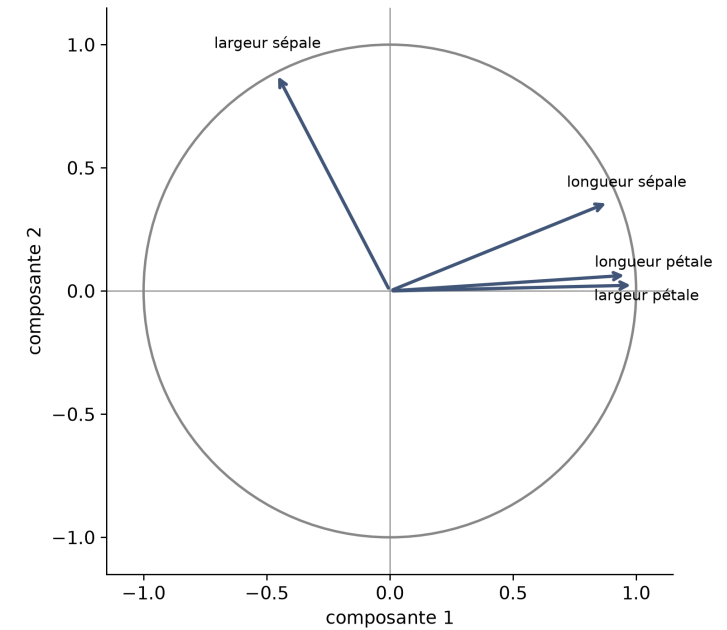
Variable standardisée z_j (colonne de Z), composante $c_k = Zu_k$,
 $S = \frac{1}{n} Z^T Z$; $\widehat{\text{cor}}$: corrélation empirique.

$$\begin{aligned}\widehat{\text{cov}}(z_j, c_k) &= \frac{1}{n} z_j^T Z u_k = (S u_k)_j = \lambda_k u_{jk}, \\ \widehat{\text{var}}(z_j) &= 1, \quad \widehat{\text{var}}(c_k) = u_k^T S u_k = \lambda_k, \\ \widehat{\text{cor}}(z_j, c_k) &= \frac{\lambda_k u_{jk}}{\sqrt{1 \cdot \lambda_k}} = \sqrt{\lambda_k} u_{jk}.\end{aligned}$$

Proximité, opposition, orthogonalité – lecture par angles.

31

Cercle des corrélations



32

Qualité d'un individu

Part de sa distance au centre visible dans le premier plan.

$$\cos^2(i, F_q) = \frac{\|P_q(x_i - \bar{x})\|^2}{\|x_i - \bar{x}\|^2} = \frac{\sum_{j=1}^q c_{ij}^2}{\|x_i - \bar{x}\|^2}.$$

33

Contribution d'un individu

Part de l'inertie de l'axe j due à l'individu i .

$$\text{CTR}(i, j) = \frac{c_{ij}^2}{n\lambda_j}, \quad \sum_{i=1}^n \text{CTR}(i, j) = 1.$$

Qualité : représentation de i . Contribution : influence de i sur l'axe.

34

Reconstruction par rang

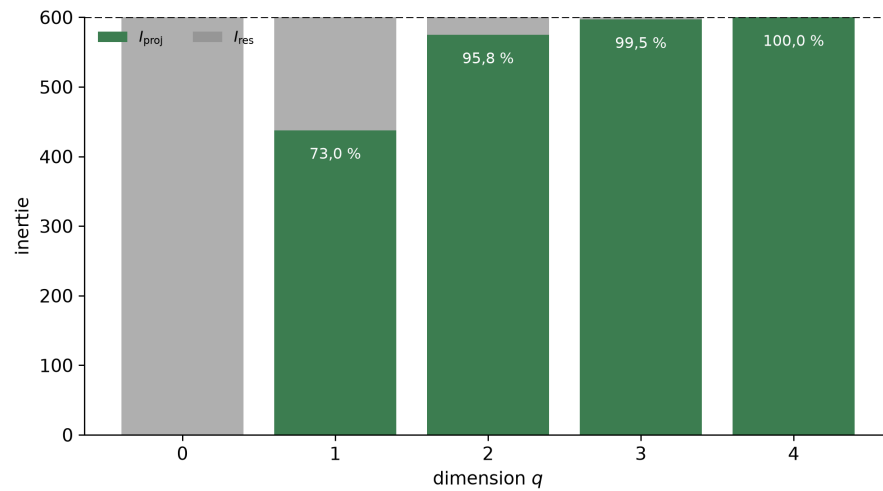


Figure 7

35

Le choix des unités

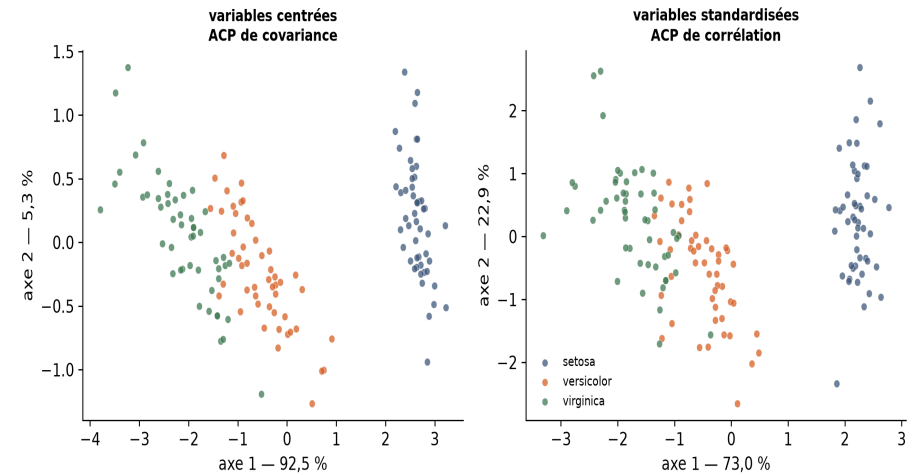


Figure 8

Covariance : variabilité physique. Corrélation : variabilité relative.

Deux théorèmes de la variance

36

37

Pythagore pour k -means

Pour x_i dans le groupe G_k , centre μ_k , centre global μ .

$$x_i - \mu = (x_i - \mu_k) + (\mu_k - \mu),$$

$$\sum_{i \in G_k} (x_i - \mu_k) = 0.$$

Le terme croisé disparaît après sommation dans chaque groupe.

38

Inertie intra + inter

Théorème de Huygens, analogue multivarié de la variance totale.

$$\underbrace{\sum_i \|x_i - \mu\|^2}_{I_T} = \underbrace{\sum_k \sum_{i \in G_k} \|x_i - \mu_k\|^2}_{I_W} + \underbrace{\sum_k n_k \|\mu_k - \mu\|^2}_{I_B}.$$

39

Deux partitions de I_T

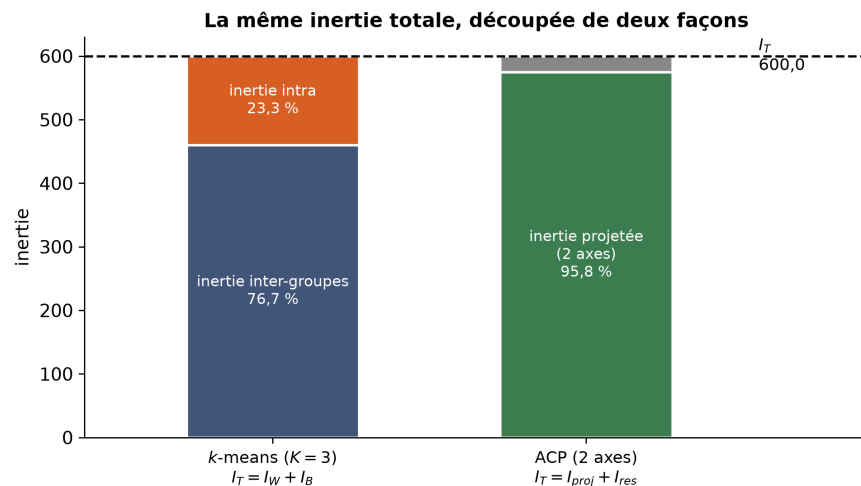


Figure 9

40

Deux questions

ACP

Sous-espace linéaire

I_{res} minimal

Compression, visualisation

Même inertie totale. Orthogonalité individuelle pour l'ACP ; centrage par groupe pour k -means.

k -means

Partition discrète

I_W minimal

Groupes, quantification

41

En bref

1. ACP : meilleure projection orthogonale de rang q .
2. Spectre : inertie projetée axe par axe.
3. Pythagore : $I_T = I_{\text{proj}} + I_{\text{res}}$.
4. k -means : $I_T = I_W + I_B$.
5. Modèle gaussien latent : même sous-espace, lecture probabiliste et EM.