

## Table des matières

<b>1</b>	<b>Introduction à l'Analyse en Composantes Principales (ACP)</b>	<b>2</b>
1.1	Qu'est-ce que l'ACP ? . . . . .	2
1.2	Historique . . . . .	2
1.3	Pour les programmeurs . . . . .	2
1.4	Pour les statisticiens . . . . .	2
1.5	Pour les mathématiciens de l'optimisation . . . . .	3
1.6	Pour les mathématiciens des opérateurs . . . . .	3
<b>2</b>	<b>Perspective Géométrique</b>	<b>4</b>
2.1	Objectif . . . . .	4
<b>3</b>	<b>Perspective Algorithmique</b>	<b>5</b>
3.1	Algorithme Commun de l'ACP . . . . .	5
<b>4</b>	<b>Transformation ACP</b>	<b>6</b>
<b>5</b>	<b>ACP et Réduction de Dimensionnalité</b>	<b>7</b>
<b>6</b>	<b>Exercices</b>	<b>8</b>
6.1	Exercice 1 . . . . .	8
6.2	Exercice 2 . . . . .	8

# 1 Introduction à l'Analyse en Composantes Principales (ACP)

## 1.1 Qu'est-ce que l'ACP ?

L'Analyse en Composantes Principales (ACP) est une méthode statistique utilisée pour réduire la dimensionnalité des données tout en conservant autant que possible la variance présente dans les données originales. Le terme "principal" se réfère à l'ordre d'importance des composantes, tandis que "composantes" désigne les coefficients ou représentations des données. L'ACP permet une analyse mathématique quantifiable et visuelle des données.

### Intuition

L'ACP cherche à identifier les directions principales dans lesquelles les données varient le plus. Cela se fait en trouvant une base orthonormale qui maximise la variance des données projetées.

## 1.2 Historique

L'ACP a été inventée en 1901 par Karl Pearson comme une analogie du théorème de l'axe principal en mécanique. Elle a été développée indépendamment et nommée par Harold Hotelling dans les années 1930. Selon le domaine d'application, elle est également connue sous le nom de transformation discrète de Karhunen-Loève en traitement du signal, transformation de Hotelling en contrôle de qualité multivarié, décomposition orthogonale propre en ingénierie mécanique, et décomposition en valeurs singulières en algèbre linéaire, entre autres.

### Vulgarisation simple

L'ACP est partout et porte différents noms selon le contexte. Elle est utilisée pour simplifier les données complexes en les réduisant à leurs éléments essentiels.

## 1.3 Pour les programmeurs

Pour un programmeur, l'ACP est un algorithme de réduction de dimensionnalité. Par exemple, en utilisant la bibliothèque `scikit-learn` en Python, l'ACP peut être appliquée pour projeter les données dans un espace de dimension inférieure.

```
1 from sklearn.decomposition import PCA
2 pca = PCA(n_components=2)
3 X_reduced = pca.fit_transform(X)
```

## 1.4 Pour les statisticiens

Pour un statisticien, l'ACP est une méthode pour trouver des représentations décorréées qui expliquent la majorité de la variance dans les données. Elle permet de réduire le nombre de variables tout en préservant l'information essentielle.

## 1.5 Pour les mathématiciens de l'optimisation

Pour un mathématicien de l'optimisation, l'ACP est la solution à un problème de maximisation itératif de type  $L^2$ . Le  $k$ -ème composant peut être trouvé en soustrayant les  $k - 1$  premiers composants principaux de  $\mathbf{X}$ .

$$\hat{\mathbf{X}}_k = \mathbf{X} - \sum_{s=1}^{k-1} \mathbf{X} \mathbf{w}_{(s)} \mathbf{w}_{(s)}^T \quad (1)$$

$$\mathbf{w}_{(k)} = \arg \max_{\|\mathbf{w}\|=1} \left\| \hat{\mathbf{X}}_k \mathbf{w} \right\|^2 \quad (2)$$

## 1.6 Pour les mathématiciens des opérateurs

Pour un mathématicien des opérateurs, l'ACP est la solution à un problème d'approximation min-max sous une énergie de type Dirichlet bornée pour un certain opérateur de corrélation.

### Fiche Récapitulative

**Objectif :** Résumer le concept principal de cette section.

**Principe central :** L'ACP réduit la dimensionnalité des données tout en maximisant la variance expliquée.

**Points essentiels :** - Réduction de dimensionnalité - Maximisation de la variance - Représentation orthonormale

**Formules clés :**

- $\mathbf{X} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T$
- $\mathbf{Y} = \mathbf{U}^T (\mathbf{X} - \mu)$

## 2 Perspective Géométrique

### 2.1 Objectif

L'objectif est de trouver les directions principales intrinsèques des données. Cela implique de trouver une base orthonormale le long des directions principales.

#### Intuition

L'ACP trouve itérativement la direction principale de "dispersion" des données.

## 3 Perspective Algorithmique

### 3.1 Algorithme Commun de l'ACP

Les données  $\mathbf{X} \sim d(\mu, \mathbf{C}, ?)$ , où  $\mathbf{X} \in R^D$ , sont issues d'une distribution avec des statistiques d'ordre connues :

- Moyenne :  $E(\mathbf{X}) = \mu$
- Matrice de covariance :  $E((\mathbf{X} - \mu)(\mathbf{X} - \mu)^T) = \mathbf{C}$

Les "directions principales" sont les vecteurs propres triés de la matrice de covariance.

$$\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T \quad (3)$$

À retenir

Les colonnes de  $\mathbf{U}$  sont les directions principales, et les valeurs propres dans  $\mathbf{\Lambda}$  représentent la variance des données le long de chaque direction.

## 4 Transformation ACP

Réécriture des données dans la nouvelle base ACP :

$$\mathbf{X} = \begin{pmatrix} x_1 \\ \vdots \\ x_D \end{pmatrix} \iff \mathbf{Y} = \begin{pmatrix} y_1 \\ \vdots \\ y_D \end{pmatrix} \quad (4)$$

Le changement de base est effectué par :

$$\mathbf{Y} = \mathbf{U}^T(\mathbf{X} - \mu) \quad (5)$$

## 5 ACP et Réduction de Dimensionnalité

L'ACP est une approximation en termes de  $k$  dans la base ACP :

$$\mathbf{X} \approx \mathbf{Y}_d = \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix}, \quad d \ll D \quad (6)$$

Les composantes principales sont :

$$\mathbf{U}_d = (u_1 \quad \cdots \quad u_d) \quad (7)$$

### Vulgarisation simple

L'ACP simplifie les données en réduisant le nombre de dimensions tout en conservant l'essentiel de l'information.

## 6 Exercices

### 6.1 Exercice 1

Considérons un signal aléatoire  $\varphi \in R^3$  avec une moyenne  $E(\varphi) = 0$  et une matrice d'autocorrélation :

$$\mathbf{R}_\varphi = \begin{pmatrix} 1 - \alpha & -\frac{\alpha}{2} & -\frac{\alpha}{2} \\ -\frac{\alpha}{2} & 1 - \alpha & -\frac{\alpha}{2} \\ -\frac{\alpha}{2} & -\frac{\alpha}{2} & 1 - \alpha \end{pmatrix} \quad (8)$$

Pour un certain scalaire  $0 < \alpha < 1$ .

#### Intuition

La matrice d'autocorrélation est circulante et peut être diagonalisée dans la base de Fourier.

### 6.2 Exercice 2

Considérons un signal aléatoire  $\varphi \in R^3$  avec une moyenne  $E(\varphi) = 0$  et une matrice d'autocorrélation :

$$\mathbf{R}_\varphi = \begin{pmatrix} 1 - \alpha & -\alpha & 0 \\ -\alpha & 1 & -\alpha \\ 0 & -\alpha & 1 - \alpha \end{pmatrix} \quad (9)$$

Pour un certain scalaire  $0 < \alpha < 1$ .

#### À retenir

La matrice d'autocorrélation n'est pas circulante et nécessite un calcul manuel des valeurs propres et vecteurs propres.