

Feuille de TD 6 : Apprentissage non-supervisé

Exercice 1 (Réduction de dimension : Analyse en composantes principales)

On dispose d'un tableau de données $X \in \mathbb{R}^{n \times p}$ avec possiblement $p > n$. On supposera que les colonnes de la matrice $X = \begin{bmatrix} X_1^\top \\ \vdots \\ X_n^\top \end{bmatrix}$ sont centrées. Une technique classique pour gérer ce type de situation est de projeter les $X_i \in \mathbb{R}^p$ sur un espace vectoriel V de petite dimension d . Pour ne pas perdre trop d'information, il est nécessaire que les projections des X_i sur V approximent "au mieux" les X_i . On va rechercher des espaces $V_d = \text{Vect}\{\alpha^{(1)}, \dots, \alpha^{(d)}\}$ tels que

$$\sum_{i=1}^n \|X_i - \text{Proj}_{V_d}(X_i)\|^2$$

est minimal et les $\alpha^{(1)}, \dots, \alpha^{(d)} \in \mathbb{R}^p$ sont de norme 1 et orthogonaux entre eux.

1. On s'intéresse d'abord au cas où $d = 1$. Montrez que chercher $\alpha \in \mathbb{R}^p$ de norme 1 minimisant

$$\sum_{i=1}^n \|X_i - \text{Proj}_{\langle \alpha \rangle}(X_i)\|^2$$

revient à chercher $\alpha \in \mathbb{R}^p$ maximisant

$$\max_{\|\alpha\|=1} \alpha^\top (X^\top X) \alpha.$$

2. Comment obtient-on $\alpha^{(1)}$?
3. Pour $d \leq \text{rang}(X)$, comment obtient-on $\alpha^{(1)}, \dots, \alpha^{(d)}$?
4. Montrer que les C_j définis par $C_j = X \alpha^{(j)}$ sont orthogonaux entre eux.
5. En notant $C_j[i]$ la i -ème composante de C_j , montrer l'égalité :

$$\text{Proj}_{V_d}(X_i) = \sum_{j=1}^d C_j[i] \alpha^{(j)}.$$

6. Exprimer en fonction des valeurs propres de $X^\top X$ le rapport

$$\frac{\sum_{i=1}^n \|\text{Proj}_{V_k}(X_i)\|^2}{\sum_{i=1}^n \|X_i\|^2}.$$