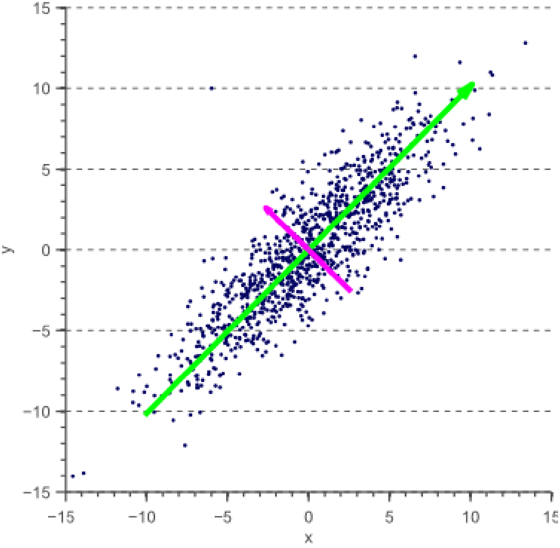
Machine Learning

Réduction de la dimensionalité : Projections linéaires

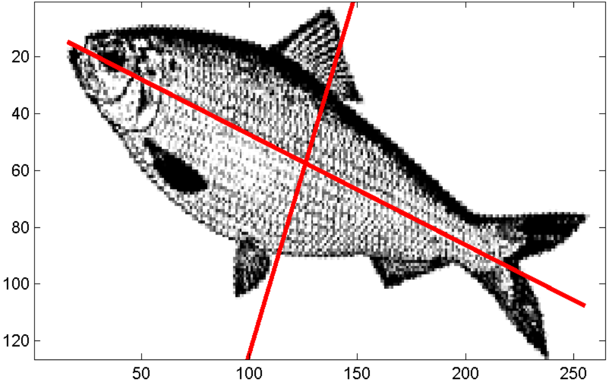
Réduction de la dimensionalité : Projections linéaires

- Principal Component Analysis (Non-supervisée)
- Linear Discriminant Analysis (Supervisée)

Réduction de la dimensionalité : PCA



Réduction de la dimensionalité : PCA



Réduction de la dimensionalité : PCA

$$X = \begin{bmatrix} X_{1,1} & \dots & X_{1,D} \\ \vdots & \ddots & \vdots \\ X_{N,1} & \dots & X_{N,D} \end{bmatrix}$$

Réduction de la dimensionalité : PCA

chaque dimension est centrée (et réduite) :

$$\bar{X} = \begin{bmatrix} X_{1,1} - \bar{X}_1 & \dots & X_{1,D} - \bar{X}_D \\ \vdots & \ddots & \vdots \\ X_{N,1} - \bar{X}_1 & \dots & X_{N,D} - \bar{X}_D \end{bmatrix}$$

ou

$$\tilde{X} = \begin{bmatrix} \frac{X_{1,1} - \bar{X}_1}{\sigma(X_1)} & \dots & \frac{X_{1,D} - \bar{X}_D}{\sigma(X_D)} \\ \vdots & \ddots & \vdots \\ \frac{X_{N,1} - \bar{X}_1}{\sigma(X_1)} & \dots & \frac{X_{N,D} - \bar{X}_D}{\sigma(X_D)} \end{bmatrix}$$

Réduction de la dimensionalité : PCA

Matrice de covariance (resp. corrélation) :

$$\frac{1}{N} * \bar{X}^T * \bar{X}\,,\ (\frac{1}{N} * \tilde{X}^T * \tilde{X})$$

ACP :
Retrouver les valeurs et vecteurs propres de de la matrice de covariance (resp. corrélation), donc diagonaliser la matrice carrée obtenue.

Vecteur propre : vecteur permettant de projeter les données
Valeur propre : “proportion d’information” conservée par la projection suivant le vecteur propre correspondant
Réduction de dimension : On ne projette que suivant le nombre de vecteurs propres voulus

Réduction de la dimensionalité : LDA

Linear Discriminant Analysis