Exercice 3

Soient $X = (x_1, \dots, x_n)$ et $Y = (y_1, \dots, y_n)$ deux vecteurs dans $\mathbb{R}^n$

En supposant que les x_i ne sont pas tous egaux,

montrer que les coefficients a et b qui minimisent la quantité

$\Delta(a,b) = \|Y - aX - b \cdot 1\|^2$ où $1 = (1, 1, \dots, 1)$

sont donnés par

$$a = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{n(\sum x_i^2) - (\sum x_i)^2} \quad \text{et} \quad b = \frac{(\sum x_i^2)(\sum y_i) - (\sum x_i)(\sum x_i y_i)}{n(\sum x_i^2) - (\sum x_i)^2}$$

a) en interprétant la question comme un problème de projection orthogonale sur le plan engendré par X et 1 ;

b) en calculant les points critiques de la fonction de 2 variable $\Delta(a,b)$

Solution

moindre carrés et projection orthogonale

L'idée est de considérer les données comme deux vecteurs de $\mathbb{R}^n$ $X = (x_1, x_2, \dots, x_n)$ et $Y = (y_1, y_2, \dots, y_n)$ et un troisième vecteur utile est $V = (1, 1, \dots, 1)$, alors le vecteur $Z = Y - aX - bV$ a pour coordonnées $Z_i = y_i - a x_i - b$

En particulier, la norme euclidienne au carré de Z

$$\|Z\|^2 = \sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} (y_i - a x_i - b)^2 = \Sigma(a,b) \text{ ou } = \Delta(a,b)$$

est précisement la quantité que l'on cherche à minimiser

on remarque que le vecteur $aX + bV$ est dans le plan P engendré par X et V, P est un plan lorsque X n'est pas proportionnel à V, i.e les x_i ne sont pas tous egaux

Chebib Mouzher

NB

dans le cas où les x_i ne sont pas tous egaux, le problème de la regression linéaire a une interprétation géométrique claire dans $\mathbb{R}^n$; il s'agit de trouver dans le plan P, le vecteur $aX + bV$ le plus proche de Y

Les coefficients a et b sont tels que le vecteur $aX+bV$ est la projection orthogonale de Y sur le plan P engendré par X et V

a et b sont les solutions du système linéaire à deux équations

$$\begin{cases} \langle aX+bV, X\rangle = \langle y, X\rangle \\ \langle aX+bV, V\rangle = \langle y, V\rangle \end{cases} \Leftrightarrow \begin{cases} a\|X\|^2 + b\langle V, X\rangle = \langle y, X\rangle \\ a\langle X, V\rangle + b\|V\|^2 = \langle y, V\rangle \end{cases}$$

avec $\|X\|^2 = \sum x_i^2$, $\langle X, V\rangle = \sum x_i$, $\|V\|^2 = n$

$$\langle y, X\rangle = \sum x_i y_i \text{ et } \langle y, V\rangle = \sum y_i$$

on a donc $\begin{pmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}$

$$\Leftrightarrow \begin{pmatrix} a \\ b \end{pmatrix} = \frac{1}{n\sum x_i^2 - (\sum x_i)^2}\begin{pmatrix} n & -\sum x_i \\ -\sum x_i & \sum x_i^2 \end{pmatrix}\begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}$$

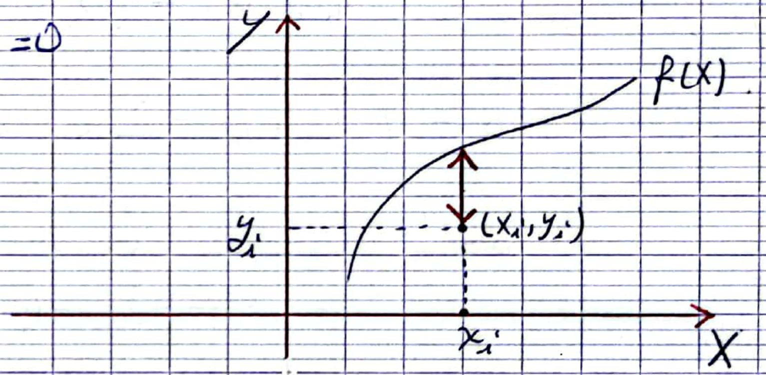
Rappel Méthode de moindres carrés (m.c)

Soit $y = f_{\theta_1, \theta_2}(X)$ une equation d'une courbe qui dépend des deux paramètres θ_1 et θ_2, selon la méthode de m.c. la courbe d'ajustement à un ensemble des points (x_i, y_i) $i = 1, \dots, n$; est celle qui minimise la somme des carrés des distances (verticales) aux divers points

Le problème revient alors à chercher les valeurs θ_1 et θ_2 qui minimisent la fonction : $d = \sum (y_i - f(x_i))^2$

Ces valeurs sont les solutions des equations normales suivantes

$$\frac{\partial d}{\partial \theta_1} = 0 \quad \text{et} \quad \frac{\partial d}{\partial \theta_2} = 0$$



b) en calculant les points critiques de la fonction de 2 variable $\Delta(a,b)$

$$\Delta(a,b) = \|y - aX - b \cdot 1\|^2 = \sum_{i=1}^{n} (y_i - a x_i - b)^2$$

$$\frac{\partial \Delta}{\partial a} = 0 \quad \text{et} \quad \frac{\partial \Delta}{\partial b} = 0$$

$$\frac{\partial \Delta}{\partial a} = 0 \Leftrightarrow 2 \sum (y_i - a x_i - b)(-x_i) = 0 \Leftrightarrow \sum x_i (y_i - a x_i - b) = 0$$

$$\frac{\partial \Delta}{\partial b} = 0 \Leftrightarrow 2 \sum (y_i - a x_i - b)(-1) = 0 \Leftrightarrow \sum (y_i - a x_i - b) = 0$$

chebil rouyher

$$\sum x_i y_i - a \sum x_i^2 - b \sum x_i = 0 \Leftrightarrow \left(\sum x_i^2\right) a + \left(\sum x_i\right) b = \sum x_i y_i$$

$$\sum y_i - a \sum x_i - n b = 0 \Leftrightarrow \left(\sum x_i\right) a + n b = \sum y_i$$

$$\begin{cases} \left(\sum x_i^2\right) a + \left(\sum x_i\right) b = \sum x_i y_i \\ \sum x_i \, a + n \, b = \sum y_i \end{cases}$$

Scanné avec CamScanner

$$\begin{array}{l} \times n \\ (-\Sigma x_i) \end{array} \left\{ \begin{array}{l} (\Sigma x_i^2) a + (\Sigma x_i) b = \Sigma x_i y_i \\ \Sigma x_i a + n b = \Sigma y_i \end{array} \right.$$

$$\left\{ \begin{array}{l} (n \Sigma x_i^2) a + n \Sigma x_i b = n \Sigma x_i y_i \\ -(\Sigma x_i)^2 a - n \Sigma x_i b = - \Sigma x_i \Sigma y_i \end{array} \right.$$

$$a \left(n \Sigma x_i^2 - (\Sigma x_i)^2 \right) = n \Sigma x_i y_i - (\Sigma x_i)(\Sigma y_i)$$

$$a = \frac{n \Sigma x_i y_i - (\Sigma x_i)(\Sigma y_i)}{n \Sigma x_i^2 - (\Sigma x_i)^2}$$

$$b = \frac{1}{n} \left[\Sigma y_i - \Sigma x_i a \right] = \frac{1}{n} \left[\Sigma y_i - \Sigma x_i \left(\frac{n \Sigma x_i y_i - (\Sigma x_i)(\Sigma y_i)}{n \Sigma x_i^2 - (\Sigma x_i)^2} \right) \right]$$

$$b = \frac{1}{n} \left[\frac{n \Sigma x_i^2 \Sigma y_i - (\Sigma x_i)^2 \Sigma y_i - n \Sigma x_i \Sigma x_i y_i + (\Sigma x_i)^2 (\Sigma y_i)}{n \Sigma x_i^2 - (\Sigma x_i)^2} \right]$$

$$b = \frac{1}{n} \left[\frac{n \Sigma x_i^2 \Sigma y_i - n \Sigma x_i \Sigma x_i y_i}{n \Sigma x_i^2 - (\Sigma x_i)^2} \right.$$

$$b = \frac{\Sigma x_i^2 \Sigma y_i - \Sigma x_i \Sigma x_i y_i}{n \Sigma x_i^2 - (\Sigma x_i)^2}$$

Scanné avec CamScanner