Ajustement exponentiel par descente de gradient

On cherche à ajuster les données par une fonction exponentielle :

$$f(x) = ae^{bx}$$

Les données sont :

$$x : -5,\ -4,\ -3,\ -2,\ -1,\ 0,\ 1,\ 2,\ 3,\ 4,\ 5,\ 6$$

$$y : 0.02,\ 0.03,\ 0.06,\ 0.04,\ 0.1,\ 0.08,\ 0.14,\ 0.15,\ 0.2,\ 0.21,\ 0.27,\ 0.35$$

(1) Définition de la fonction $D(a, b)$

On veut minimiser l'erreur entre les données réelles et la fonction. On utilise la somme des carrés des erreurs :

$$D(a, b) = \sum_{i=1}^{n} (y_i - f(x_i))^2$$

Or :

$$f(x_i) = ae^{bx_i}$$

Donc :

$$D(a, b) = \sum_{i=1}^{n} (y_i - ae^{bx_i})^2$$

C'est la fonction d'erreur que l'on veut minimiser.

(2) Calcul du gradient

On veut calculer :

$$D_a(a, b) = \frac{\partial D}{\partial a} \qquad D_b(a, b) = \frac{\partial D}{\partial b}$$

On part de :

$$D(a, b) = \sum (y_i - ae^{bx_i})^2$$

Posons :

$$E_i = y_i - ae^{bx_i}$$

Alors :

$$D = \sum E_i^2$$

Dérivée par rapport à a

Rappel :

$$\frac{d}{du}(u^2) = 2u$$

Donc :

$$\frac{\partial D}{\partial a} = \sum 2E_i \frac{\partial E_i}{\partial a}$$

Or :

$$E_i = y_i - ae^{bx_i}$$

Donc :

$$\frac{\partial E_i}{\partial a} = -e^{bx_i}$$

Ainsi :

$$\frac{\partial D}{\partial a} = \sum 2(y_i - ae^{bx_i})(-e^{bx_i})$$

On obtient :

$$D_a(a,b) = -2\sum(y_i - ae^{bx_i})e^{bx_i}$$

Forme équivalente :

$$D_a(a,b) = 2\sum(ae^{bx_i} - y_i)e^{bx_i}$$

Dérivée par rapport à b

On utilise :

$$\frac{\partial D}{\partial b} = \sum 2E_i \frac{\partial E_i}{\partial b}$$

Or :

$$E_i = y_i - ae^{bx_i}$$

Rappel de dérivation :

$$\frac{d}{db}(e^{bx_i}) = x_i e^{bx_i}$$

Donc :

$$\frac{\partial E_i}{\partial b} = -ax_i e^{bx_i}$$

Ainsi :

$$\frac{\partial D}{\partial b} = \sum 2(y_i - ae^{bx_i})(-ax_i e^{bx_i})$$

On obtient :

$$D_b(a,b) = -2\sum(y_i - ae^{bx_i})ax_i e^{bx_i}$$

Forme équivalente :

$$D_b(a,b) = 2\sum(ae^{bx_i} - y_i)ax_i e^{bx_i}$$

(3) Méthode de descente du gradient

Principe :

1. Choisir des valeurs initiales a_0 et b_0.
2. Calculer le gradient (D_a, D_b).
3. Mettre à jour les paramètres :

$$a_{n+1} = a_n - \alpha D_a$$

$$b_{n+1} = b_n - \alpha D_b$$

4. Répéter jusqu'à convergence.

Le paramètre α est le pas d'apprentissage.

(4) Résultat numérique

Conditions :

$$a_0 = 0.2, \quad b_0 = 0.1, \quad \alpha = 0.001$$

Après convergence :

$$a \approx 0.0963$$

$$b \approx 0.2128$$

Fonction ajustée :

$$f(x) \approx 0.0963\, e^{0.2128x}$$

(5) Commande gnuplot

```
f(x)=0.0963*exp(0.2128*x)
plot "data.txt" with points, f(x) with lines
```