2019 Fall MA 511
Ronglong Fang

Problem 1. Given data points $(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N) \in \mathbb{R} \times \mathbb{R}$, find the least squares linear regression function to fit them. That is, find $a, b \in \mathbb{R}$ to minimize $\sum_{j=1}^{N}(ax_j + b - y_j)^2$.

Proof. Let $f(a,b) = \sum_{j=1}^{N}(ax_j + b - y_j)^2$, our goal is to find a, b that minimize f. So we need to solve the equation.

$$\begin{cases} \frac{\partial f}{\partial a}(a,b) = 0 \\ \frac{\partial f}{\partial b}(a,b) = 0 \end{cases}$$

That is

$$\begin{cases} \sum_{j=1}^{N} 2 * (ax_j + b - y_j) * x_j = 0 \\ \sum_{j=1}^{N} 2 * (ax_j + b - y_j) = 0 \end{cases}$$

which can be simplified as

$$\begin{cases} a * \sum_{j=1}^{N} x_j^2 + b * \sum_{j=1}^{N} x_j - \sum_{j=1}^{N} y_j x_j = 0 \\ a * \sum_{j=1}^{N} x_j + Nb - \sum_{j=1}^{N} y_j = 0 \end{cases}$$

we can get $a = \frac{N\sum_{j=1}^{N} x_j y_j - \sum_{j=1}^{N} y_j x_j}{N\sum_{j=1}^{N} x_j^2 - (\sum_{j=1}^{N} x_j)^2}$, and $b = \frac{\sum_{j=1}^{N} y_j}{N} - a * \frac{\sum_{j=1}^{j=1} x_j}{N}$ □

Problem 2 (*)**.** Given data points $(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \ldots (\boldsymbol{x}_N, y_N) \in \mathbb{R}^d \times \mathbb{R}$, find the least squares linear regression function to fit them. That is, find $\boldsymbol{a} \in \mathbb{R}^d$ and $b \in \mathbb{R}$ to minimize $\sum_{j=1}^{N}(\boldsymbol{a} \cdot \boldsymbol{x}_j + b - y_j)^2$, where $\boldsymbol{s} \cdot \boldsymbol{t} = \sum_{j=1}^{d} s_j t_j$.

Proof. This problem is similar as the first problem, we can use the same method. Let

$$f(\boldsymbol{a}, b) = \sum_{j=1}^{N}(\boldsymbol{a} \cdot \boldsymbol{x_j} + b - y_j)^2$$

We need calculate the derivative of f.

$$\begin{cases} \triangledown_{\boldsymbol{a_1}} f = \sum_{j=1}^{N}(\boldsymbol{a} \cdot \boldsymbol{x_j} + b - y_j) \cdot \boldsymbol{x_j}^{(1)} = 0 \\ \triangledown_{\boldsymbol{a_2}} f = \sum_{j=1}^{N}(\boldsymbol{a} \cdot \boldsymbol{x_j} + b - y_j) \cdot \boldsymbol{x_j}^{(2)} = 0 \\ \vdots \\ \triangledown_{\boldsymbol{a_N}} f = \sum_{j=1}^{N}(\boldsymbol{a} \cdot \boldsymbol{x_j} + b - y_j) \cdot \boldsymbol{x_j}^{(N)} = 0 \\ \triangledown_b f = \sum_{j=1}^{N} \boldsymbol{a} \cdot \boldsymbol{x_j} + b - y_j = 0 \end{cases}$$

where $\triangledown_t$ is the derivative of t. We can rewrite the equation with the matrix form.

$$\begin{pmatrix} \sum_{j=1}^{N} \mathbf{x_j}\mathrm{T} \cdot \boldsymbol{x_j}^{(1)} & \sum_{j=1}^{N} \boldsymbol{x_j}^{(1)} \\ \sum_{j=1}^{N} \mathbf{x_j}\mathrm{T} \cdot \boldsymbol{x_j}^{(2)} & \sum_{j=1}^{N} \boldsymbol{x_j}^{(1)} \\ \vdots & \\ \sum_{j=1}^{N} \mathbf{x_j}\mathrm{T} \cdot \boldsymbol{x_j}^{(N)} & \sum_{j=1}^{N} \boldsymbol{x_j}^{(N)} \\ \sum_{j=1}^{N} \mathbf{x_j}\mathrm{T} & \sum_{j=1}^{N} 1 \end{pmatrix} \begin{pmatrix} a^{(1)} \\ a^{(2)} \\ \vdots \\ a^{(N)} \\ b \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{N} \boldsymbol{x_j}^{(1)} \cdot y_j \\ \sum_{j=1}^{N} \boldsymbol{x_j}^{(2)} \cdot y_j \\ \vdots \\ \sum_{j=1}^{N} \boldsymbol{x_j}^{(N)} \cdot y_j \\ \sum_{j=1}^{N} y_j \end{pmatrix}$$

□

let $\mathbf{X}$ denote the left matrix, where $\mathbf{X}$ is the size of $(N+1) \times (N+1)$; let the $\widetilde{\boldsymbol{a}}$ denote the left vector, thus the size is $(N+1) \times 1$; let $\mathbf{b}$ denote the right vector, and the size is also $(N+1) \times 1$. We can use computer to solve the equation, the solution is $\widetilde{\boldsymbol{a}} = X^{-1} \cdot \boldsymbol{b}$, where X^{-1} is the inverse of the matrix X.