LINEAR REGRESSION

Đạt Nguyễn Ngọc | 11200745

October 2022

1 Problems

1.1 Biến đổi lại linear regression trên lớp ra latex, từ $t = y(x, w) + noise \Rightarrow w = (X^T X)^{-1} X^T t$

Solution

Suppose that the observations are drawn independently from a Gaussian distribution

$$t = y(x, w) + N(0, \sigma^2) = N(y(x, w), \sigma^2)$$

With $\beta^{-1} = \frac{1}{\sigma^2}$

$$p(t|x, w, \beta) = \prod_{n=1}^{N} N(t_n|y(x_n, w), \beta^{-1})$$

It is convenient to maximize the logarithm of the likelihood function:

$$\begin{aligned}
logp(t|x, w, \beta) &= \sum_{n=1}^{N} log(N(t_n|y(x_n, w), \beta^{-1})) \\
&= \sum_{n=1}^{N} log(\frac{1}{\sqrt{2\pi\beta^{-1}}} e^{\frac{-(t_n - y(x_n, w))^2 \beta}{2}}) \\
&= \sum_{n=1}^{N} (\frac{-1}{2} log(2\pi\beta^{-1}) - (t_n - y(x_n, w))^2 - \frac{\beta}{2}) \\
&= -\sum_{n=1}^{N} (t_n - y(x_n - w))^2
\end{aligned}$$

Minimize $\sum_{n=1}^{N} (t_n - y(x_n - w))^2$

Set: $L = \frac{1}{N} \sum_{n=1}^{N} (t_n - y(x_n - w))^2$

with $\mathbf{x} = \begin{pmatrix} 1 & x_1 \\ 2 & x_2 \\ 3 & x_3 \\ ... & ... \\ 1 & x_n \end{pmatrix}, \mathbf{w} = \begin{pmatrix} w_0 \\ w_1 \end{pmatrix}, \mathbf{y} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ ... \\ y_n \end{pmatrix} = \begin{pmatrix} w_1x_1 + w_0 \\ w_1x_2 + w_0 \\ w_1x_3 + w_0 \\ ... \\ w_1x_n + w_0 \end{pmatrix} =$
xw

$\mathbf{t} - \mathbf{y} = \begin{pmatrix} t_1 - y_1 \\ t_2 - y_2 \\ t_3 - y_3 \\ ... \\ t_n - y_n \end{pmatrix}$

$\Rightarrow$||t - xw||$^2 = (t_1 - y_1)^2 + ... + (t_n - y_n)^2 = \sum_{i=1}^{N}(t_i - y_i)^2 = L$

$\Rightarrow L = ||t - y||_2^2 = ||t - xw||_2^2$

$\Rightarrow \frac{\partial(L)}{\partial(w)} = 2X^T(t - Xw) = 0$

$\Rightarrow X^Tt = X^TXw(X^TXisinvertable)$

$\Rightarrow w = (X^TX)^{-1}X^Tt$

1.2 Proof that X^TX invertible when X full rank.

Solution

The key observation is that for $v \in R^m, Xv = 0$ if and only if $X^TXv = 0$. For the non-trivial implication if $X^TXv = 0$ then $v^TX^TXv = 0$ that is $(Xv)^TXv = 0$ which implies that $Xv = 0$

If the rank of X is m, this mean that X is independent. So by the observation, X^TX is independent, which makes it invertable (as it is square)

2 Coding

2.1 Viết code numpy, tìm model linear regression cho bai toán dữ đoán giá nhà, với dataset này

2.1.1 Vẽ model dự đoán (đường thẳng) và dữ liệu (point - scatter).

2.1.2 Dự đoán giá các căn nhà có diện tích 50, 100, 150.

2.2 Viết code numpy, tìm model linear regression cho bai toán dữ đoán giá nhà với dataset này

Solution

Code file solution for part 2