linear regression

Nguyen Thi Hoai Linh

September 2021

1 EX1:

We have a data set of observations x = $(x_1, x_2, ..., x_N)^T$
Representing N observations of the scalar variable x and their corresponding target values t = $(t_1, t_2, ..., t_N)^T$
Suppose that the observations are drawn independently from a Gaussian distribution. Data points that are drawn independently from the same distribution are said to be independent and identically distributed
t = y(x, w) + N(0, β^{-1})$t = N(y(x,w), \beta^{-1})$
Precision parameter $\beta^{-1} = \sigma^2$
p(t|x, w, β) $= N(t|y(x,w), \beta^{-1})$
the likelihood function:
p(t|x, w, β) $= \prod_{n=1}^{N} N(t_n|y(x_n,w), \beta^{-1}) = \frac{1}{\sqrt{2\pi\beta^{-1}}} e^{\frac{-(t-y(x,w))^2\beta}{2}}$
it is convenient to maximize the logarithm of the likelihood function

$$log \prod_{n=1}^{N} N(t_n|y(x_n,w), \beta^{-1}) = \sum_{n=1}^{N} log\left(\frac{1}{\sqrt{2\pi\beta^{-1}}} e^{\frac{-(t_n-y(x_n,w))^2\beta}{2}}\right)$$

$$= log(2\pi\beta^{-1})^{\frac{N}{2}} + log(e^{-\frac{\beta}{2}\sum_{n=1}^{N} N(t_n|y(x_n,w))^2)})$$

$$= -\frac{N}{2} log 2\pi\beta^{-1} - \frac{\beta}{2}\sum_{n=1}^{N}(t_n - y(x_n,w))^2 = -\frac{N}{2} log 2\pi + \frac{N}{2} log\beta - \frac{\beta}{2}\sum_{n=1}^{N}(t_n - y(x_n,w))^2$$

max log p(t|x,w,β) $= -max\frac{\beta}{2}\sum_{n=1}^{N}(t_n - y(x_n,w))^2$

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

L=$\sum_{n=1}^{N}(t_n - y(x_n,w))^2$
We have:
y$(x_n, w) = w_1 x_n + w_0$

$$x = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ ... & ... \\ 1 & x_n \end{bmatrix}, t = \begin{bmatrix} t_1 \\ t_2 \\ ... \\ t_n \end{bmatrix}, t = \begin{bmatrix} w_0 \\ w_1 \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{bmatrix} = \begin{bmatrix} w_1 x_1 + w_0 \\ w_1 x_2 + w_0 \\ .... \\ w_1 x_n + w_0 \end{bmatrix} = xw$$

$\Rightarrow t - y = \begin{bmatrix} t_1 - y_1 \\ t_2 - y_2 \\ .... \\ t_n - y_n \end{bmatrix}$
$\Rightarrow ||t-y||^2 = (t_1 - y_1)^2 + (t_2 + y_2)^2 + ... + (t_n - y_n)^2 = \sum_{n=1}^{N} (t_n - y_n)^2$
$\Rightarrow L = ||t-y||_2^2 = ||t - xw||_2^2$

$$\frac{\partial L}{\partial w} = 2x(t - xw) = 0$$

$\Rightarrow x^T t = x^T x w$
$\Rightarrow w = (x^T x)^{-1} x^T t$