ML WEEK4

phuong nguyen

October 2021

1 Posterior Bien Doi

Follow the flow of Bayes Theorem, we get:

$I = P(w/D) = p(w/x, t, \alpha, \beta) = \frac{p(t/x,w,\beta)p(w/x)}{p(x,t,\alpha,\beta)}$

I now is called Posteriror, we concentate on A,B :

$B = p(t/x, w, \beta) = \prod N(t_i/y(x_i, w), \beta^{-1})$

Suppose A is normal distribution :

$A = p(w/x) = N(w, 0, \alpha^{-1})$

We expect to max I <=> we want to max A.B because I depends on w, getting max of A.B <=> get min of -(A.B)

$A = \frac{1}{\sqrt{(2\pi)^D.|\alpha^{-1}|}} e^{\frac{-1}{2} w^T(\alpha^{-1}.I_n)^{-1} w}$

Since $(\alpha^{-1}.I_n)^{-1} = (\frac{1}{\alpha} I_n)^{-1} = \alpha I_n$

$A = \frac{1}{\sqrt{(2\pi)^D.|\alpha^{-1}|}} e^{\frac{-1\alpha}{2} w^T w}$

$=> lnA = \frac{-\alpha}{2} w^T w$

$B = \prod N(t_i/y(x_i, w), \beta^{-1})$

$B = \frac{1}{(\beta^{-1}\sqrt{2\pi})^N} e^{\sum \frac{-\beta}{2}(t_i - y(x_i,w))^2}$

$=> lnB = \sum \frac{-\beta}{2}(t_i - y(x_i, w))^2$

Finding min of -(A.B) by taking logarit of them.

min=ln(-(A.B))=-(lnA+lnB)= $\frac{\alpha}{2} w^T w + \sum \frac{\beta}{2}(t_i - y(x_i, w))^2$

min=$\frac{\alpha}{\beta}w^T w$+$\sum(t_i - y(x_i, w))^2$

let $\frac{\alpha}{\beta} = \lambda$, X be matrix with entries are data and one more column filled by one and t be target matrix, w is cofficients matrix, then:

min=$\|Xw - t\|_2^2$+$\|w\|_2^2$

Using matrix calculus, we get: $\frac{d(min)}{d(w)} = 2X^T.(Xw - t) + 2\lambda w$

$\frac{d(min)}{d(w)} = 0 <=> X^T Xw - X^T t + \lambda w = 0 <=> w = (X^T.X + \lambda I_n)^{-1}.X^T t$