30/04/2019 Inferência e aprendizagem

Giovanni Amorim

1. Questão 3 da prova.

   (a) $\bar{g}(x) = E_D(g^{(D)}(x))$
   O modelo linear se ”encaixa”nos dois pontos, formando a reta que os contém:

$$y = mx + k$$

$$m = \frac{y_1 - y_2}{x_1 - x_2}$$

$$y_1 = \frac{y_1 - y_2}{x_1 - x_2}x_1 + k$$

$$k = y_1 - \frac{y_1 - y_2}{x_1 - x_2}x_1$$

$$k = \frac{x_1y_2 - y_1x_2}{x_1 - x_2}$$

   Logo:

$$g^{(D)}(x) = \frac{y_1 - y_2}{x_1 - x_2}x + \frac{x_1y_2 - y_1x_2}{x_1 - x_2}$$

$$\bar{g}(x) = E_D[\frac{y_1 - y_2}{x_1 - x_2}x + \frac{x_1y_2 - x_2y_1}{x_1 - x_2}]$$

$$g^{(D)}(x) = (\frac{1}{2})^2 \int_{-1}^{1}\int_{-1}^{1} \frac{x_1^2 - x_2^2}{x_1 - x_2} x dx_1 dx_2 + (\frac{1}{2})^2 \int_{-1}^{1}\int_{-1}^{1} \frac{x_1x_2^2 - x_2x_1^2}{x_1 - x_2} dx_1 dx_2$$

$$g^{(D)}(x) = \frac{1}{4}\int_{-1}^{1}\int_{-1}^{1} [(x_1 + x_2)x + x_1x_2] dx_1 dx_2$$

$$g^{(D)}(x) = 0.x + 0 = 0$$

   (b) $E_{out} = E_x[(g(x) - f(x))^2]$
   Considerando

$$g(x) = mx + k$$

   :

$$E_{out} = E_x[(mx + k - x^2)^2]$$

$$E_{out} = E_x(x^4) - 2kE_x(x^3) + (k^2 - 2m)E_x(x^2) + 2kmE_x(x) + m^2$$

$$E_{out} = \frac{1}{2}\int_{-1}^{1} x^4 dx - 2m\frac{1}{2}\int_{-1}^{1} x^3 dx + (m^2 - 2k)\frac{1}{2}\int_{-1}^{1} x^2 dx + 2mk\frac{1}{2}\int_{-1}^{1} x dx + k^2$$

$$E_{out} = \frac{1}{2}.\frac{2}{5} - m.\frac{0}{4} + (m^2 - 2k)\frac{1}{2}.\frac{2}{3} + 2mk\frac{1}{2}.\frac{0}{2} + k^2$$

$$E_{out} = \frac{1}{5} + \frac{(m^2 - 2k)}{3} + k^2$$

Agora, os valores de m e k podem ser definidos pelos pontos $(x_1, y_1), (x_2, y_2)$:

$$m = \frac{y_1 - y_2}{x_1 - x_2}$$

$$m = \frac{x_1^2 - x_2^2}{x_1 - x_2}$$

$$m = x_1 + x_2$$

$$k = \frac{x_1 y_2 - y_1 x_2}{x_1 - x_2}$$

$$k = \frac{x_1 x_2^2 - x_1^2 x_2}{x_1 - x_2}$$

$$k = x_1 x_2$$

Logo:

$$E_D[E_{out}] = E_D[\frac{1}{5} + \frac{(m^2 - 2k)}{3} + k^2]$$

$$E_D[E_{out}] = \frac{1}{5} + \frac{1}{3}E_D[(x_1 + x_2)^2 + 2x_1x_2] + E_D[x_1^2 x_2^2]$$

$$E_D[E_{out}] = \frac{1}{5} + \frac{1}{3}.(\frac{1}{2^2}).\int_{-1}^{1}\int_{-1}^{1}(x_1^2 + x_2^2 + 4x_1x_2)dx_1dx_2 + (\frac{1}{2^2})\int_{-1}^{1}\int_{-1}^{1} x_1^2 x_2^2 dx_1 dx_2$$

$$E_D[E_{out}] = \frac{1}{5} + \frac{1}{3}.\frac{1}{4}.\frac{8}{3} + \frac{1}{4}.\frac{4}{9} = \frac{8}{15}$$

$$bias(x) = (\bar{g}(x) - f(x))^2 = f(x)^2 = x^4$$

$$E_x(bias(x)) = E_x(x^4) = \frac{1}{2}\int_{-1}^{1} x^4 dx = \frac{1}{5}$$

Finalmente, sabendo que: $E_{out}(x) = bias(x) + var(x)$

$$\frac{8}{15} = \frac{1}{5} + var(x)$$

$$var(x) = \frac{1}{3}$$

2. Questão 5 da prova.
   A partir de uma abordagem Bayesiana, começamos com:

$$P(w|y,X) = \frac{P(w,y|X)}{P(y|X)}$$

   Logo, podemos maximizar $P(w,y|X)$:

$$max_w P(w,y|X) = max_w P(y|X)P(w|y,X)$$

$$\hat{w} = argmax_w[\prod_{n=1}^{N} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y_n - (\sum_{i=0}^{d} w_i x^i))^2}{2\sigma^2}} + \prod_{j=0}^{d} \frac{1}{\alpha\sqrt{2\pi}} e^{-\frac{w_j^2}{2\alpha^2}}]$$

   Aplicando agora a função log (não modifica a lógica da maximização):

$$\hat{w} = argmax_w \frac{-1}{2\sigma^2}[\sum_{n=1}^{N}(y_n - (w_0 + w_1 x_{n,1} + w_2 x_{n,2} + ... + w_d x_{n,d}))^2 + \frac{\sigma^2}{\alpha^2}\sum_{j=0}^{d} w_j^2]$$

$$\hat{w} = argmin_w[(y - Xw)^T(y - Xw) + \frac{\sigma^2}{\alpha^2} w^T w$$

$$\hat{w} = argmin_w ||y - Xw||^2 + \lambda||w||^2$$

   Obviamente, o termo no qual chegamos é o mesmo que se encontra na minimização do E_{aug} apresentado.
   Dessa forma, fica simples de perceber que a Regressão Ridge, que é basicamente a introdução de um termo de regularização, pode ser interpretada como uma ”desvalorização”do conhecimento a priori da distribuição do seu parâmetro w.
   A escolha de um λ maior automaticamente resulta em uma escolha de α menor, que se traduz em menor valor de variância do w a priori. Ou seja, 'o que definimos anteriormente como um w parâmetro inicial' ganha mais importância.
   Por outro lado, escolher um valor reduzido de λ tem como consequência um aumento em α: nosso w a priori tem variância maior e, por isso, perde importância no resultado do modelo.
   Concluímos que o fator adicionado λ pode ser interpretado como uma medição de quão 'bayesiano' você quer ser em seu modelo.