Machine Learning Worksheet 06

Linear Regression

1 Ridge regression

Problem 1: Show that the following holds: The ridge regression estimates can be obtained by ordinary least squares regression on an augmented dataset: Augment the design matrix $\boldsymbol{\Phi}$ with p additional rows $\sqrt{\lambda}\boldsymbol{I}$ and augment $\boldsymbol{z}$ with p zeros.

2 Bayesian Linear Regression

Problem 2: We have seen that, as the size of a data set increases, the uncertainty associated with the posterior distribution over model parameters decreases (see tower equalities). Prove the following matrix identity

$$(\boldsymbol{M} + \boldsymbol{v}\boldsymbol{v}^T)^{-1} = \boldsymbol{M}^{-1} - \frac{(\boldsymbol{M}^{-1}\boldsymbol{v})(\boldsymbol{v}^T\boldsymbol{M}^{-1})}{1 + \boldsymbol{v}^T\boldsymbol{M}^{-1}\boldsymbol{v}}$$

and, using it, show that the uncertainty $\sigma_N^2(\boldsymbol{x})$ associated with the Bayesian linear regression function given by Eq. (26) on the slides satisfies

$$\sigma_{N+1}^2(\boldsymbol{x}) \leq \sigma_N^2(\boldsymbol{x}) \tag{1}$$

3 Facebook advertisements

You want to boost your Facebook page and therefore you book Facebook advertisements. A simple linear model for the number of new likes per week (y), depending on the money spent (x) could be:

$$\begin{aligned} y &= a_0 + a_1 x + \epsilon \\ \text{where } y &= \text{number of new likes per week} \\ x &= \text{money spent in that week, in units of 1 EUR} \\ \epsilon &= \text{normal (Gaussian) distributed fluctuations} \end{aligned}$$

After taking a lot of measurement data you fit the parameters. You find:

$$\begin{aligned} a_0 &= 10 \\ a_1 &= 5 \\ \mathbb{E}[y] &= 0 \\ \text{Var}[y] &= 4 \end{aligned}$$

The full model is therefore given by

$$\begin{aligned} y &= 10 + 5x + \mathcal{N}(0, 4) \\ &= 10 + 5x + (8\pi)^{-1/2} \exp(-x^2/8) \end{aligned}$$

Problem 3: Assume you spend no money, what is the probability that you get more than 10 likes per week?

Problem 4: Now you spend 1 EUR on advertisements. What is the expected value of likes?