Machine Learning and Computational Statistics

Theodoros Georgiopoulos

Homework 4

Exercises for Unit 4: Bias - Variance and Estimators

Solution for exercise 1

(a) A reasonable measure to quantify the performance of an estimator is its mean-square deviation from the optimal one. This quantity can be analysed as

$$\begin{aligned} E_D\left[(f(\boldsymbol{x};D) - E[y|\boldsymbol{x}])^2\right] &= E_D\left[(f(\boldsymbol{x};D) - E_D[f(\boldsymbol{x};D)] + E_D[f(\boldsymbol{x};D)] - E[y|\boldsymbol{x}])^2\right] \\ &= \underbrace{E_D\left[(f(\boldsymbol{x};D) - E_D[f(\boldsymbol{x};D)])^2\right]}_{\text{variance}} + \underbrace{\left(E_D[f(\boldsymbol{x};D)] - E[y|\boldsymbol{x}]\right)^2}_{\text{bias}^2}. \end{aligned}$$

The first term is the variance of the estimator arround its own mean value and the second term is the bias squared. To minimize the whole quantity we have to minimize variance and bias simultaneously. It turns out that this is not possible.

(b) For a fixed number of data points, trying to minimize the bias results in an increase of variance and vice versa. This happens because, in order to reduce the bias, you have to increase the complexity of the estimator. This, in turn, results in higher variance as we change the training sets. The only way to reduce both terms simultaneously is to increase the number of the training data points and at the same time increase the complexity of the model carefully.

Solution for exercise 2

We have the regression task $y = g(x) + \eta$, where y and x are modeled by the joint pdf

$$p(x,y) = \frac{3}{2}, x \in (0,1), y \in (x^2, 1)$$

(a) We have the following integral

$$\begin{aligned} \int_x \int_y p(x,y)dxdy &= \int_x \left[\frac{3}{2}y\right]_{x^2}^1 dx \\ &= \int_x \left(\frac{3}{2} - \frac{3}{2}x^2\right) dx \\ &= \left[\frac{3}{2}x - \frac{1}{2}x^3\right]_0^1 dx \\ &= 1. \end{aligned}$$

So, the $p(x,y)$ is a pdf.

(b) The marginal pdf of x is

$$p(x) = \int_y p(x,y)dy = \left[\frac{3}{2}y\right]_{x^2}^{1} = \frac{3}{2}(1-x^2).$$

(c) The conditional pdf of y given x is

$$p(y|x) = \frac{p(x,y)}{p(x)} = \frac{1}{1-x^2}$$

(d) We have

$$E[y|x] = \int_y yp(y|x)dy = \int_{x^2}^{1} y\frac{1}{1-x^2}dy = \frac{1}{1-x^2}\left[\frac{y^2}{2}\right]_{x^2}^{1} = \frac{1+x^2}{2}$$