Problem 1.   2.6

In equation 2.13, we summed up the squares of the differences between the actual value and the estimated value. This error function is the one most frequently used, but it is one of several possible error functions. Because it sums up the squares of the differences, it is not robust to outliers. What would be a better error function to implement *robust regression*?

Solution

Problem 2.   2.7

Derive equation 2.17:

$$w_0 = \bar{r} - w_1\bar{x}$$
$$w_1 = \frac{\sum_t x^t r^t - \bar{x}\bar{r}N}{\sum_t (x^t)^2 - N\bar{x}^2}$$

Solution

We seek to minimize the error function:

$$E(w_1, w_0|X) = \frac{1}{N}\sum_{t=1}^{N}\left[r^t - \left(w_1 x^t + w_0\right)\right]$$

To do this, we set the derivative $\frac{d}{dx}E(X)$ equal to zero:

Problem 3.   2.9

Show that the VC dimension of the triangle hypothesis class is 7 in two dimensions. (Hint: For best separation, it is best to place the seven points equidistant on a circle.)

Solution

Problem 4.   2.10

Assume as in exercise 8 that our hypothesis class is the set of lines. Write down an error function that not only minimizes the number of misclassifications but also maximizes the margin.

Solution

Problem 5.   2.11

One source of noise is error in the labels. Can you propose a method to find datapoints that are highly likely to be mislabeled?

Solution