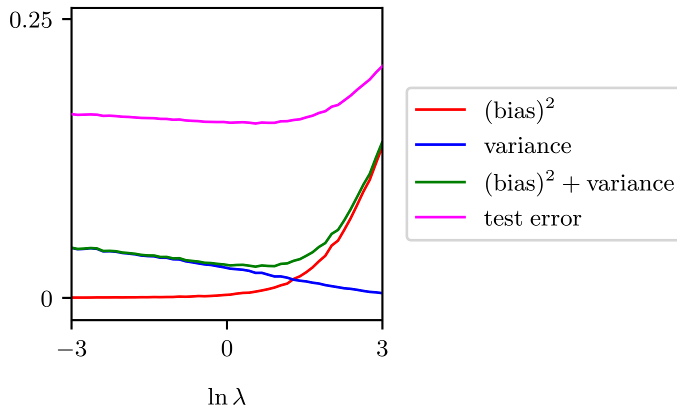

Figure 4.8 Plot of squared bias and variance, together with their sum, corresponding to the results shown in Figure 4.7. Also shown is the average test set error for a test data set size of 1,000 points. The minimum value of $(\text{bias})^2 + \text{variance}$ occurs around $\ln \lambda = 0.43$, which is close to the value that gives the minimum error on the test data.

Exercises

4.1 ($\star$) Consider the sum-of-squares error function given by (1.2) in which the function $y(x, \mathbf{w})$ is given by the polynomial (1.1). Show that the coefficients $\mathbf{w} = \{w_i\}$ that minimize this error function are given by the solution to the following set of linear equations:

$$\sum_{j=0}^{M} A_{ij} w_j = T_i \tag{4.53}$$

where

$$A_{ij} = \sum_{n=1}^{N} (x_n)^{i+j}, \qquad T_i = \sum_{n=1}^{N} (x_n)^i t_n. \tag{4.54}$$

Here a suffix i or j denotes the index of a component, whereas $(x)^i$ denotes x raised to the power of i.

4.2 ($\star$) Write down the set of coupled linear equations, analogous to (4.53), satisfied by the coefficients w_i that minimize the regularized sum-of-squares error function given by (1.4).

4.3 ($\star$) Show that the $\tanh$ function defined by

$$\tanh(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}} \tag{4.55}$$

and the logistic sigmoid function defined by (4.6) are related by

$$\tanh(a) = 2\sigma(2a) - 1. \tag{4.56}$$

Hence, show that a general linear combination of logistic sigmoid functions of the form

$$y(x, \mathbf{w}) = w_0 + \sum_{j=1}^{M} w_j \sigma\left(\frac{x - \mu_j}{s}\right) \tag{4.57}$$