ECE 530 (Fall 2023) **Instructor:** S. Bose

Homework 3

Assigned: Sep 18 Due: Sep 27

Instructions: *You are encouraged to discuss with your peers. You need to, however, turn in your own solutions. Mention the names of your collaborators. Crediting other classmates does not take away any points from you. Any submissions made by one week after the due date will be marked on 50% of the total points. You don't need to submit after that.*

Problem 1. Can you fit a line? [5+3+3+4 = 15 points]

Consider N data-points $(\mathbf{x}_i, y_i)$ for $i = 1, \ldots, N$ obtained from an experiment, where $\mathbf{x}_i \in \mathbb{R}^n$ and $y_i \in \mathbb{R}$. Let $N > n + 1$. The goal of linear regression is to find the 'best' linear fit, i.e., find $\mathbf{c} \in \mathbb{R}^n, d \in \mathbb{R}$ such that

$$y_i \approx \mathbf{c}^\top \mathbf{x}_i + d, \quad \text{for } i = 1, \ldots, N.$$

(a) Suppose each measurement is corrupted by independent Gaussian noise with identical variances and 0 mean. Find $\mathbf{A} \in \mathbb{R}^{N \times (n+1)}$ in terms of $\mathbf{x}_1, \ldots, \mathbf{x}_N$, such that the maximum likelihood estimate of $\widehat{\mathbf{c}}$ and $\widehat{d}$ is

$$\begin{pmatrix} \widehat{\mathbf{c}} \\ \widehat{d} \end{pmatrix} = \left(\mathbf{A}^\top \mathbf{A}\right)^{-1} \mathbf{A}^\top \mathbf{y}, \qquad \text{where} \quad \mathbf{y} := \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}.$$

(b) Consider the case with $n = 1$, i.e., $x_1, \ldots, x_N$ are scalars. Finding the best linear fit amounts to finding $(\hat{c}, \hat{d})$ that minimizes

$$J(c, d) := \sum_{i=1}^{N} (y_i - cx_i - d)^2. \tag{1}$$

Compute a stationary point $(\hat{c}, \hat{d})$ by setting the derivative of J with respect to c and d to zero.

(c) Use the second derivative test to conclude that your $(\hat{c}, \hat{d})$ computed in part (b) is indeed a local minimizer of J. Can you conclude from this second derivative test alone that it is a global minimizer?

(d) Consider the following (x, y) pairs: $(1.00, 1.10)$, $(1.50, 1.62)$, $(2.00, 1.98)$, $(2.50, 2.37)$, $(3.00, 3.23)$, $(3.50, 3.69)$, $(4.00, 3.97)$. Draw a scatter plot of these points. Then, plot the best linear fit to these points using your formulae in part (b).

Problem 2. Estimate with confidence. [15+10 = 25 points]

Suppose the true parameters of a system are described by $\mathbf{x} = (1, 1)^\top$, and the measurements are given by

$$\mathbf{y} = \underbrace{\begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \\ 4 & 5 \end{pmatrix}}_{:=\mathbf{A}} \mathbf{x} + \mathbf{e},$$

where $\mathbf{e} \in \mathbb{R}^4$ is sampled from a multi-variate Gaussian distribution with zero mean and covariance matrix $\mathbf{R} := \mathrm{diag}(0.1, 0.2, 0.3, 0.4)$. That is, $\mathbf{e} \sim \mathcal{N}(0, \mathbf{R})$.

(a) Write an appropriate code for the following experiment.

  (i) Generate an error vector $\mathbf{e}$ according to $\mathcal{N}(0, \mathbf{R})$. Compute $\mathbf{y}$ from it and report it.

  (ii) In class, we derived a formula to compute the maximum likelihood estimate $\hat{\mathbf{x}}$ from $\mathbf{y}$. Using $\mathbf{y}$ in (i), compute and report your estimate $\hat{\mathbf{x}}$.

  (iii) The estimated error is given by $\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}$. With $\mathbf{y}$ from (i) and $\hat{\mathbf{x}}$ from (ii), compute and report

$$J(\hat{\mathbf{x}}) := (\mathbf{y} - \mathbf{A}\hat{\mathbf{x}})^\top \mathbf{R}^{-1} (\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}).$$

  Recall that we claimed (without proof) that $J(\hat{\mathbf{x}}) \sim \chi_2^2$. Then, we have

$$\mathbb{P}\{J(\hat{\mathbf{x}}) \leq 5.9915\} = 0.95,$$

  where $\mathbb{P}\{\mathcal{E}\}$ denotes the probability of an event $\mathcal{E}$. Conclude with 95% confidence level, whether your estimated error conforms to the error model we assumed.

(b) Repeat the experiment in (a) 10K times. Each time you run the experiment, record the value of $J(\hat{\mathbf{x}})$. Plot a histogram of $J(\hat{\mathbf{x}})$. Comment qualitatively, how the histogram of $J(\hat{\mathbf{x}})$ compares to the probability distribution function of a χ_2^2 random variable. Also, report the percentage of times you concluded that your measurement did not conform to the error model. Compare this fraction to your confidence level in (a)(iii).