Math 309 Assignment 2 Solution

Problem 1. Let

$$L(\mu,\sigma) = \ln f(\mu,\sigma) = -\frac{m}{2}\ln 2\pi - m\ln\sigma - \frac{1}{2\sigma^2}\sum_{j=1}^{m}(\mu - x_j)^2$$

The *Maximum Likelihood estimator* $(\bar\mu, \bar\sigma)$ is given by

$$\frac{\partial L}{\partial \mu}(\bar\mu,\bar\sigma) = -\frac{1}{\bar\sigma^2}\sum_{j=1}^{m}(\bar\mu - x_j) = -\frac{1}{\bar\sigma^2}\left(m\bar\mu - \sum_{j=1}^{m}x_j\right) = 0 \tag{1a}$$

$$\frac{\partial L}{\partial \sigma}(\bar\mu,\bar\sigma) = -\frac{m}{\bar\sigma} + \frac{1}{\bar\sigma^3}\sum_{j=1}^{m}(\bar\mu - x_j)^2 = 0 \tag{1b}$$

Therefore, the esitmator is given by

$$\bar\mu = \frac{1}{m}\sum_{j=1}^{m}x_j, \quad \bar\sigma = \left(\frac{1}{m}\sum_{j=1}^{m}(x_j - \bar\mu)^2\right)^{1/2}.$$

The Hessian matrix is

$$\nabla^2 L(\mu,\sigma) = \begin{pmatrix} -\frac{m}{\sigma^2} & \frac{2}{\sigma^3}\sum_{j=1}^{m}(\mu - x_j) \\ \frac{2}{\sigma^3}\sum_{j=1}^{m}(\mu - x_j) & \frac{m}{\sigma^2} - \frac{3}{\sigma^4}\sum_{j=1}^{m}(\mu - x_j)^2 \end{pmatrix}.$$

and

$$\nabla^2 L(\bar\mu,\bar\sigma) = \begin{pmatrix} -\frac{m}{\bar\sigma^2} & 0 \\ 0 & \frac{m}{\bar\sigma^2} - \frac{3}{\bar\sigma^4}\sum_{j=1}^{m}(\bar\mu - x_j)^2 \end{pmatrix} = \begin{pmatrix} -\frac{m}{\bar\sigma^2} & 0 \\ 0 & -\frac{2m}{\bar\sigma^2}. \end{pmatrix}$$

The two eigenvalues for the (diagonal) Hessian matrix $\nabla^2 L(\bar\mu,\bar\sigma)$ are $\lambda_1 = -m/\bar\sigma^2, \lambda_2 = -2m/\bar\sigma^2$, both negative. Therefore, $\nabla^2 L(\bar\mu,\bar\sigma)$ is negative definite and $(\bar\mu,\bar\sigma)$ is a local maximizer.

Remark. In general, we don't know whether the estimator $(\bar\mu,\bar\sigma)$ is a global maximizer or not, since neither f nor L is convex. However, since there is only one solution $(\bar\mu,\bar\sigma)$ to the equation $\nabla L = 0$, it must be the only global minimizer.

Problem 2. First we need to simplify the function by looking at all points x such that

$$x^2 = |2 - x|.$$

When $x > 2$, then $x^2 = x - 2$ which has no solution (in fact we have $x^2 > x - 2$ for all $x \in \mathbb{R}$. For $x < 2$ then $x^2 = 2 - x$ has two solutions $x_1 = -2$ and $x_2 = 1$. Moreover, for $x \in (-2, 1)$ we have $x^2 < 2 - x$. Therefore, the original function can be written as

$$f(x) = \begin{cases} 2 - x, & x \in (-2, 1), \\ x^2, & x \le -2 \text{ or } x > 1. \end{cases}$$

On the interval $(-\infty, -2]$, $f'(x) = 2x < 0$ and thus f is decreasing. The minimizer on this interval is obtained at $x_1^* = -2$ with value $f(x_1^*) = 4$. On the interval $[-2, 1]$, $f'(x) = -1$ and thus f is decreasing. The minimizer is obtained at $x_2^* = 1$ with function value $f(x_2^*) = 1$. On the interval $[1, \infty)$, $f'(x) = 2x > 0$ and thus f is increasing. The minimizer is obtained at $x_2^* = 1$ with value $f(x_2^*) = 1$. Put all together, the global minimizer is at $x^* = 1$ with the minimal value $f(x^*) = 1$.

Problem 3. The characteristic polynomial is

$$p(\lambda) = \det(\lambda I - M) = \lambda^2 - (a + c)\lambda + ac - b^2 = \lambda^2 - \mathrm{tr}(M)\lambda + \det(M).$$

The two roots are

$$\lambda_\pm = \frac{a + c \pm \sqrt{(a + c)^2 - 4(ac - b^2)}}{2} = \frac{a + c \pm \sqrt{(a - c)^2 + 4b^2}}{2},$$

bot of which are real. Since $\lambda_+ > 0$, $\lambda_- = \det M/\lambda_+ > 0$. Therefore, M has two positive eigenvalues, and therefore is positive definite.

Problem 4. To show that one matrix N is the inverse of another matrix M, we only need to show that $NM = I$ (or $MN = I$), the identity matrix.

$$\begin{aligned} &(A + UCV)(A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}) \\ &= AA^{-1} + AA^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1} - UCVA^{-1} - UCVA^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1} \\ &= I + U(C^{-1} + VA^{-1}U)^{-1}VA^{-1} - UCVA^{-1} - UC\mathbf{VA^{-1}U(C^{-1} + VA^{-1}U)^{-1}}VA^{-1} \end{aligned} \tag{2}$$

Using the fact that

$$\begin{aligned} I &= (C^{-1} + VA^{-1}U)(C^{-1} + VA^{-1}U)^{-1} \\ &= C^{-1}(C^{-1} + VA^{-1}U)^{-1} + VA^{-1}U(C^{-1} + VA^{-1}U)^{-1}, \end{aligned}$$

we have $VA^{-1}U(C^{-1} + VA^{-1}U)^{-1} = I - C^{-1}(C^{-1} + VA^{-1}U)^{-1}$. Substituting it into (2),

$$\begin{aligned} &(A + UCV)(A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}) \\ &= I + U(C^{-1} + VA^{-1}U)^{-1}VA^{-1} - UCVA^{-1} - UC(\mathbf{I - C^{-1}(C^{-1} + VA^{-1}U)^{-1}}VA^{-1} \\ &= I. \end{aligned} \tag{3}$$

Therefore, $A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$ is the inverse of $A + UCV$.

Problem 5. See the completed code on webct (under "matlab demo").