ASE 381P Problem Set #3

Posting Date: September 26, 2019

You need not hand in anything. Instead, be prepared to answer any of these problems on an upcoming take-home exam. You may discuss your solutions with classmates up until the time that the exam becomes available, but do not swap work.

1. Let $Z^k = \{z_1, ..., z_k\}$ be a set of k independent random samples from the normal distribution $N(\mu, \sigma^2)$. Find the maximum likelihood estimates $\hat{\mu}_{ML}$ and $\hat{\sigma}^2_{ML}$ of the parameters μ and σ^2. Make $\hat{\sigma}^2_{ML}$ depend on $\hat{\mu}_{ML}$ instead of on μ, which is assumed to be unknown. Viewing $\hat{\mu}_{ML}$ and $\hat{\sigma}^2_{ML}$ as random variables, find $E[\hat{\mu}_{ML}]$, $E[\hat{\sigma}^2_{ML}]$, and $\mathrm{var}(\hat{\mu}_{ML})$. Are the ML estimates for μ and σ^2 both unbiased? What effect does the number of samples k have on the estimates?

2. Write a Matlab function to solve the weighted least-squares problem:

   Minimize $J(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{z} - H\boldsymbol{x})^T R^{-1}(\boldsymbol{z} - H\boldsymbol{x})$

   where R is an $m \times m$ positive definite symmetric matrix representing the measurement error covariance. Use the square-root information approach described in lecture wherein the Cholesky and QR factorizations are invoked. Write your Matlab function to adhere to the interface given below. Test your function with the values for $\boldsymbol{z}$, H, and R given below. Compare the answer for $\hat{\boldsymbol{x}}$ returned by your function to the answer provided by the standard weighted least squares technique (i.e., by the normal equations). Provide numerical answers for both methods; also print out `Rotilde` for the square-root method. Explain the differences between the answers and indicate which one should be trusted most. Note: you should be able to copy and past the text and numbers below into Matlab.

```
function [xhat,Rotilde] = sribls(Hprime,zprime,R)
% sribls : Square-root-information-based least squares routine.  Given the
%          linear measurement model:
%
%            zprime = Hprime*x + w,  w ~ N(0,R)
%
%          sribls returns xhat, the least squares (and Maximum Likelihood)
%          estimate of x, and Rotilde, the associated square root
%          information matrix.
%
% INPUTS
%
% Hprime ----- nz-by-nx measurement sensitivity matrix.
%
% zprime ----- nz-by-1 measurement vector.
%
% R ---------- nz-by-nz measurement noise covariance matrix.
```

```
%
%
% OUTPUTS
%
% xhat ------- nx-by-1 Maximum Likelihood estimate of x.
%
% Rotilde ---- nx-by-nx square root information matrix, where P =
%              inv(Rotilde'*Rotilde) is the estimation error covariance matrix.
%
%+------------------------------------------------------------------------------+
% References:
%
%
% Author:
%+==============================================================================+
```

```
z =
   5.505751578229795
   9.754581690707774
  -3.380491358445003
  -1.240740004272318
   1.402850573448269

H =
  -0.809498694424876  -0.809498871800032  -1.618995978525818
  -2.944284161994896  -2.944284358048384  -5.888569324509237
   1.438380292815098   1.438381712125249   2.876762701564763
   0.325190539456198   0.325190831040572   0.650382205584935
  -0.754928319169703  -0.754928121358650  -1.509856684243494

R =
   9.599563453667082   1.695803491633456   3.374804722838387  -2.036285115123790  -1.912094793736365
   1.695803491633456  22.490148367908073   2.451654084712032   9.799176417244743   7.078780745346799
   3.374804722838387   2.451654084712032  12.073999822869279   1.904627078020524  -3.804148586359607
  -2.036285115123790   9.799176417244743   1.904627078020524   6.270434397020487   4.070651550638881
  -1.912094793736365   7.078780745346799  -3.804148586359607   4.070651550638881   5.320048568546946
```

3. Suppose that one defines the following least-squares estimation cost function for k samples of data:

$$J(\boldsymbol{x}, k) = \sum_{j=1}^{k} \left[\boldsymbol{z}(j) - H(j)\boldsymbol{x}\right]^T R^{-1}(j) \left[\boldsymbol{z}(j) - H(j)\boldsymbol{x}\right]$$

where $R(j) = E[\boldsymbol{w}(j)\boldsymbol{w}^T(j)]$ is the usual measurement error covariance matrix. Suppose also that $\hat{\boldsymbol{x}}(k|Z^k)$ and $\hat{P}(k|Z^k)$ are, respectively, the least-squares estimate that minimizes $J(\boldsymbol{x}, k)$ and the estimation error covariance for this estimate. Then prove that

$$J(\boldsymbol{x}, k) = [\boldsymbol{x} - \hat{\boldsymbol{x}}(k|Z^k)]^T \hat{P}^{-1}(k|Z^k)[\boldsymbol{x} - \hat{\boldsymbol{x}}(k|Z^k)] + J[\hat{\boldsymbol{x}}(k|Z^k), k]$$

Hints: Both forms of $J(\boldsymbol{x}, k)$ are quadratic in the argument $\boldsymbol{x}$. Equality of the two forms of $J(\boldsymbol{x}, k)$ can be proved by showing that the respective coefficients of the constant, linear, and quadratic $\boldsymbol{x}$ terms are equal. You may need to try several alternate equivalent formulas for $\hat{P}(k|Z^k)$ in order to find the one that is most useful for purposes of this proof.

4. An object is moving along one dimension with a constant velocity and its position is measured at regular intervals. The first measurement is taken at the initial position and the last measurement is taken at the final position. Assume that the number of measurements n is odd and that the measurement errors are mutually uncorrelated, zero-mean, and Gaussian with variance σ^2.

   (a) Consider the case for which the constant velocity is unknown. Calculate the variance of the last position estimate (based on all n measurements).

   (b) Now consider the case for which the constant velocity is known exactly. Only the object's position is unknown and must be estimated. Calculate the variance of the last position estimate (based on all n measurements).

   (c) Show that the ratio of the position estimate variance for the unknown velocity case to the position estimate variance for the known velocity case is $(4n-2)/(n+1)$.

   (d) For the unknown velocity case, calculate the position estimate variance at the center position [i.e., the one corresponding to the measurement with index $i_c = (n+1)/2$ for indexing that starts with $i = 1$] based on all n measurements. Is there any other position estimate that has a lower variance in this case? Give an intuitive explanation.

5. Consider the problem of estimating an unknown frequency parameter ω from complex measurements

$$z_k = \exp(j\omega kT) + n_k, \quad k = 0, 1, ..., N-1$$

where $j = \sqrt{-1}$ and where $n_k = a_k + jb_k$ is a sequence of independent, identically-distributed complex zero-mean Gaussian noise samples with variance σ^2:

$$a_k, b_k \sim \mathcal{N}(0, \sigma^2), \quad E[a_k a_j] = \sigma^2 \delta_{kj}, \quad E[b_k b_j] = \sigma^2 \delta_{kj}, \quad E[a_k b_j] = 0 \ \ \forall k, j$$

Here, δ_{kj} is the Kronecker delta (equal to unity for $k = j$ and otherwise zero).

Suppose the N measurements are stacked as a complex vector $\boldsymbol{z} = \boldsymbol{x} + j\boldsymbol{y}$, where $\boldsymbol{z} = [z_0, z_1, ..., z_{N-1}]^T$, and where $\boldsymbol{x} = [x_0, x_1, ..., x_{N-1}]^T$ and $\boldsymbol{y} = [y_0, y_1, ..., y_{N-1}]^T$ are the real and imaginary components of $\boldsymbol{z}$.

We wish to find the CRLB for an ML estimate of ω. Since $\boldsymbol{z}$ is Gaussian, the likelihood function is simply

$$\Lambda(\omega) \triangleq p(\boldsymbol{z}|\omega) = \left(\frac{1}{2\pi\sigma^2}\right)^N \exp\left(-\frac{1}{2\sigma^2}[\boldsymbol{z} - \boldsymbol{\mu}_z]^H[\boldsymbol{z} - \boldsymbol{\mu}_z]\right)$$

where $\boldsymbol{\mu}_z = E[\boldsymbol{z}]$ and where $\boldsymbol{x}^H$ represents the conjugate transpose of the vector $\boldsymbol{x}$. Expressing the inner product as a summation, and recognizing that $\exp(j\omega kT) = \cos(\omega kT) + j\sin(\omega kT)$, $\Lambda(\omega)$ may be written

$$\Lambda(\omega) = \left(\frac{1}{2\pi\sigma^2}\right)^N \exp\left(-\frac{1}{2\sigma^2}\sum_{k=0}^{N-1}[x_k - \cos(\omega kT)]^2 + [y_k - \sin(\omega kT)]^2\right)$$

Find the CRLB for an ML estimate of ω.

6. Consider a standard least-squares parameter estimation problem with an unknown constant parameter vector $\boldsymbol{x} \in \mathbb{R}^n$, measurements $\boldsymbol{z} \in \mathbb{R}^m$, and zero-mean Gaussian measurement noise $\boldsymbol{w} \in \mathbb{R}^m$:

$$\boldsymbol{z} = H\boldsymbol{x} + \boldsymbol{w}$$

Assume the problem has been normalized such that $E[\boldsymbol{w}\boldsymbol{w}^T] = I_{m\times m}$. Also assume that H is full column rank.

(a) Give an expression for $\hat{\boldsymbol{x}}$, the least squares solution for $\boldsymbol{x}$.

(b) Let $\tilde{\boldsymbol{x}} \triangleq \boldsymbol{x} - \hat{\boldsymbol{x}}$. Give an expression for the error covariance matrix $P = E[\tilde{\boldsymbol{x}}\tilde{\boldsymbol{x}}^T]$ in terms of H.

(c) Find the log likelihood function $\log \Lambda(\boldsymbol{x}) \triangleq \log p(\boldsymbol{z}|\boldsymbol{x})$, where log refers to the natural logarithm. Express $\log \Lambda(\boldsymbol{x})$ in the form

$$\log \Lambda(\boldsymbol{x}) = C - \tfrac{1}{2}\sum_{l=1}^{m} \gamma_l^2$$

where C is a constant and γ_l is a scalar that you express in terms of $\boldsymbol{z}$, H, and $\boldsymbol{x}$.

(d) The covariance P cannot be smaller (in the positive definite sense) than the so-called Cramer-Rao lower bound (CRLB), which is the inverse of the Fisher information matrix J. Let J_{ij} be the ijth element of J, and let x_i be the ith element of $\boldsymbol{x}$. Then $J_{ij} = E[G_i G_j]$ where

$$G_i \triangleq \frac{\partial \log \Lambda(\boldsymbol{x})}{\partial x_i}$$

Find an expression for J_{ij} in terms of the elements of H.

(e) Given your answers to 2 and 4 above, is the least squares estimator *efficient*? Explain.

Also do the following problems from Bar Shalom: 2-8 (should be easy given chapter 3 and the equivalence between MAP and MMSE estimation for linear/Gaussian problems), 2-9, 2-11, 2-14 (Clarification: Let the final question be restated as follows: "For what value of N are you willing to risk your job by saying there is a 5% or less probability of getting $\hat{\sigma}^2 \leq 80$ while the true value of σ^2 is 100 or higher?"), 3-7 (in parts (2) and (3) replace MMSE with MSE), 3-9 (orthogonal means uncorrelated here), 3-12, 3-13, 3-16.