ASEN 5044 Statistical Estimation for Dynamical Systems
Fall 2018
Homework 6

Out: Friday 10/26/2018 (posted on Canvas)

Due: Friday 11/2/2018, 1 pm (Canvas - **no credit for illegible work**)

Show all your work and explain your reasoning.

1. A certain linear system is known to satisfy the following differential equation,

$$\ddot{z} + 10\dot{z} + 100z = f(t), \quad z(0) = \dot{z}(0) = 0,$$

where $z(t)$ is the response variable (e.g. position) and $f(t)$ is a white noise forcing function with a PSD of 10 'units' (assume the units are consistent throughout).

(a) Let z and $\dot{z}$ be the state variables for this system. Write out the dynamics for this CT system in stochastic LTI form.

(b) Suppose this process is sampled at a uniform rate beginning at time $t = 0$, with $\Delta T = 0.2$ sec. Convert the CT stochastic LTI model from part (a) into DT form, and report the corresponding F and Q matrices (you may use software for this part).

(c) Suppose the units of z are in mand $\dot{z}$ are in m/s. Provide the corresponding units of the elements of Q, and explain what Q says physically about the elements of the DT process noise vector $w(k)$ and its effect on the DT state $x(k)$.

(d) Suppose a sensor is attached to this system, with the CT model $y(t) = z + \dot{z}\Delta T + \tilde{v}(t)$, where $\tilde{v}(t)$ is a white noise input with PSD of 3 'units'. Report the H and R matrices for the corresponding DT model; also report the units of R if z has units of meters.

2. Simon Problem 3.4.

3. Simon Problem 3.13 (use software, but show and explain all your work).

4. (Use software for this problem, but show and explain all your work) Consider the 3D static robot GPS localization problem defined in Lecture 19. Suppose the AWGN sensor error covariance matrix is given by (in units of m^2)

$$R = \begin{bmatrix} 8 & 5.15 & 6.5 \\ 5.15 & 5 & -4.07 \\ 6.5 & -4.07 & 50 \end{bmatrix}$$

(a) Suppose the true robot position is known to be at $\xi_0 = \eta_0 = z_0 = 1$ m. Simulate $T = 100$ measurements for the robot at this location, and provide three separate 2D scatter plots of the resulting simulated noisy y data: one showing all $y_k(1)$ vs. $y_k(2)$ data, another showing all $y_k(1)$ vs. $y_k(3)$ data, and another showing all $y_k(2)$ vs. $y_k(3)$ data (be sure to use the same scales on both axes on each plot).

(b) Compute the sample covariance matrix for the simulated measurement vectors from part (a) – how does this compare to R?

(c) Use the data simulated from (a) to estimate the robot's 3D position via batch weighted LS using the first 3 samples you simulated, then the first 10 samples, and then using all 100 samples you simulated. Report the **estimation error covariance** matrix in all 3 cases (be sure to provide units); how accurate are these estimates compared to the ground truth? (Hint: recall from Simon Chapter 3 that the estimation error covariance is NOT the same as the matrix R from part (b)...)

(d) Using the data log of $y(k)$ vectors over $T' = 30$ time steps in the `hw6problem5data.csv` file provided on D2L, estimate the robot's position using batch weighted LS and report the estimated error covariance matrix (be sure to provide units; note the data log has a full 3×1 data vector $y(k)$ in each column).

(e) Repeat (d) using batch unweighted LS, and compare the final result and its covariance to the results from (d).

(f) Repeat (d) using recursive weighted LS, and report the results by showing a separate plot for each of the estimated states (using solid lines) and their associated $\pm 2\sigma$ error bounds (using dashed lines) evolving vs. time step k, i.e. as each new measurement vector $y(k)$ comes in for $k = 1$ to T' (be sure to provide units).

Advanced Questions ***All students are welcome to try any of these for extra credit (only given if all regular problems turned in on time as well). Submit your responses for these questions with rest of your homework, but make sure these are clearly labeled and start on separate pages – indicate in the .pdf file name (per instructions posted on Canvas) and on the front page of your assignment if you answered these questions, so they can be spotted, graded and recorded more easily.***

AQ1. The *Cramér-Rao Lower Bound* (CRLB) implies that there is a fundamental limit to how certain we can be of an *unbiased* estimate of a parameter. Let $\mathbf{X} = [X_1 \dots X_N]^T$ be some random vector with PDF

$$p(\mathbf{x}; \theta)$$

where θ is some parameter of interest that influences the PDF of $\mathbf{X}$. For example, for some normally distributed scalar random variable Y with mean μ we may write

$$Y \sim p(y; \mu) = \mathcal{N}(\mu, \sigma^2)$$

Let $\hat{\mathbf{X}}$ be a realization (i.e. a measurement) of $\mathbf{X}$ and let $\hat{\theta}$ be an *unbiased* estimate of θ based on $\hat{\mathbf{X}}$. Then the CRLB states that

$$\mathrm{cov}(\hat{\mathbf{X}}; \hat{\theta}) \geq \mathbf{I}^{-1}(\theta)$$

where $\mathbf{I}(\theta)$ is defined as the Fisher Information Matrix, given by

$$\mathbf{I}(\theta) = \mathbb{E}\left[\left(\frac{\partial \ln p(\mathbf{X}; \theta)}{\partial \theta} \right)^T \left(\frac{\partial \ln p(\mathbf{X}; \theta)}{\partial \theta} \right) \right]$$

Note that the expectation is taken with respect to $\mathbf{X}$. The above implies that the uncertainty in the estimate is bounded from below by the inverse of the Fisher Information Matrix, where "bounded" in this context means $\text{cov}(\hat{\mathbf{X}}; \hat{\theta}) - \mathbf{I}^{-1}(\theta)$ is positive semi-definite. An estimator that performs at the CRLB, i.e. achieves $\text{cov}(\hat{\mathbf{X}}; \hat{\theta}) = \mathbf{I}^{-1}(\theta)$ is said to be *fully efficient.*

This problem will explore one primary application of the CRLB, namely the evaluation of the efficiency of an estimator.

(a) Given $N > 2$ observations of (x_i, y_i) of a simple linear model of the form

$$y_i = ax_i + b + w_i$$

where $\mathbf{w} = \begin{bmatrix} w_1 & \dots & w_N \end{bmatrix}^T \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$, state the equations for a batch linear least squares estimator that solves for an optimal estimate $\hat{\mathbf{X}}$ of the parameters $\mathbf{X} = \begin{bmatrix} a & b \end{bmatrix}^T$.

(b) Derive the CRLB of $\hat{\mathbf{X}}$ using the following procedure.

i. Find an expression for the PDF of $\hat{\mathbf{Y}} = \begin{bmatrix} y_1 & \dots & y_N \end{bmatrix}^T$ as a function of $\mathbf{X}$. Call this $p(\hat{\mathbf{Y}}; \mathbf{X})$.

ii. Compute the *score* of the PDF

$$\frac{\partial \ln p(\hat{\mathbf{Y}}; \mathbf{X})}{\partial \mathbf{X}}$$

iii. Compute the Fisher Information Matrix

$$\mathbf{I}(\theta) = \mathbb{E}\left[\left(\frac{\partial \ln p(\hat{\mathbf{Y}}; \mathbf{X})}{\partial \mathbf{X}}\right)^T \left(\frac{\partial \ln p(\hat{\mathbf{Y}}; \mathbf{X})}{\partial \mathbf{X}}\right)\right]$$

(c) Show that the batch LLS estimator derived in Part (a) achieves the CRLB found in Part (b).