ASEN 5044 Statistical Estimation for Dynamical Systems
Fall 2018
Midterm Exam 2

Out: Thursday 11/08/2018 (posted on Canvas)

Due: Thursday 11/15/2018, 11 am (Canvas)

This exam is open notes and open book. You may ask Prof. Ahmed and TA for clarification, ***but you may not consult with each other (honor system applies and will be enforced)****. Show all your work and explain your reasoning for full credit.* ***Use computer software only as/when explicitly indicated in the problem statement.***

1. [20 pts] Simon, Problem 4.2. (**Hint:** read the problem very carefully: it says the variance of x_k is σ^2 for all time steps k; it also does *not* say x_k is zero mean for all time k, so do *not* make this assumption)

2. [20 pts] Given the random vector $x \in \mathbb{R}^n$ (with some finite mean and covariance matrix) and the constant non-random matrix $A \in \mathbb{R}^{n \times n}$ where $A = A^T$, find $\mathbb{E}[x^T A x]$. Be sure to show and properly explain all steps used to get to your result (**Hint**: Lec 2 slide 4 – think about the size of the result of the quadratic form to see the relevant connection. To use this fact, you should first prove that, if Z is a square matrix whose elements are random variables, then $\mathbb{E}[tr(AZ)] = tr(A\mathbb{E}[Z])$).

3. [35 pts] Consider the coordinated turning aircraft problem from HW 7. Assume again that the dynamics are free of process noise,

$$x(k+1) = Fx(k), \tag{1}$$

for the F matrix and states as defined in HW 7 (for some known turning rate(s) Ω and discretization step ΔT), but now assume that noisy measurements of the form

$$y(k+1) = Hx(k+1) + v(k), \tag{2}$$

$$E[v(k)] = 0, \quad E[v(k)v^T(j)] = \delta(k,j)R(k) \tag{3}$$

are available with additive white noise $v(k)$ and non-stationary noise covariance $R(k)$. For each part of this problem, it is desired to estimate the initial state $x(0) \in \mathbb{R}^n$ of a dynamical system consisting of either one or two turning aircraft at time $k = 0$ from noisy measurements $y(k) \in \mathbb{R}^p$ taken at time steps $k = 1, 2, ..., T$.

a. Derive the analytical expression for the batch estimator $\hat{x}(0)$ which minimizes the cost function,

$$J(T) = \sum_{k=1}^{T} (y_k - \hat{y}_k)^T [R(k)]^{-1} (y_k - \hat{y}_k) \tag{4}$$

where $\hat{y}_k$ is the time k estimator-based predicted measurement (a function of $\hat{x}(0)$) and y_k is the actual measurement at time k. Be sure to precisely and carefully define all matrices and vectors used by the estimator, for the case where x_k and y_k obey eqs. (1) and (2).

b. Suppose a ground tracking station monitors aircraft A which is turning with $\Omega_A = 0.045$ rad/s , and converts 3D range and bearing data into 2D 'pseudo-measurements' $y_A(k)$ with the following DT measurement model

$$y_A(k) = Hx_A(k) + v_A(k),$$
$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \; R_A = \begin{bmatrix} 75 & 7.5 \\ 7.5 & 75 \end{bmatrix} + \begin{bmatrix} 12.5\sin(k/10) & 25.5\sin(k/10) \\ 25.5\sin(k/10) & 12.5\cos(k/10) \end{bmatrix}$$

where R_A has units of m^2. Using the data posted in '`midterm2_problem3b.mat`', use your result from part (a) to compute the estimate $x_A(0)$, and report the final state estimation error covariance matrix. Note that the data provided is for time steps $k \geq 1$, and each column k corresponds to a single y_k vector at step k.

c. Suppose now there is also a second aircraft B which is turning with $\Omega_B = -0.045$ rad/s. The tracking station can only directly sense one aircraft at a time, and thus cannot sense B while it senses A. However, a transponder between A and B provides a noisy measurement $y_D(k)$ of the difference in their 2D positions as they are turning, $r_A = [\xi_A, \eta_A]^T$ and $r_B = [\xi_B, \eta_B]^T$,

$$y_D(k) = r_A(k) - r_B(k) + v_D(k),$$
$$R_D = \begin{bmatrix} 8000 & 500 \\ 500 & 8000 \end{bmatrix},$$

where R_D has units of m^2 and $v_D(k) \sim \mathcal{N}(0, R_D)$ follows a stationary white noise process.

Using the data posted in '`midterm2_problem3c.mat`' in the array 'yaugHist' (where each column contains a concatenated 4×1 data vector $[y_A^T(k), y_D^T(k)]^T$ that includes a new set of $y_A(k)$ measurements for time steps $k \geq 1$), compute an RLLS estimate for $x(0) = [x_A(0), x_B(0)]^T$. In separate plots, show the evolution of each aircraft's state estimate vs. k (use 4 subplots per aircraft); also on separate plots, show positive 2σ bounds for each estimated state vs. k (use 4 subplots per aircraft). Be sure to explain how you set up the RLLS estimator to estimate $x(0)$ and how you initialized RLLS.

4. [25 pts] Consider the pure prediction update for the state covariance of a DT linear-Gaussian system with state transition matrix F and process noise covariance matrix Q,

$$P_{k+1} = FP_kF^T + Q. \tag{5}$$

If F is stable, then as $k \to \infty$, this equation leads to a DT *Lyapunov equation* for the steady-state covariance P_∞ ,

$$P_\infty = FP_\infty F^T + Q. \tag{6}$$

a. Suppose $F = f$ is some scalar constant, where $|f| < 1$, and $Q = q$ is an arbitrary process noise variance, with $q > 0$. Derive an analytical solution for the scalar state variance $P_\infty = \sigma_\infty^2$ at steady state.

b. Run and report (using suitably labeled/scaled plots) a simple simulation of the covariance prediction update for 50 time steps at values of $f = 0.8$, $q = 10$, and $\sigma_0^2 = 50$, and compare this to the analytical result from part (a). Repeat and report the comparison for $\sigma_0^2 = 10$. What do these sets of results tell you about the sensitivity of the steady state variance on the initial variance at time 0?

c. Now suppose F is an arbitrary non-diagonal but stable 2×2 state transition matrix, and Q is an arbitrary non-diagonal 2×2 covariance matrix. Manually derive an analytical expression for a system of linear equations that (when solved) provides the solution for the elements of P_∞ in terms of the elements of F and Q. You do not need to explicitly solve for each unknown P_∞ element, but you must specify all coefficients for each equation corresponding to the unknown P_∞ elements, and express these together in an appropriate matrix-vector form (DO NOT use Matlab's Symbolic toolbox or other computer aids). (**Hint**: the fact that P_∞ and Q are both covariance matrices makes it easy to describe their elements).

d. Use the result from part (c) to solve for the elements of P_∞ for the following F and Q values (you may solve the system of equations computationally):

$$F = \begin{bmatrix} 0.99 & 0.2 \\ 0 & -0.76 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 0.37 \\ 0.37 & 2.5 \end{bmatrix}.$$

e. Use suitably labeled/scaled plots to validate your results from part (d) by comparing to both: (a) a sufficiently long predicted covariance simulation for $P_0 = 10 \cdot I_{2\times 2}$, and (b) the output of the `dlyap.m` command in Matlab.

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. Hidden Markov models (HMMs) are useful for devising approximations to the general Bayes filter when a stochastic dynamical system's states can be restricted to a finite domain. Consider the problem of localizing an object that moves and is observed stochastically in a bounded domain, e.g. an old jalopy moving on a long, windy, and unpaved road that is monitored by noisy unattended ground sensors. Let the jalopy's state is given by its scalar position, $x(k) \in [0, L]$, where L is the length of the road. Suppose the road is discretized into N grid cells with spacing Δx, and that the jalopy moves with some random constant velocity over time interval Δt between each time step $k = 0, 1, 2, ...$ In this case, the random state variable $x(k)$ can be replaced by $\zeta(k) \in \{1, \cdots, N\}$, which indexes a grid cell for $x(k)$ at time k and whose dynamics are well-approximated by an HMM [1].

For grid cells $i, j \in \{1, \cdots, N\}$, let the HMM conditional state index transition probabilities $P(\zeta_{k+1} = i|\zeta_k = j)$ (which approximately encode the spatially and temporally

[1] e.g., see: N. Ahmed, D. Casbeer, Y. Cao, and D. Kingston, "Multi-target Localization on Road Networks with Hidden Markov Rao-Blackwellized Particle Filters," AIAA Journal of Aerospace Information Systems, Vol. 14, No. 11 (2017), pp. 573-596.

discretized system dynamics) be defined by the following time-invariant parameters,

$$P(\zeta_{k+1} = i|\zeta_k = j) = \begin{cases} 0.3, \text{ if } \; i = j \text{ and } \; i \in \{2, \cdots, N-1\} \\ 0.5, \text{ if } \; i = j+1 \text{ and } \; i \in \{3, \cdots, N-1\} \text{ and } \; j \geq 2 \\ 0.2, \text{ if } \; i = j-1 \text{ and } \; i \in \{2, \cdots, N-2\} \text{ and } \; j \leq N-1 \\ 0.1, \text{ if } \; i = j = 1 \\ 0.1, \text{ if } \; i = j = N \\ 0.9, \text{ if } \; i = 2 \text{ and } \; j = 1 \\ 0.9, \text{ if } \; i = N-1 \text{ and } \; j = N \\ 0, \text{ otherwise.} \end{cases}$$

The Chapman-Kolmogorov equation for recursive probabilistic state prediction can thus be approximated in terms of the marginal probability mass distribution over ζ_{k+1},

$$P(\zeta_{k+1} = i) = \sum_{j=1}^{N} P(\zeta_{k+1} = i|\zeta_k = j)P(\zeta_k = j).$$

Finally, suppose that an unattended ground sensor with resolution $2r$ provides noisy grid cell index observations $z_k = i \in \{1, \cdots, N\}$ for the jalopy conditioned on $\zeta_k = j$, as given by the probabilistic observation likelihood

$$P(z_k = i|\zeta_{k+1} = j) = \mathcal{U}_{i|j}[a(r), b(r)],$$
$$a(r) = \max(1, j-r), \quad b(r) = \min(N, j+r),$$

For this problem, assume $r = 25$, $L = 1000$ m, $\Delta x = 10$ m, $\Delta T = 5$ s, and $N = 100$.

a. Suppose the vector $\rho^-_{k+1} \equiv [P(\zeta_{k+1} = 1), \cdots, P(\zeta_{k+1} = N)]^T$. Find the square matrix T such that $\rho^-_{k+1} = T\rho^-_k$, and determine whether a *stationary distribution* ρ^-_∞ exists for the system, i.e. such that $\rho^-_\infty = T\rho^-_\infty$. (**Hint:** If ρ^-_∞ exists, what would be special about the relationship between ρ^-_∞ and T?)

b. Suppose ρ^-_0 is such that $P(\zeta_0 = 3) = P(\zeta_0 = 19) = P(\zeta_0 = 35) = \frac{1}{3}$. Derive an expression in terms of ρ^-_0 which gives the predicted grid index distribution ρ^-_k at some future time step $k > 0$, and plot ρ^-_k vs. ζ for $k = 10, 50, 180, 300$ and 600 time steps (e.g. using the 'barplot' function in Matlab). If ρ^-_∞ exists, approximately how long does it take for ρ^-_k to get 'close to' ρ^-_∞?

c. Use the prior in part (b) and the measurement data log 'zMeasHist' posted in the file `'midterm2_problemAQ1.mat'` to implement the Bayes filter, i.e. compute a recursive posterior distribution $\rho^+_k \equiv [P(\zeta_k = 1|z_{1:k}), \cdots, P(\zeta_k = N|z_{1:k})]^T$ to solve the filtering problem at each time step in the data log (the first value of 'zMeasHist' corresponds to $z(0)$, which is deliberately set to NaN). Turn in plots for ρ^+_k vs. $\zeta(k)$ at time steps $k = 7, 50, 180, 300,$ and 350; report the corresponding MMSE and MAP estimates for $\hat{\zeta}^+_k$ at each of these specified time steps, and compare/discuss the results with respect to the ground truth ζ_k values provided in the array 'zetaHist' (where the initial value corresponds to the true $\zeta(0)$ initial condition).