

## Problem 1

Suppose that a dynamic scalar system is given as  $x_{k+1} = f_{x_k} + w_k$  where  $w_k$  is zero-mean white noise with variance  $q$ . Show that if the variance of  $x_k$  is  $\sigma^2$  for all  $k$ , then it must be true that  $f^2 = (\sigma^2 - q)/\sigma^2$ . **Hint:** The problem says the variance of  $x_k$  is  $\sigma^2$  for all time steps  $k$ . It does *not* say  $x_k$  is zero mean for all  $k$ .

If we start with the assumption that  $x_k \sim \mathcal{N}(\mu_k, \sigma^2)$  and  $w_k \sim \mathcal{N}(0, q)$ , then we can find the variance of  $x_{k+1}$  as the variance of the linear combination  $f x_k + w_k$ :

$$\sigma_{x_{k+1}}^2 = f^2 \sigma_{x_k}^2 + q$$

But since  $x_k$  has the same variance at all timesteps  $k$  the expression becomes  $\sigma^2 = f^2 \sigma^2 + q$ . If we solve this equation for  $f^2$  we get

$$f^2 = \frac{\sigma^2 - q}{\sigma^2}$$

## Problem 2

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 your result. **Hint:** Ref lecture 2, slide 2. 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  $\mathbb{E}[\text{tr}(AZ)] = \text{tr}(A\mathbb{E}[Z])$ .

$\mathbb{E}[x^T A x]$  can be broken down as follows:

$$\begin{aligned} \mathbb{E}[x^T A x] &= \mathbb{E}\left[\sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j\right] \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} \mathbb{E}[x_i x_j] \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} \left[ \text{cov}(x_i x_j) + \mathbb{E}[x_i] \mathbb{E}[x_j] \right] \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} (\sigma_{ji} + \mu_i \mu_j) \quad \text{noting: } \text{cov}(x_i x_j) = \text{cov}(x_j x_i) \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{ij} \sigma_{ji} + \sum_{i=1}^n \sum_{j=1}^n a_{ij} \mu_i \mu_j \\ &= \text{tr}(AP) + \mu^T A \mu \end{aligned}$$

where  $\sigma_{ij}$  is the covariance between  $x_i$  and  $x_j$ ,  $P$  is the full covariance matrix of  $x$ , and  $\mu$  is the expected value of  $x$ .

### Problem 3

Consider the coordinated turning aircraft problem from homework 7. Assume again the dynamics are free of process noise,

$$x(k+1) = Fx(k)$$

for the  $F$  matrix and states as defined in homework 7 for some unknown turning rate  $\Omega$  and discretization step  $\Delta T$ . Now add measurements of the form

$$\begin{aligned} y(k+1) &= Hx(k+1) + v(k) \\ \mathbb{E}[v(k)] &= 0, \quad \mathbb{E}[v(k)v(j)^T] = \delta(k, j)R(k) \end{aligned}$$

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  $k=1, 2, \dots, T$ .

#### part (a)

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

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

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 equations (1) and (2).

The cost function is

$$\begin{aligned} J(T) &= \sum_{k=1}^T (y_k - \hat{y}_k)^T [R(k)]^{-1} (y_k - \hat{y}_k) \\ &= \sum_{k=1}^T (y_k - Hx(k))^T [R(k)]^{-1} (y_k - Hx(k)) \\ &= \sum_{k=1}^T (y_k - HF^k x(0))^T [R(k)]^{-1} (y_k - HF^k x(0)) \\ &= \sum_{k=1}^T \left( y_k^T [R(k)]^{-1} y_k - y_k^T [R(k)]^{-1} HF^k x(0) \right. \\ &\quad \left. - x(0)^T [F^k]^T H^T [R(k)]^{-1} y_k + x(0)^T [F^k]^T H^T [R(k)]^{-1} HF^k x(0) \right) \\ &= \sum_{k=1}^T (y_k^T R(k)^{-1} y_k - M_1(k)x(0) - x(0)^T M_2(k) + x(0)^T M_3(k)x(0)) \end{aligned}$$

Where  $M_1(k) = y_k^T R(k)^{-1} y_k HF^k$ ,  $M_2(k) = (F^k)^T H^T R(k)^{-1} y_k$ , and  $M_3(k) = (F^k)^T H^T R(k)^{-1} HF^k$ . We should note that because  $R(k)$  is symmetric its inverse is also symmetric. This means  $M_1(k)^T = M_2(k)$ . The batch LLS estimate for  $\hat{x}(0)$  is

$$\hat{x}(0) = \operatorname{argmin}_{x(0)} J(T)$$

so we need to find the  $x(0)$  that produces the minimum value for the cost function  $J(T)$ . We do this by finding the value of  $x(0)$  at which the derivative of the cost function with respect to  $x(0)$ ,  $\nabla_{x(0)} J(T)$  is equal to zero. Noting, as we did in lecture 21, slide 6 that

$$\begin{aligned}\frac{\partial}{\partial x}(x^T M x) &= 2Mx \quad \text{when } M \text{ is symmetric} \\ \frac{\partial}{\partial x} M^T x &= \frac{\partial}{\partial x} x^T M = M \quad \text{for any } M\end{aligned}$$

This means the derivative of our cost function is

$$\begin{aligned}\nabla_{x(0)} J(T) &= \sum_{k=1}^T (2M_3(k)x(0) - 2M_2(k)) \\ &= \sum_{k=1}^T (2[F^k]^T H^T [R(k)]^{-1} H F^k x(0) - 2[F^k]^T H^T R(k)^{-1} y_k)\end{aligned}$$

We can get rid of the summations by stacking the  $y_k$  vectors into a  $\mathbb{Y}$  vector, rearranging the  $R(k)^{-1}$  matrices into a block diagonal  $\mathbb{R}^{-1}$  matrix, and stacking the  $H F^k$  matrices (with dimension  $p \times n$ ) into a  $\mathbb{H}$  matrix (with dimension  $pT \times n$ ). This results in

$$\nabla_{x(0)} J(t) = 2\mathbb{H}^T \mathbb{R}^{-1} \mathbb{H} x(0) - 2\mathbb{H}^T \mathbb{R}^{-1} \mathbb{Y}$$

To find the LLS estimate  $\hat{x}(0)$ , the above equation for  $\nabla_{x(0)} J(T)$  is set equal to zero and solved for  $x(0)$ :

$$\begin{aligned}0 &= 2\mathbb{H}^T \mathbb{R}^{-1} \mathbb{H} \hat{x}(0) - 2\mathbb{H}^T \mathbb{R}^{-1} \mathbb{Y} \\ 2\mathbb{H}^T \mathbb{R}^{-1} \mathbb{Y} &= 2\mathbb{H}^T \mathbb{R}^{-1} \mathbb{H} \hat{x}(0) \\ \hat{x}(0) &= (2\mathbb{H}^T \mathbb{R}^{-1} \mathbb{H})^{-1} 2\mathbb{H}^T \mathbb{R}^{-1} \mathbb{Y}\end{aligned}$$

### part (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

$$y_A(k) = H x_A(k) + v_A(k)$$

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

where  $R_A$  has units of  $\text{m}^2$ . Using the data posted in `midterm2.problem3b.mat`, use your result from part (a) to compute the estimate  $\hat{x}_A(0)$ , and report the final state estimation error covariance matrix. Note the data provided is for time steps  $k \geq 1$  and each column  $k$  corresponds to a single  $y_k$  vector at step  $k$ .

The estimate for the initial state is  $\hat{x}_A(0) = [7121, -13647, -180, 774]^T$ . The final state estimation error covariance is

$$\text{cov}(e_{\text{LS}}) = (\mathbb{H}^T \mathbb{R}^{-1} \mathbb{H})^{-1} = \begin{bmatrix} 0.906 & -1.81 & 0.116 & -0.212 \\ -1.81 & 310 & -0.252 & 41.4 \\ 0.116 & -0.252 & 0.896 & -1.79 \\ -0.2124 & 41.1 & -1.79 & 352 \end{bmatrix}$$

**part (c)**

Suppose now there is 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  $r_B = [\xi_B, \eta_B]$ ,

$$y_D(k) = r_A(k) - r_B(k) + v_D(k)$$

$$R_D = \begin{bmatrix} 8000 & 500 \\ 500 & 8000 \end{bmatrix}$$

where  $R_D$  has units of  $\text{m}^2$  and  $v_D \sim \mathcal{N}(0, R_D)$  follows a stationary white noise process. Using the data posted in `midterm2_problem3c.mat` in the array 'yaugHist' (where each column contains a concatenated  $4 \times 1$  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  $k$  using 4 subplots per aircraft. Also on 4 separate subplots per aircraft show positive  $2\sigma$  bounds ofr each estimated state vs  $k$ . Be sure to explain how you set up the RLLS estimator to estimate  $x(0)$  and how you initialized the estimator.

We can first combine the measurements into a single measurement vector  $y(k) = [y_A(k), y_D(k)]^T$ . We can say in this case that  $y(k) = H_D x(k)$  where  $x(k) = [x_A(k), x_B(k)]^T$  and

$$H_D = \begin{bmatrix} H & 0 \\ H & -H \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \end{bmatrix}$$

We can put this in terms of  $x(0)$  if  $H_D(k)$  is

$$H_D(k) = \begin{bmatrix} HF_A^k & 0 \\ HF_A^k & -HF_B^k \end{bmatrix}$$

So the full measurement function is  $y(k) = H_D(k)x(0) + v(k)$  where  $v(k) \sim \mathcal{N}(0, R(k))$

$$R(k) = \begin{bmatrix} R_A(k) & 0 \\ 0 & R_D \end{bmatrix}$$

So we can set up our RLLS estimator for  $x(0)$  as

$$K_k = P_{k-1} H_D(k)^T (H_D(k) P_{k-1} H_D(k)^T + R(k))^{-1}$$

$$\hat{x}_k(0) = \hat{x}_{k-1}(0) + K_k (y_k - H_D(k) \hat{x}_{k-1}(0))$$

$$P_k = (I - K_k H_D(k)) P_{k-1} (I - K_k H_D(k))^T + K_k R(k) K_k^T$$

and we initialize our estimator using a batch estimate from the first 2 measurements, because the system is observable with 2 measurements.

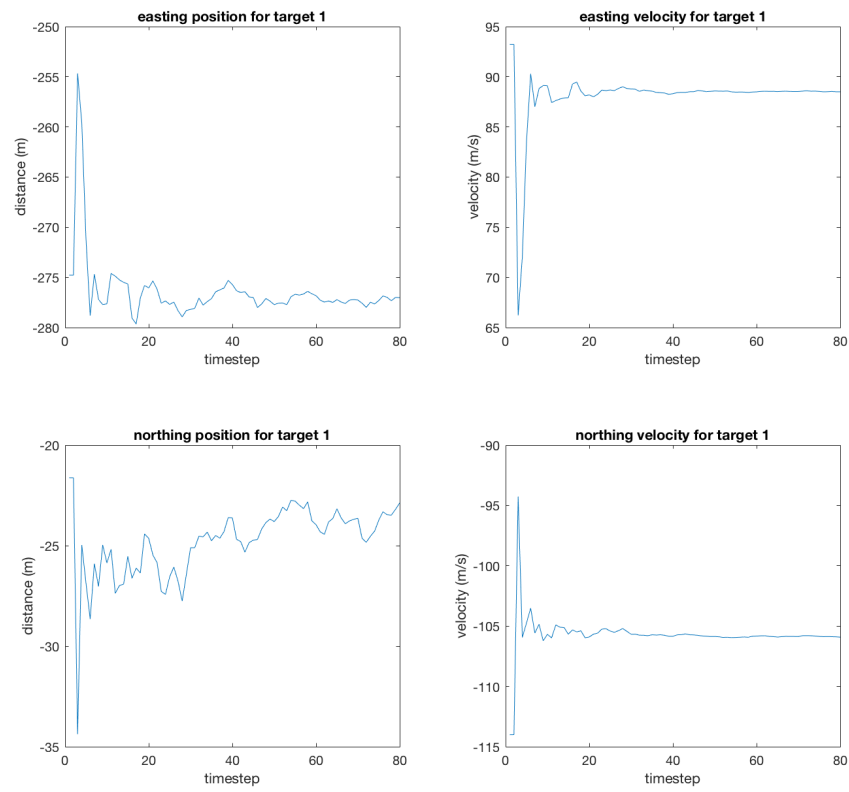


Figure 1: Estimate of initial conditions for aircraft 1

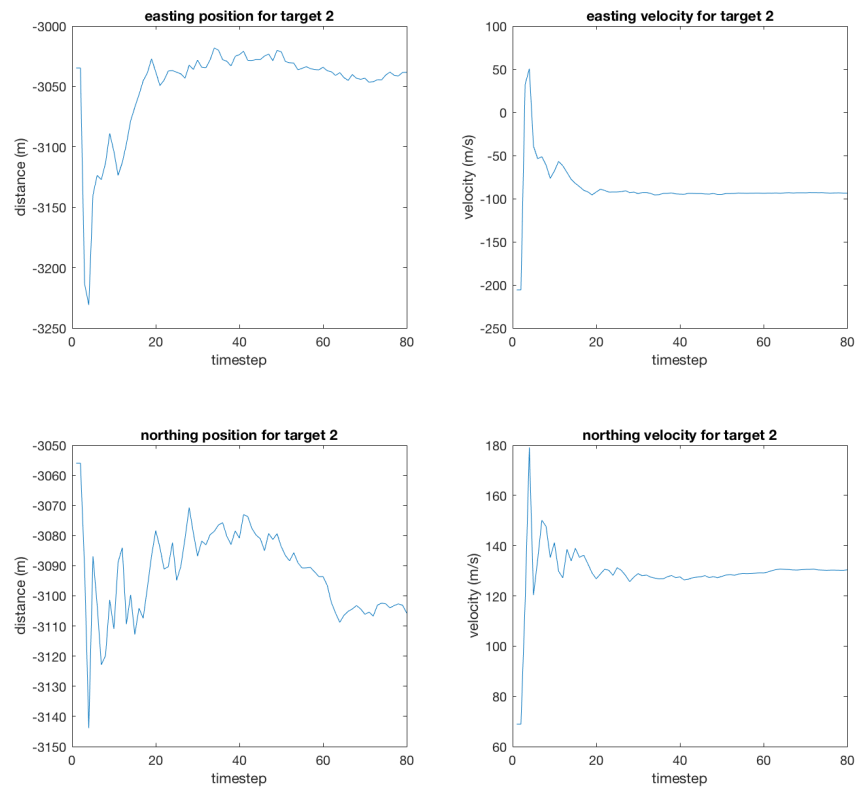


Figure 2: Estimate of initial conditions for aircraft 2

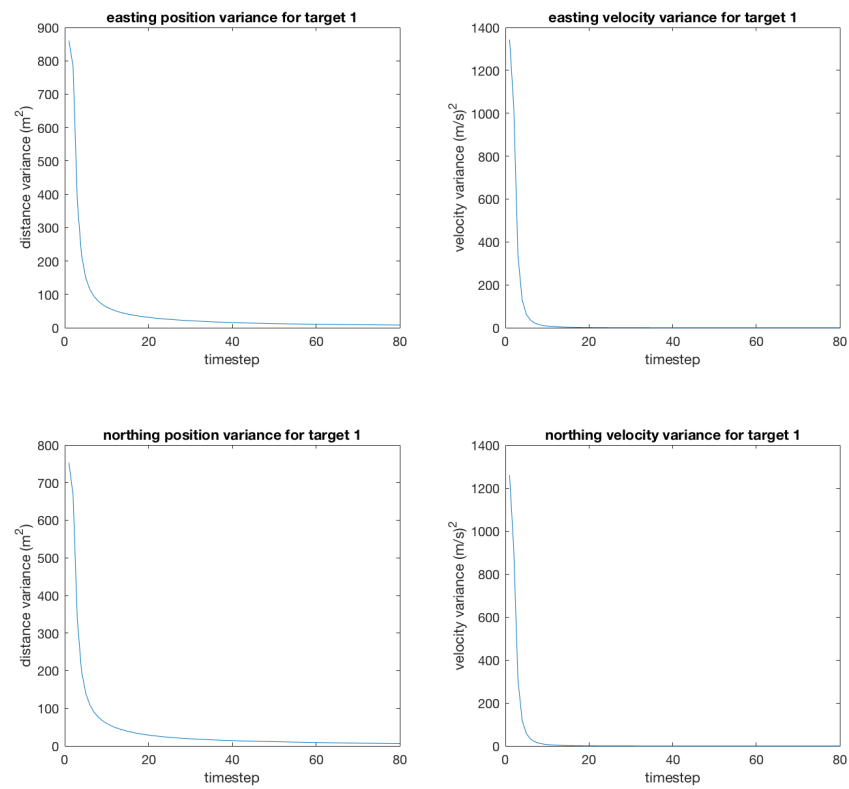


Figure 3: Variance of initial condition estimate for aircraft 1

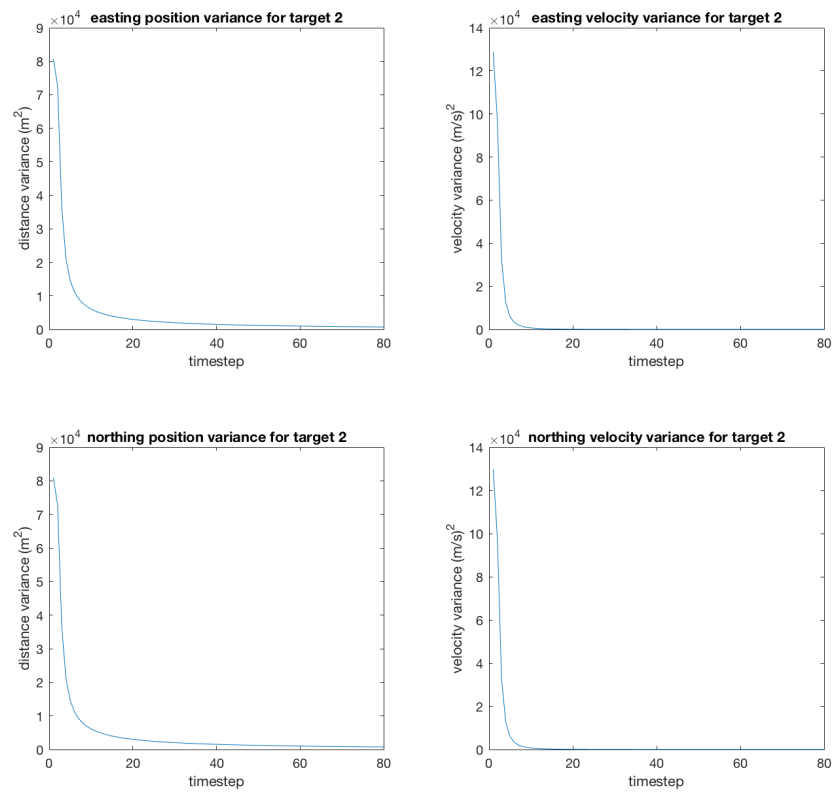


Figure 4: Variance of initial condition estimate for aircraft 2