Quiz Solutions

Table of Contents

Question 1

Problem statement

Three (3) sensors

- Each has a single noise measurement
- $z_i = x + v_i$
- Unknown *x* is a constant
- Measurement noises v_i have a mean of zero but are not necessarily uncorrleated
- Estimate of x is a linear combination of the measurements that is not a function of the unknown to be estimated, *x*
- The estimate will be unbiased – its ensemble mean will be equal to *x*

Part A (50%)

Show the equation for the gain vector $\underline{k}$ in terms of the covariance matrix R_v

Show the equation for the minimum mean square error J_{MIN} in terms of the covariance matrix R_v

Part B (50%)

Given that the measurement error covariance is

$$R_v = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{bmatrix}$$

Show the explicit expressions for the elements of the gain vector k_1, k_2, and k_3

Show the explicit expression for the minimum mean square error J_{MIN}

Solution

This problem asks you to show your understanding of Gelb's Example 1.0-1, pages 5 and 6. See the class handout notes for February 2, Section 3 for an explanation of this important fundamental example.

Problem 2

Problem statement

Part A (30%)

Show that the matrix

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$$

satisfies the polynomial equation

$$A^2 - 4 \cdot A - 5 \cdot I = 0$$

Part B (20%)

Find the eigenvalues of the matrix A of Part A.

Part C (50%)

Using the Cayley-Hamilton theorem, find the explicit forms for $a_1(t)$ and $a_2(t)$ in the equation

$$\exp(A \cdot t) = a_1(t) \cdot I + a_2(t) \cdot A$$

in terms of the eigenvalues λ_1 and λ_2. Then, give it explicitly in terms of the numbers for the eigenvalues that you found in Part B.

Solution

This is a modification of Gelb's Problem 2.5 page 47, with a new matrix. The characteristic equation of the matrix A above is given as the polynomial equation, and the eigenvalues are +5 and -1. See the class handout notes for February 9, 2006, Section 1.5, for Part C.

Problem 3

Problem Statement

We are defining the software functions for the data association problem for a tracker to be used in a dense target environment – a lot of airplanes close together, at different altitudes and flying in different directions with not a lot of ground distance between them. We are considering three cost functions for use in thresholding and tie-breaking in the data association function:
number:

Hyperpolygon

$$J_{HP} = \sum_{i=1}^{M} \frac{\left| y_i - h_i\left(\underline{\tilde{x}}\right)\right|}{\sqrt{\sigma^2\left(y_i\right) + \sigma^2\left(h_i\left(\underline{\tilde{x}}\right)\right)}} < thr_{HP}$$

Hypercube

$$J_{HC} = \underset{1 \le i \le M}{\text{Max}} \left(\frac{\left| y_i - h_i\left(\underline{\tilde{x}}\right)\right|}{\sqrt{\sigma^2\left(y_1\right) + \sigma^2\left(h_i\left(\underline{\tilde{x}}\right)\right)}} \right) < thr_{HC}$$

Hypersphere (Bhatacharya distance)

$$J_{HS} = \left(\underline{y} - \underline{h}\left(\underline{\tilde{x}}\right)\right)^T \cdot \left(H \cdot \tilde{P} \cdot H^T + R\right)^{-1} \cdot \left(\underline{y} - \underline{h}\left(\underline{\tilde{x}}\right)\right) < thr_{HS}$$

Part A (50%)

We have two measurements from a radar, range and azimuth. Show the role of the correlation between these in the equation for the Bhatacharya distance. For simplicity, use the form

$$J_{HS} = \underline{\Delta x}^T \cdot P_E^{-1} \cdot \underline{\Delta x}$$

and use this equation for P_E :

$$P_E = \begin{bmatrix} \sigma_{x1}^2 & \sigma_{x1} \cdot \sigma_{x2} \cdot \rho_{12} \\ \sigma_{x1} \cdot \sigma_{x2} \cdot \rho_{12} & \sigma_{x2}^2 \end{bmatrix}$$

Show the expression for J_{HS} expanded into its scalar terms, not as a quadratic form. Show the expanded expression again with ρ_{12} taken as zero. Explain the importance of ρ_{12} when the variance of one component of $\underline{\Delta x}$ is much larger than the other, and the correlation in the errors ρ_{12} is high (close to +1 or -1).

Part B (50%)

We have data from two radars and we want to merge the track files. We must associate track files from one radar with track files from the other radar. This operation is identical to the problem of association of radar returns to track files, except that state vectors from the first radar replace the measurement vector in the operation.

The state vectors from both radars as provided to the track file merging function are latitude, longitude, and altitude, plus velocity North, velocity East, and climb rate. Thus we have six states in the state vectors.

Evaluate the use of the hypercube or hyperpolygon instead of the hypersphere (Bhatacharya distance) in defining the cost function for association of these radar tracks. Use this as a rationale for requiring that the entire covariance matrix be provided by each radar, instead of simply the variances of each state. State this rationale as simply as possible, and use as few sentences as possible.

Solution

This problem is treated in the lecture notes handout for February 16, 2006, section 4.1.5. We begin by writing the inverse of the covariance matrix P_E ,

$$\begin{aligned} P_E^{-1} &= \frac{1}{\sigma_{x1}^2 \cdot \sigma_{x2}^2 \cdot \left(1-\rho_{12}^2\right)} \cdot \begin{bmatrix} \sigma_{x2}^2 & -\sigma_{x1} \cdot \sigma_{x2} \cdot \rho_{12} \\ -\sigma_{x1} \cdot \sigma_{x2} \cdot \rho_{12} & \sigma_{x1}^2 \end{bmatrix} \\ &= \frac{1}{1-\rho_{12}^2} \cdot \begin{bmatrix} \dfrac{1}{\sigma_{x1}^2} & -\dfrac{\rho_{12}}{\sigma_{x1} \cdot \sigma_{x2}} \\ -\dfrac{\rho_{12}}{\sigma_{x1} \cdot \sigma_{x2}} & \dfrac{1}{\sigma_{x1}^2} \end{bmatrix} \end{aligned} \tag{3.1}$$

The quadratic form requested is

$$J_{HS}\left(\rho_{12}\right) = \frac{1}{1-\rho_{12}^2} \cdot \left(\frac{\left(\Delta x_1\right)^2}{\sigma_{x1}^2} - 2 \cdot \frac{\left(\Delta x_1\right) \cdot \left(\Delta x_2\right) \cdot \rho_{12}}{\sigma_{x1} \cdot \sigma_{x2}} + \frac{\left(\Delta x_2\right)^2}{\sigma_{x2}^2} \right) \tag{3.2}$$

With the correlation taken as zero the quadratic form becomes

$$J_{HS}\left(0\right) = \frac{\left(\Delta x_1\right)^2}{\sigma_{x1}^2} + \frac{\left(\Delta x_2\right)^2}{\sigma_{x2}^2} \tag{3.3}$$

The contour traced out on a plot of Δx_2 versus Δx_1 when J_{HS} is set to a value of R is an ellipse of area

$$A_{HS}\left(R\right) = \pi \cdot R^2 \cdot \sigma_{x1} \cdot \sigma_{x2} \cdot \sqrt{1-\rho_{12}^2} \tag{3.4}$$

When ρ_{12} is high (close to +1 or -1) this area is far smaller than when ρ_{12} is small. In addition, since the localization ellipses intersect the Δx_1 and Δx_2 axes at the same points, the localization ellipse covers area outside the simple ellipses with semi-axes $R \cdot \sigma_{x1}$ and $R \cdot \sigma_2$ .when ρ_{12} is large. This means that the areas included in

$$J_{HS}\left(\rho_{12}\right) \le R^2 \tag{3.5}$$

are dramatically different for small ρ_{12} than for large ρ_{12} .

The long axis of the ellipse is approximately, when ρ_{12} is large, along the line

$$\frac{\Delta x_2}{\Delta x_1} = \frac{\sigma_{x2}}{\sigma_{x1}} \cdot \rho_{12} \tag{3.6}$$

When one variance is much larger than the other, the ellipse will lie along the axis of the variable with the larger variance.

Problem 4

Problem statement

Give the Kalman filter equations outlined below and describe each term in them.

Part A (50%)

Give the meaning of

- The state vector $\underline{x}$,
- Its extrapolated value $\underline{\tilde{x}}$, and
- Its estimated value $\underline{\hat{x}}$.

Give the Kalman filter equations as defined in the study guide for the following quantities. Explain each term.

1. The measurement model; the equation beginning with $\underline{y} = \ldots$
2. The state vector extrapolation model; the equation beginning with $\underline{\dot{x}} = \ldots$
3. The covariance extrapolation approximation; the equation beginning with $\tilde{P} \approx \ldots$
4. The Kalman gain; the equation beginning with $K = \ldots$
5. The state vector update; the equation beginning with $\underline{\hat{x}} = \ldots$
6. The covariance update; any of the equations beginning with $P = \ldots$ or $P^{-1} = \ldots$

Solution for Part A

This part is addressed in the handout lecture notes of February 16, 2006, section 2.1.

Part B (50%)

Submit only one of the two alternatives below. Designate your reply "Alternative 1" or Alternative 2." Only responses for the alternative that you designate will be graded.

Alternative 1

Give the form for the covariance update equation that you would use when first blocking out a Kalman filter in software for testing and simulation. Explain why you would use that form as opposed to each of the other alternative algebraic forms for the covariance update.

Solution for Part B Alternative 1

See the lecture notes for February 16, 2006, section 2.2. Without knowing more about any special problems that a particular application will have, the first baseline for a Kalman filter should use the Joseph stabilized form of the covariance update. If, during development, simulation, analysis, or evaluation, problems are detected that may be caused by numerical problems with their roots in the covariance update, then double precision and the normal form can be substituted. If this solves the problem, then a square root filter should be considered for that application.

Alternative 2

Use the Matrix Inversion Lemma (as given on the web site http://rowan.jkbeard.com on the Files page) to show that the Joseph stabilized form is equal to the normal form when the Kalman gain K is equal to the optimum value.

Solution for Part B Alternative 2

The Joseph stabilized form for the covariance update is

$$P = \left(I - K \cdot H\right) \cdot \tilde{P} \cdot \left(I - K \cdot H\right)^T + K \cdot R \cdot K^T \tag{4.1}$$

and the optimum value of the Kalman gain K is

$$K = \tilde{P} \cdot H^T \cdot \left(H \cdot \tilde{P} \cdot H^T + R\right)^{-1} \tag{4.2}$$

Following Gelb, p109, equation (4.2-16), we have, by first expanding (4.1), then substituting for K from (4.2), then collecting terms and then using (4.2) again to group its right hand side to again use K in the equation,

$$\begin{aligned} P &= \tilde{P} - K \cdot H \cdot \tilde{P} - \tilde{P} \cdot H^T \cdot K^T + K \cdot \left(H \cdot \tilde{P} \cdot H^T + R\right) \cdot K^T \\ &= \tilde{P} - 2 \cdot \tilde{P} \cdot H^T \cdot \left(H \cdot \tilde{P} \cdot H^T + R\right)^{-1} \cdot H \cdot \tilde{P} + \tilde{P} \cdot H^T \cdot \left(H \cdot \tilde{P} \cdot H^T + R\right)^{-1} \cdot H \cdot \tilde{P} \\ &= \tilde{P} - \tilde{P} \cdot H^T \cdot \left(H \cdot \tilde{P} \cdot H^T + R\right)^{-1} \cdot H \cdot \tilde{P} \\ &= \left(I - \tilde{P} \cdot H^T \cdot \left(H \cdot \tilde{P} \cdot H^T + R\right)^{-1} \cdot H\right) \cdot \tilde{P} \\ &= \left(I - K \cdot H\right) \cdot \tilde{P} \end{aligned} \tag{4.3}$$

which final result Gelb states is obtained "after some manipulation."

We have the matrix inversion lemma,

$$\left(A - B \cdot D^{-1} \cdot C\right)^{-1} = A^{-1} + A^{-1} \cdot B \cdot \left(D - C \cdot A^{-1} \cdot B\right)^{-1} \cdot C \cdot A^{-1} \tag{4.4}$$

When (4.2) is substituted into (4.1), a lot of expressions similar to each side of (4.4) present themselves. One of the simplest is to look at $\left(I - K \cdot H\right)$ using the optimum Kalman gain K from (4.2):

$$I - K \cdot H = I - \tilde{P} \cdot H^T \cdot \left(H \cdot \tilde{P} \cdot H^T + R\right)^{-1} \cdot H \tag{4.5}$$

If we define the matrices in (4.4) as

$$\begin{aligned} A &= I \\ B &= \tilde{P} \cdot H^T \\ C &= H \\ D &= -R \end{aligned} \tag{4.6}$$

then we immediately find that

$$\begin{aligned} I - K \cdot H &= \left(I + \tilde{P} \cdot H^T \cdot R^{-1} \cdot H \right)^{-1} \\ &= \left(\tilde{P}^{-1} + H^T \cdot R^{-1} \cdot H \right)^{-1} \cdot \tilde{P}^{-1} \end{aligned} \tag{4.7}$$

We recognize the normal form as the first term on the right hand side of (4.7) so we immediately have our result, but if we acknowledge that the normal form is derived from (4.1) and that we have not yet shown it, we can do so by combining (4.3) and (4.7).

.

Note that, if we pose the Kalman update as an estimation process from $\underline{\tilde{x}}$ and its error covariance $\tilde{P}$, the data measurements $\underline{y}$ and its error covariance R and the measurement model $\underline{h}(\underline{x})$, the normal form of the covariance is the first equation produced for the covariance update as the Cramer-Rao bound. The form given by (4.1) follows from the update equation

$$\begin{aligned} \underline{\hat{x}} &= \underline{\tilde{x}} + K \cdot \left(\underline{y} - \underline{h}\left(\underline{\tilde{x}} \right) \right) \\ &= \underline{\tilde{x}} + K \cdot \left(\underline{y} - \underline{h}\left(\underline{x} \right) - H \cdot \left(\underline{\tilde{x}} - \underline{x} \right) \right) \\ &= \underline{x} + \left(I - K \cdot H \right) \cdot \left(\underline{\tilde{x}} - \underline{x} \right) + K \cdot \underline{v} \end{aligned} \tag{4.8}$$

by subtracting the true states $\underline{x}$ from both sides and taking the covariance to yield (4.1). Then, with (4.7), we have the short form.