Homework 8 Solutions

1. ***Recursive estimation for navigation***

We have a ship at position

$$x = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

and beacons located at

$$q_1 = \begin{bmatrix} 500 \\ 0 \end{bmatrix} \qquad q_2 = \begin{bmatrix} 0 \\ 500 \end{bmatrix} \qquad q_3 = \begin{bmatrix} 500 \\ -200 \end{bmatrix} \qquad q_4 = \begin{bmatrix} -500 \\ 500 \end{bmatrix}$$

We take measurements at times $t = 1, \dots, 6$. At each time t, $r(t)$ is a vector of range measurements, and $b(t)$ is a vector indicating which beacons were used. e.g.

$$r(1) = \begin{bmatrix} 498.373 \\ 497.219 \\ 498.116 \end{bmatrix} \qquad b(1) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

indicates that at time 1, the distance to beacon 1 was measured at 498.373, the distance to beacon 2 was measured at 497.18 and the distance to beacon 1 was measured again at 498.116.

Measurements to beacon q_i are corrupted by Gaussian noise of covariance c_i, where

$$c_1 = 0.1 \qquad c_2 = 0.15 \qquad c_3 = 0.2 \qquad c_4 = 0.1$$

Each scalar measurement is corrupted by noise independent from that affecting any other scalar measurement. So in the above example, we would have a covariance matrix for the measurement of

$$\Sigma = \begin{bmatrix} 0.1 & & \\ & 0.15 & \\ & & 0.1 \end{bmatrix}$$

The data is

$$b(1) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad b(2) = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad b(3) = 2 \quad b(4) = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} \quad b(5) = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \quad b(6) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

and

$$r(1) = \begin{bmatrix} 498.3732 \\ 497.2192 \\ 498.1156 \end{bmatrix} \quad r(2) = \begin{bmatrix} 498.4453 \\ 537.5662 \end{bmatrix} \quad r(3) = 496.9780$$

$$r(4) = \begin{bmatrix} 536.9246 \\ 537.6320 \\ 537.7702 \end{bmatrix} \quad r(5) = \begin{bmatrix} 706.2695 \\ 496.9436 \end{bmatrix} \quad r(6) = \begin{bmatrix} 497.5740 \\ 496.8541 \\ 538.5130 \end{bmatrix}$$

Denote by $\hat{x}(t)$ the minimum-variance unbiased estimate of x given measurements $y(1), \dots, y(t)$, and by $Q(t)$ the measurement covariance.

(a) Find the MMSE estimate $\hat{x}(1)$ and $Q(1)$ assuming the prior covariance is infinitely large.

(b) Find, using recursive estimation, $\hat{x}(2), \dots, \hat{x}(6)$, and $Q(2), \dots, Q(6)$. The 90% confidence ellipsoid at time t is

$$E(t) = \left\{ z \in \mathbb{R}^2 \mid (z - \hat{x}(t))^T Q(t)^{-1} (z - \hat{x}(t)) \le F^{-1}_{\chi^2_2}(0.9) \right\}$$

For each t plot $\hat{x}(t)$ and the 90% confidence ellipsoid.

(c) Form the vector y and matrix A such that

$$y = Ax + w$$

where y contains all measurements. Compute the (non-recursive) MMSE estimate for x, and verify that it equals $\hat{x}(6)$ that you computed recursively above.

(d) Recall that an $n \times n$ matrix J is called a *permutation* matrix if every row and every column contains exactly one 1, with every other entry 0. If we have

$$y_{\text{perm}} = Jy$$

Then y_{perm} just contains the entries of y in a different order. Use this to show that if we receive measurements in a different order, the final estimate $\hat{x}(6)$ is unchanged.

(e) Is it true that $E(t) \subset E(t+1)$? That is, when we receive a new measurement, will the new confidence ellipsoid be inside the old one? Give a proof, or a counterexample.

Solution.

(a) For each t, we construct

$$A_t = \begin{bmatrix} \frac{q_{b(t)_1}^T}{\|q_{b(t)_1}\|} \\ \frac{q_{b(t)_2}^T}{\|q_{b(t)_2}\|} \\ \vdots \end{bmatrix} \qquad y(t) = \begin{bmatrix} \|q_{b(t)_1}\| - r(t)_1 \\ \|q_{b(t)_2}\| - r(t)_2 \\ \vdots \end{bmatrix}$$

so that, for the first time step, we have

$$y(1) = A_1 x + w(1)$$

Since both x and $w(0)$ are independent Gaussians, we can find the MMSE estimate for x given $y(1)$, as

$$x|y(1) \sim N(\hat{x}(1), Q(1))$$

where

$$\begin{aligned} \hat{x}(1) &= \hat{x}(0) + Q(0)A_1^T(A_1 Q(0) A_1^T + \Sigma_{w(1)})^{-1}(y(1) - A_1\hat{x}(0)) \\ &= \hat{x}(0) + (Q(0)^{-1} + A_1^T \Sigma_{w(1)}^{-1} A_1)^{-1} A_1^T \Sigma_{w(1)}^{-1}(y(1) - A_1\hat{x}(0)) \\ Q(1) &= Q(0) - Q(0)A_1^T(A_1 Q(0) A_1^T + \Sigma_{w(1)})^{-1} A_1 Q(0) \\ &= (Q(0)^{-1} + A_1^T \Sigma_{w(1)}^{-1} A_1)^{-1} \end{aligned}$$

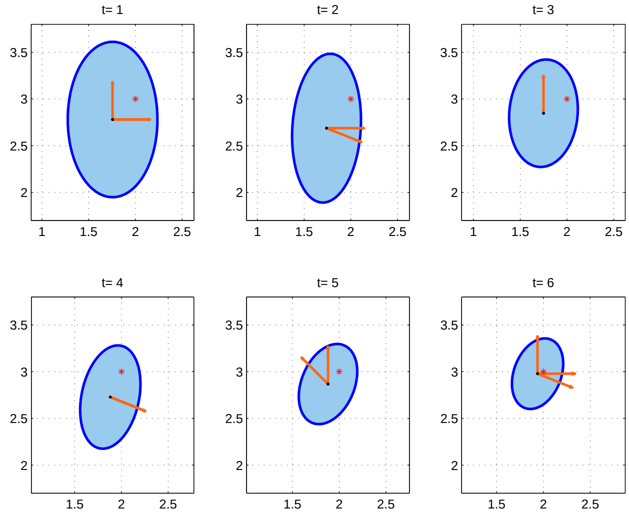
By assumption, $\hat{x}(0) = 0$ and $Q(0)$ is infinitely large. Thus, our posterior distribution reduces to

$$\begin{aligned} \hat{x}(1) &= (A_1^T \Sigma_{w(1)}^{-1} A_1)^{-1} A_1^T \Sigma_{w(1)}^{-1} y(1) \\ Q(1) &= (A_1^T \Sigma_{w(1)}^{-1} A_1)^{-1} \end{aligned}$$

(b) Using the same updates as above, we can recursively compute our estimates as

$$\begin{aligned} \hat{x}(t+1) &= \hat{x}(t) + Q(t)A_{t+1}^T(A_{t+1}Q(t)A_{t+1}^T + \Sigma_{w(t+1)})^{-1}(y(t+1) - A_{t+1}\hat{x}(t)) \\ Q(t+1) &= Q(t) - Q(t)A_{t+1}^T(A_{t+1}Q(t)A_{t+1}^T + \Sigma_{w(t+1)})^{-1} A_{t+1} Q(t) \end{aligned}$$

The 90% confidence ellipses, centered about the estimates, are plotted below.



(c) Define y, A, and w as

$$y = \begin{bmatrix} y(1) \\ \vdots \\ y(6) \end{bmatrix} \qquad A = \begin{bmatrix} A_1 \\ \vdots \\ A_6 \end{bmatrix} \qquad w = \begin{bmatrix} w(1) \\ \vdots \\ w(6) \end{bmatrix}$$

From our assumptions of part a, we have

$$\begin{aligned} \hat{x}(6) &= (A^T \Sigma_w^{-1} A)^{-1} A^T \Sigma_w^{-1} y \\ Q(6) &= (A^T \Sigma_w^{-1} A)^{-1} \end{aligned}$$

We can verify that the estimate here matches the estimator from part b. Namely,

$$\hat{x}(6) = \begin{bmatrix} 1.94 \\ 2.98 \end{bmatrix}$$

(d) If we use the permutation matrix J to rearrange the order of y, we obtain the system

$$y_{\text{perm}} = Jy = JAx + Jw$$

Once again, the posterior distribution for x given y_{perm} is defined by

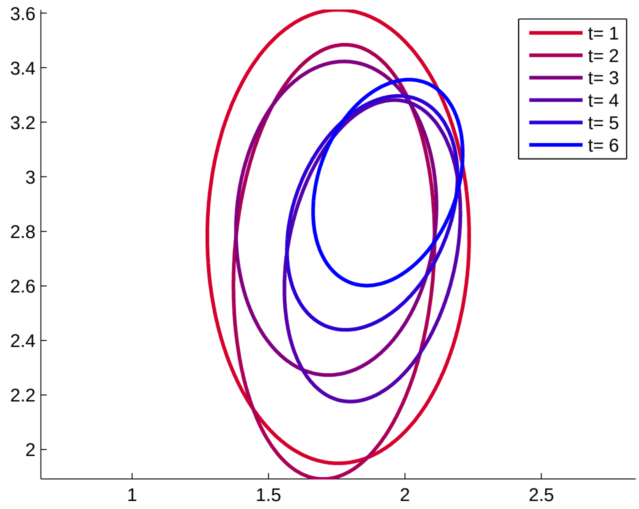
$$\begin{aligned} \hat{x}(6) &= ((JA)^T (J\Sigma_w J^T)^{-1} JA)^{-1} (JA)^T (J\Sigma_w J^T)^{-1} y_{\text{perm}} \\ &= (A^T J^T J \Sigma_w^{-1} J^T JA)^{-1} A^T J^T J \Sigma_w^{-1} J^T J y \\ Q(6) &= ((JA)^T (J\Sigma_w J^T)^{-1} JA)^{-1} \\ &= (A^T J^T J \Sigma_w^{-1} J^T JA)^{-1} \end{aligned}$$

Using the fact that $J^T J = JJ^T = I$, this reduces to

$$\begin{aligned} \hat{x}(6) &= (A^T \Sigma_w^{-1} A)^{-1} A^T \Sigma_w^{-1} y \\ Q(6) &= (A^T \Sigma_w^{-1} A)^{-1} \end{aligned}$$

which matches our previous estimate for x.

(e) It is readily apparent from our plots of part b that one ellipse need not contain the subsequent ellipses. To highlight this fact, we've plotted those confidence ellipses on top of each other below.



2. ***Downdating of estimates***

We have seen that one can use recursive estimation to update an estimate using new measurements. In this question, we look at what happens if you find out that one of your old measurements was incorrect. The effect of the incorrect measurement can be removed by *downdating*.

Suppose we have

$$y = Ax + w \qquad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \qquad A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \qquad w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

where $w \sim \mathcal{N}(0, I)$, $x \sim \mathcal{N}(0, \Sigma_x)$, $y_1 \in \mathbb{R}^{m_1}$ and $y_2 \in \mathbb{R}^{m_2}$. We form $\hat{x}_{12}$, the MMSE estimate of x given measurements y_1 and y_2 using the standard MMSE

$$\hat{x}_{12} = Q_{12} A^T y$$
$$Q_{12} = (A^T A + \Sigma_x^{-1})^{-1}$$

After doing this, we realize that y_1 was really an erroneous measurement, and we'd prefer to use the estimate $\hat{x}_2$ based on y_2 only, given by

$$\hat{x}_2 = Q_2 A_2^T y_2$$
$$Q_2 = (A_2^T A_2 + \Sigma_x^{-1})^{-1}$$

Show that one may find Q_2 and $\hat{x}_2$ using

$$Q_2 = Q_{12} + Q_{12} A_1^T (I - A_1 Q_{12} A_1^T)^{-1} A_1 Q_{12}$$
$$\hat{x}_2 = \hat{x}_{12} - Q_{12} A_1^T (I - A_1 Q_{12} A_1^T)^{-1} (y_1 - A_1 \hat{x}_{12})$$

This removes the effect of y_1, and gives the optimal estimate just based on y_2. Notice that one only needs the measurement y_1 that is being removed, not the other measurement y_2.

Solution.

From standard MMSE analysis, we note that

$$Q_{12} = \mathrm{cov}(x|y_1, y_2) \qquad Q_2 = \mathrm{cov}(x|y_2)$$

We can approach this problem as a two step estimation process. We first make an estimate of $x|y_2$, then recursively make an estimate of $x|y_1, y_2$.

Using the Kalman filter results, we know the measurement update produces

$$\begin{aligned} Q_{12} &= Q_2 - Q_2 A_1^T (A_1 Q_2 A_1^T + I)^{-1} A_1 Q_2 \\ \hat{x}_{12} &= \hat{x}_2 + Q_2 A_1^T (A_1 Q_2 A_1^T + I)^{-1} (y_1 - A_1 \hat{x}_2) \end{aligned}$$

Let us first analyze the covariance update. Rewrite it as

$$Q_{12} = (Q_2^{-1} + A_1^T A_1)^{-1}$$

Then

$$\begin{aligned} Q_2 &= (Q_{12}^{-1} - A_1^T A_1)^{-1} \\ &= Q_{12} + Q_{12} A_1^T (I - A_1 Q_{12} A_1^T)^{-1} A_1 Q_{12} \end{aligned}$$

as desired.

For the estimate update, we rearrange the expression

$$\begin{aligned} \hat{x}_{12} &= (I + Q_2 A_1^T A_1)^{-1} \hat{x}_2 + Q_2 A_1^T (A_1 Q_2 A_1^T + I)^{-1} y_1 \\ \hat{x}_2 &= (I + Q_2 A_1^T A_1) \hat{x}_{12} - (I + Q_2 A_1^T A_1) Q_2 A_1^T (A_1 Q_2 A_1^T + I)^{-1} y_1 \\ &= \hat{x}_{12} - Q_2 A_1^T (y_1 - A_1 \hat{x}_{12}) \end{aligned}$$

Plugging in for Q_2 yields

$$\begin{aligned} \hat{x}_2 &= \hat{x}_{12} - Q_{12} A_1^T - Q_{12} A_1^T (I - A_1 Q_{12} A_1^T)^{-1} A_1 Q_{12} A_1^T (y_1 - A_1 \hat{x}_{12}) \\ &= \hat{x}_{12} - Q_{12} A_1^T (I - A_1 Q_{12} A_1^T)^{-1} (y_1 - A_1 \hat{x}_{12}) \end{aligned}$$

as desired.