Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Institute for Dynamic Systems and Control
IDSC
Institut für Dynamische Systeme und Regelungstechnik

Final Exam **August 23rd, 2010**

Introduction to Recursive Filtering and Estimation (151-0566-00) **Prof. R. D'Andrea**

Solutions

Exam Duration: **150 minutes**

Number of Problems: **5**

Permitted aids: One A4 sheet of paper.

Use only the provided sheets for your solutions.

Problem 1 20%

A group of friends have plans to go sailing on Sunday, but their plans are dependent on the weather. They consider both wind and rainfall and assume rain and wind do not influence each other. The sailing conditions are best if there is wind but no rain. If it is windy and not raining, then there is a 90% chance that they will go sailing. If it is not windy and not rainy, then there is a 50% chance of going, since it is likely that wind conditions change quickly. If it is raining, then, no matter what the wind conditions are, the probability that they will go sailing drops to 10%.

The weather prediction on Saturday calls for a 40% chance of rain and a 10% chance of wind on Sunday.

Given that we know that the friends eventually did go sailing, how likely is it that it was windy and not raining on Sunday?

Solution 1

Introduce the discrete random variables r, w, and g representing rainfall, wind conditions, and the friends' decision, where $r \in \{1, 0\} = \{\text{raining, not raining}\}$, $w \in \{1, 0\} = \{\text{windy, not windy}\}$, and $g \in \{1, 0\}\{\text{going sailing, not going sailing}\}$. From the text, the following probabilities can be extracted:

$$\begin{aligned}
&\Pr(g = 1|r = 0, w = 1) = 0.9 \\
&\Pr(g = 1|r = 0, w = 0) = 0.5 \\
&\Pr(g = 1|r = 1) = 0.1
\end{aligned}$$

$$\begin{aligned}
&\Pr(r = 1) = 0.4 \\
&\Pr(w = 1) = 0.1.
\end{aligned}$$

We are interested in the probability $\Pr(r = 0, w = 1|g = 1)$. With Bayes' Theorem and assuming independence between r and w, the probability $\Pr(r = 0, w = 1|g = 1)$ can be related to the given probabilities,

$$\begin{aligned}
\Pr(r = 0, w = 1|g = 1) &= \frac{\Pr(g = 1|r = 0, w = 1) \Pr(r = 0, w = 1)}{\Pr(g = 1)} \\
&= \frac{\Pr(g = 1|r = 0, w = 1) \Pr(r = 0) \Pr(w = 1)}{\Pr(g = 1)}.
\end{aligned}$$

With the law of total probability,

$$\begin{aligned}
\Pr(g = 1) = &\ \Pr(g = 1|r = 0, w = 1) \Pr(r = 0) \Pr(w = 1) \\
&+ \Pr(g = 1|r = 0, w = 0) \Pr(r = 0) \Pr(w = 0) \\
&+ \Pr(g = 1|r = 1) \Pr(r = 1) \\
= &\ 0.9 \cdot (1 - 0.4) \cdot 0.1 + 0.5 \cdot (1 - 0.4) \cdot (1 - 0.1) + 0.1 \cdot 0.4 \\
= &\ 0.364.
\end{aligned}$$

Then,

$$\Pr(r = 0, w = 1|g = 1) = \frac{0.9 \cdot (1 - 0.4) \cdot 0.1}{0.364} = \frac{27}{182} \approx 15\%.$$

Note that

$$\Pr(r = 0, w = 1) = 0.06 = 6\%.$$

The decision of the friends to go sailing (which is based on the weather on Sunday) gives us additional information about the weather conditions on Sunday.

Problem 2 20%

In a manufacturing process, the mass m of a part is to be estimated using measurements from two different scales.

The mass can take on positive integer values $\{0, 1, 2, 3, \dots\}$ corresponding to the mass quantization units. The measurements z_1 and z_2 of the part's mass from scale 1 and scale 2, respectively, are affected by the additive measurement errors w_1 and w_2, that is

$$z_1 = m + w_1$$
$$z_2 = m + w_2.$$

The measurement errors can be modeled as independent discrete random variables taking integer values with the probability density functions given in Fig. 1.

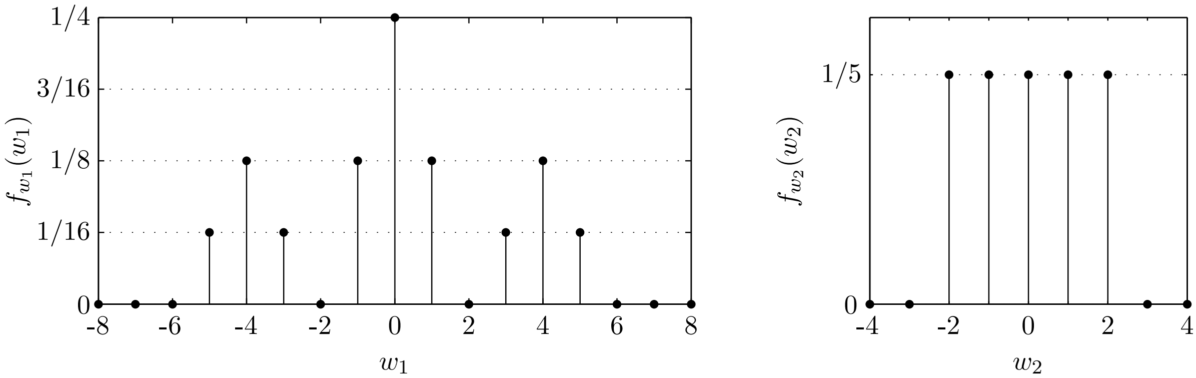

Figure 1: Measurement error probability density functions.

a) Calculate the *maximum likelihood estimate* of the part's mass for the following measurements

i) $z_1 = 8$, $z_2 = 10$

ii) $z_1 = 8$, $z_2 = 11$.

The part under consideration is taken from a set of parts whose mass is know to be distributed according to the probability density function in Fig. 2.

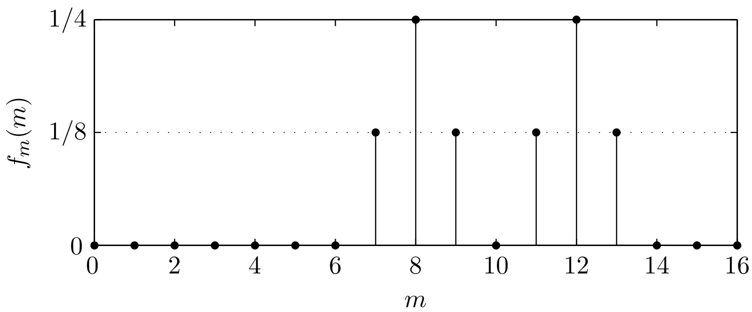

Figure 2: Mass probability density function.

b) Calculate the *maximum a-posteriori estimate* of the part's mass taking into account the part's distribution in Fig. 2 and given the measurements $z_1 = 8$, $z_2 = 11$.

Solution 2

a) The maximum likelihood estimate is defined by

$$\hat{m}_{ML} = \arg\max_{m} f_{z_1,z_2|m}(z_1, z_2|m).$$

Since the measurement errors w_1 and w_2 are independent, the measurements z_1 and z_2 are conditionally independent; hence

$$f_{z_1,z_2|m}(z_1, z_2|m) = f_{z_1|m}(z_1|m)\, f_{z_2|m}(z_2|m).$$

Using $w_i = z_i - m$, $i = 1, 2$, we obtain

$$f_{z_1,z_2|m}(z_1, z_2|m) = f_{w_1}(z_1 - m)\, f_{w_2}(z_2 - m),$$

which is to be maximized over m. In the following, we present a graphical solution for the optimization problem.

i) For the measurements $z_1 = 8$, $z_2 = 10$ the corresponding functions $f_{w_1}(z_1 - m)$, $f_{w_2}(z_2 - m)$ and their product is shown in Fig. 3. From the graph on the right, it is

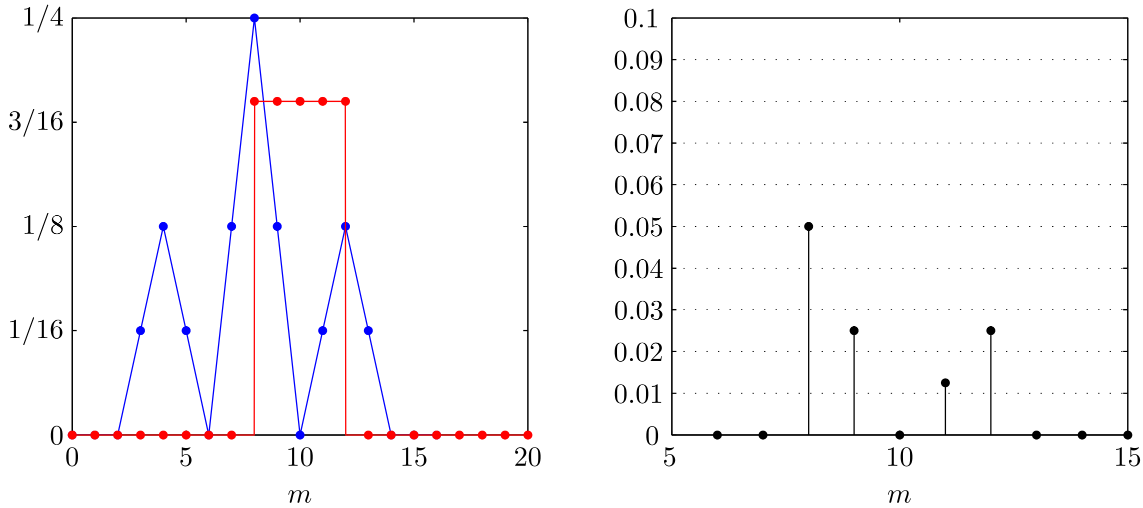

Figure 3: Left: likelihood functions $f_{w_1}(8-m)$ (blue) and $f_{w_2}(10-m)$ (red). Right: the product of the two, $f_{w_1}(8 - m)\, f_{w_2}(10 - m)$, which is to be maximized.

obvious that the maximum is attained at $m = 8$. Therefore,

$$\hat{m}_{ML} = \arg\max_{m} f_{w_1}(8 - m)\, f_{w_2}(10 - m) = 8.$$

ii) For the measurements $z_1 = 8$, $z_2 = 11$, the same graphs as in (i) are shown in Fig. 4.

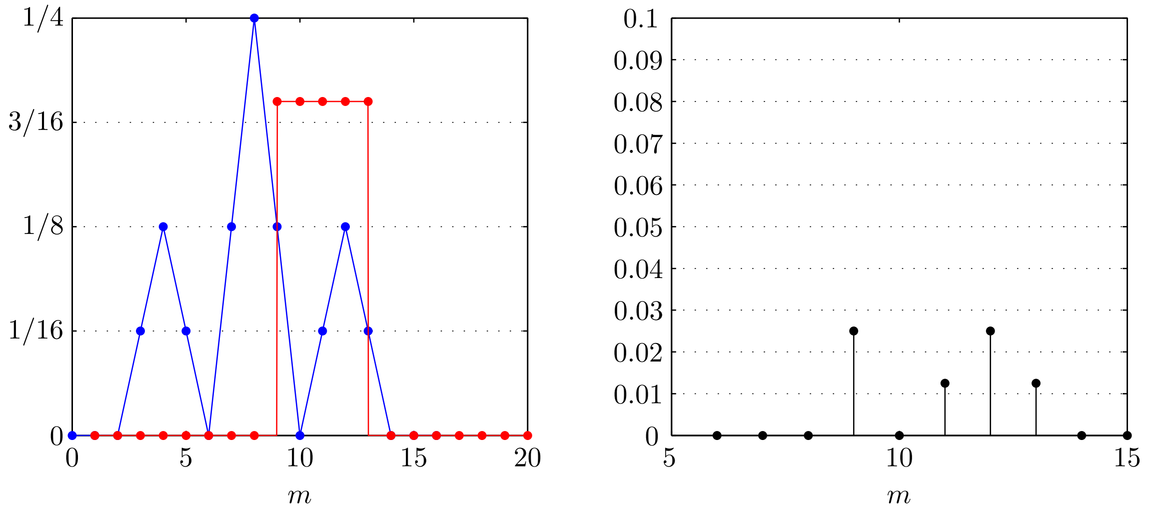

Figure 4: Left: likelihood functions $f_{w_1}(8-m)$ (blue) and $f_{w_2}(11-m)$ (red). Right: the product of the two, $f_{w_1}(8-m)\, f_{w_2}(11-m)$, which is to be maximized.

In this case, obviously,

$$\hat{m}_{ML} = \arg\max_m f_{w_1}(8-m)\, f_{w_2}(11-m) = \{9, 12\}.$$

b) The maximum a posteriori estimate is defined by

$$\hat{m}_{MAP} = \arg\max_m f_{z_1,z_2|m}(z_1, z_2|m)\, f_m(m).$$

Using the result from part (a) for $z_1 = 8$, $z_2 = 11$ in Fig. 4, we obtain the likelihood function $f_{z_1,z_2|m}(z_1, z_2|m) = f_{w_1}(z_1 - m)\, f_{w_2}(z_2 - m)$ and probability density function $f_m(m)$ as shown in Fig. 5.

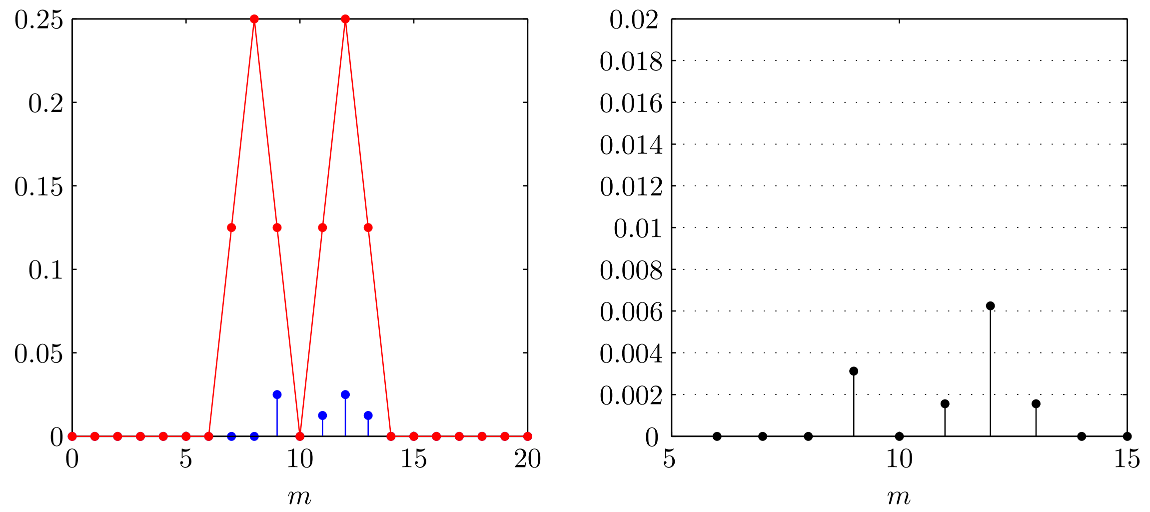

Figure 5: Left: likelihood function $f_{w_1}(8-m)\, f_{w_2}(11-m)$ (blue) and parameter distribution $f_m(m)$ (red). Right: the product of the two, $f_{w_1}(8-m)\, f_{w_2}(11-m)\, f_m(m)$, which is to be maximized.

From this, it is obvious that

$$\begin{aligned}\hat{m}_{MAP} &= \arg\max_{m} f_{z_1,z_2|m}(8, 11|m)\, f_m(m)\\ &= \arg\max_{m} f_{w_1}(8-m)\, f_{w_2}(11-m)\, f_m(m)\\ &= 12.\end{aligned}$$

Problem 3 **20%**

Recall the basic particle filtering algorithm we derived in class:

- Initialize the algorithm by randomly drawing N samples from $f_{x(0)}(x(0))$, that is, the probability density function of the initial state $x(0)$.

  Obtain $x(n, 0|0),\ n = 1, 2, \ldots, N$.

- Step 1: Simulate the N particles via the process equation.

  Obtain the *a priori* particles $x(n, k|k-1)$.

- Step 2: After a new measurement $z(k)$ at time k, scale each *a priori* particle by the measurement likelihood, normalize the resulting relative likelihoods, and obtain a corresponding weight β_n for each particle.

  Resample to get the *a posteriori* particles $x(n, k|k)$ that have equal weights. Go to Step 1.

Suppose you have a measurement $z(k) = x(k)/w(k)$, where $w(k)$ is uniformly distributed on the interval $[0.9, 1.1]$. Suppose that, at time k, five *a priori* particles $x(n, k|k-1),\ n = 1, 2, 3, 4, 5$, are given as 0.6, 0.8, 1, 1.2, and 1.4, and that the measurement is obtained as $z(k) = 1$.

a) What are the weights β_n of the particles given the measurement $z(k) = 1$?

b) Which particles are obtained after the resampling step?

c) How would you improve the basic algorithm? Give one possible improvement.

Solution 3

a) The weights β_n are given as

$$\beta_n = \alpha f(z(k)|x(n,k|k-1)) \qquad \text{with} \quad \alpha = \left(\sum_{n=1}^{5} f\left(z(k)|x(n,k|k-1)\right) \right)^{-1}.$$

From the measurement equation, we obtain the relation

$$w(k) = x(k)/z(k)$$

and, thus,

$$f_{z(k)|x(n,k|k-1)}\left(z(k)|x(n,k|k-1)\right) = f_{w(k)}\left(\, w(k) = x(n,k|k-1)/z(k)\,\right).$$

For the given particles,

n	$x(n,k\|k-1)/z(k) = x(n,k\|k-1)$	$f(z(k)\|x(n,k\|k-1))$
1	0.6	0
2	0.8	0
3	1	$1/0.2 = 5$
4	1.2	0
5	1.4	0

Finally, $\alpha = 0.2$ and $\beta_{1,2,4,5} = 0$, $\beta_3 = 1$.

b) Given that all weight is put on particle 3, resampling results in $x(n,k|k) = 1$ for $n = 1,2,3,4,5$, where, now, each particle has a probability of $1/5$.

c) In order to avoid sample impoverishment, i.e. all particles converge to the same value, we can perturb the particles $x(n,k|k)$ by an additive noise. This is called *roughening*. Another possibility is to *adapt the noise parameters*. For this problem, the noise interval $[0.9, 1.1]$ could be extended to increase the robustness of the algorithm. Other options (not covered in class) are *prior editing*, *Markov chain Monte Carlo resampling*, *regularized particle filtering*, and many more, cf. 'Optimal State Estimation' by Dan Simon.

Problem 4 20%

You have two accounts at a bank; their balances are denoted by x_1 and x_2, respectively, and assumed to be real numbers (negative values are allowed).
Every day k, there is a flow of money (inflow or outflow) on each of the accounts that you cannot control. The money flows are independent and can be modeled relatively accurately as Gaussian random variables with zero mean and unit variance.
Your accountant informs you about your financial situation every day. However, he does not disclose the actual balances of your accounts, but only tells you the sum of the two and the difference of the two, that is, the first account minus the second. Moreover, since his arithmetic skills are not very well developed, he usually miscalculates the sum and the difference. His calculation mistakes can be modeled as additive errors that are independent Gaussian random variables with zero mean and variance equal to two. The numbers the accountant tells you for the sum and the difference of your two accounts are denoted $z_1(k)$ and $z_2(k)$, respectively.
Since you are dissatisfied with the accountant's information, but you do not want to change your bank, you decide to build an estimator to keep track of the balances $x_1(k)$ and $x_2(k)$ of your two accounts.

a) Derive the Kalman filter equations for the given problem, that is, state the prior update equations and the measurement update equations for the mean and the variance of the estimates of $x_1(k)$ and $x_2(k)$.

After opening your accounts on day $k = 0$, you know for sure that there is no money on your accounts at the end of that day. On the following two days, the accountant discloses the information given in Table 1 to you.

day k	$z_1(k)$ (erroneous account sum)	$z_2(k)$ (erroneous account difference)
1	1	0
2	2	-1

Table 1: The accountant's information.

b) Using your estimator from part (a) calculate the Kalman filter estimates $\hat{x}_1(k|k)$ and $\hat{x}_2(k|k)$ for $k = 1, 2$, that is, the Kalman filter estimates of the account balances x_1 and x_2, respectively, at the end of day k using all the information the accountant has disclosed up to and including day k.

Now, you want to implement an account balance estimator on your mobile phone in order to always have your account information available. Since your phone is an older model you are concerned with the computational efficiency of your algorithm. You decide to implement a steady-state estimator, that is, an estimator of the form

$$\begin{bmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \end{bmatrix} = A \begin{bmatrix} \hat{x}_1(k-1) \\ \hat{x}_2(k-1) \end{bmatrix} + B \begin{bmatrix} z_1(k) \\ z_2(k) \end{bmatrix},$$

where A and B are constant matrices of appropriate dimensions and $\hat{x}_1(k)$ and $\hat{x}_2(k)$ are the estimates of $x_1(k)$ and $x_2(k)$.

c) How would you design this estimator, that is, how would you choose the matrices A and B?

Hint: It is enough to state the equations that uniquely define the matrices A and B; you do not have to solve them.

Solution 4

a) The process equation for the given problem reads

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

where $x(k) := [x_1(k), x_2(k)]^T$ is the state vector and $v(k)$ is process noise (the random inflow and outflow) with $E[v(k)] = 0$ and $E[v(k)(v(k))^T] = I$. The measurement equation reads

$$z(k) = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{=:H} x(k) + w(k), \tag{2}$$

where $z(k) := [z_1(k), z_2(k)]^T$ and $w(k)$ is measurement noise (the accountant's calculation errors) with $E[w(k)] = 0$ and $E[w(k)(w(k))^T] = 2I$.

Using the notation introduced in class, the Kalman filter equations for the system (1), (2) are as follows:

Prior update (S1):

$$\begin{aligned} \hat{x}(k|k-1) &= \hat{x}(k-1|k-1) \\ P(k|k-1) &= P(k-1|k-1) + I \end{aligned}$$

Measurement update (S2):

$$\begin{aligned} K(k) &= P(k|k-1)H^T\big(HP(k|k-1)H^T + 2I\big)^{-1} \\ \hat{x}(k|k) &= \hat{x}(k|k-1) + K(k)\big(z(k) - H\hat{x}(k|k-1)\big) \\ P(k|k) &= \big(I - K(k)H\big)P(k|k-1)\big(I - K(k)H\big)^T + 2K(k)K^T(k). \end{aligned}$$

b) Since both account balances are *known* to be 0 at initial time $k = 0$, we have the initial conditions $\hat{x}(0|0) = 0$ and $P(0|0) = 0$ to start the Kalman filter recursion. We obtain the following numbers for the first two recursion steps:

Step $k = 0$:

$$\begin{aligned} \hat{x}(0|0) &= 0 \\ P(0|0) &= 0 \end{aligned}$$

Step $k = 1$:

$$\begin{aligned} \hat{x}(1|0) &= \hat{x}(0|0) = 0 \\ P(1|0) &= P(0|0) + I = I \end{aligned}$$

$$\begin{aligned} K(1) &= H^T\big(HH^T + 2I\big)^{-1} \\ &= \frac{1}{4}H^T \\ \hat{x}(1|1) &= \hat{x}(1|0) + K(1)\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} - H\hat{x}(1|0)\right) \\ &= \frac{1}{4}H^T \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\ P(1|1) &= \big(I - \frac{1}{4}H^TH\big)I\big(I - \frac{1}{4}H^TH\big)^T + 2\frac{1}{4}H^T\frac{1}{4}H \\ &= \frac{1}{2}I \end{aligned}$$

Step $k = 2$:

$$\hat{x}(2|1) = \hat{x}(1|1) = \frac{1}{4}\begin{bmatrix}1\\1\end{bmatrix}$$
$$P(2|1) = P(1|1) + I = \frac{3}{2}I$$

$$K(2) = \frac{3}{2}H^T\big(3I + 2I\big)^{-1}$$
$$= \frac{3}{10}H^T$$
$$\hat{x}(2|2) = \frac{1}{4}\begin{bmatrix}1\\1\end{bmatrix} + \frac{3}{10}H^T\left(\begin{bmatrix}2\\-1\end{bmatrix} - H\frac{1}{4}\begin{bmatrix}1\\1\end{bmatrix}\right)$$
$$= \begin{bmatrix}0.4\\1\end{bmatrix}$$

Hence, the sought estimates are $\hat{x}_1(1|1) = 0.25$, $\hat{x}_2(1|1) = 0.25$, $\hat{x}_1(2|2) = 0.4$, and $\hat{x}_2(2|2) = 1$.

c) We design a steady-state Kalman filter. This filter yields the same performance as the standard Kalman filter in steady state.

The steady-state Kalman filter for the given problem is given by

$$\hat{x}(k) = \big(I - KH\big)\hat{x}(k-1) + Kz(k),$$

where K is defined as

$$K := P_\infty H^T\big(HP_\infty H^T + 2I\big)^{-1},$$

and P_∞ is the symmetric, positive semi-definite solution to the discrete algebraic Riccati equation (DARE)

$$P_\infty = P_\infty - P_\infty H^T\big(HP_\infty H^T + 2I\big)^{-1}HP_\infty + I.$$

Note that the DARE has a unique solution $P_\infty \geq 0$ since (I, H) is detectable (in fact, observable) and (I, I) is stabilizable (in fact, controllable).

Problem 5 **20%**

Consider the scalar system

$$\begin{aligned} x(k) &= x(k-1) + v(k) \\ z(k) &= x(k) + w(k)\,, \qquad\qquad k = 1,\, 2,\, \ldots\,, \end{aligned}$$

where $v(k), w(k) \in [-1, 1]$ are uniformly distributed random variables. The initial state $x(0)$ is uniformly distributed on the interval $[-1, 1]$. All random variables $v(k)$, $w(k)$, and $x(0)$ are assumed to be mutually independent and independent over time.

Suppose the first measurement is $z(1) = 1$.

a) Use Bayesian tracking to find $f_{x(1)|z(1)}(x(1)|z(1))$, that is, the probability distribution of $x(1)$ given $z(1)$.

b) Calculate the Kalman filter estimate $\hat{x}(1|1)$, that is, the Kalman filter estimate of $x(1)$ given $z(1)$.

c) How is $\hat{x}(1|1)$ related to $f_{x(1)|z(1)}(x(1)|z(1))$ for the given problem?

Solution 5

a) Bayesian tracking proceeds in two steps:

(i) a prior update of the probability density function of $x(k)$ based on the process equation, given all measurements up to and including time $(k-1)$, and

(ii) the measurement update of the probability density function of $x(k)$ given the measurement $z(k)$ at time k.

For the given problem, that is:

(i) For $k = 1$, the prior update is based on the initial distribution of $x(0)$,

$$\begin{aligned} f\left(x(1)\right) &= \int_{x(0)} f\left(x(1)|x(0)\right) f\left(x(0)\right)\, dx(0) \\ &= \frac{1}{2}\int_{-1}^{1} f\left(x(1)|x(0)\right)\, dx(0)\,. \end{aligned}$$

Note that $x(k)$ is a continuous random variable and, thus, the formulas taught in class (that consider a discrete random variable $x(k)$) have to be adapted accordingly. Basically, the sums are replaced by integrals.
Recalling the process equation, we get

$$f\left(x(1)|x(0)\right) = \begin{cases} 1/2 & \text{for } \ x(1)-1 \le x(0) \le x(1)+1 \\ 0 & \text{otherwise.} \end{cases}$$

Therefore,

$$f\left(x(1)\right) = \begin{cases} 0 & \text{for } \ x(1) < -2 \ \text{ and } \ x(1) \ge 2 \\ 1/4\left(x(1)+2\right) & \text{for } \ -2 \le x(1) < 0 \\ -1/4\left(x(1)-2\right) & \text{for } \ 0 \le x(1) < 2\,, \end{cases}$$

which is a triangular distribution as shown in Fig. 6.

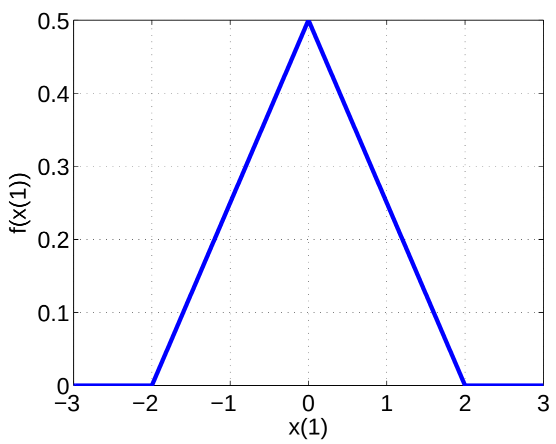

Figure 6: The prior probability distribution of $x(1)$.

(ii) Given $z(1) = 1$, the measurement update can be calculated as

$$f\left(x(1)|z(1)=1\right) = \frac{f\left(z(1)=1|x(1)\right) f\left(x(1)\right)}{c},$$

where c is a normalization constant and

$$f\left(z(1)=1|x(1)\right) = \begin{cases} 1/2 & \text{for } \; 0 \leq x(1) \leq 2 \\ 0 & \text{otherwise}\,. \end{cases}$$

Then,

$$f\left(x(1)|z(1)=1\right) = \begin{cases} 0 & \text{for } x(1) < 0 \text{ and } x(1) > 2 \\ -1/(8c)\,(x(1)-2) & \text{otherwise}\,. \end{cases}$$

The normalization constant is computed as

$$c = -\frac{1}{8}\int_0^2 \left(x(1)-2\right)\, dx(1) = \frac{1}{4}.$$

Fig. 7 shows $f\left(x(1)|z(1)=1\right)$.

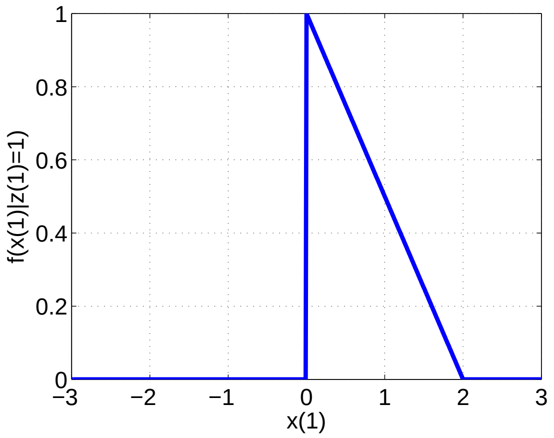

Figure 7: The updated probability distribution of $x(1)$ given $z(1) = 1$.

b) From the probability density functions of $x(0)$, $v(k)$, and $w(k)$, we obtain the statistical properties needed for computing the Kalman filter estimate, namely,

$$\hat{x}(0|0) = \mathrm{E}[x(0)] = 0$$
$$P(0|0) = \mathrm{Var}\left(x(0)\right) = \frac{1}{3}$$

and, similarly,

$$Q(k) = \mathrm{Var}\left(v(k)\right) = \frac{1}{3}$$
$$R(k) = \mathrm{Var}\left(w(k)\right) = \frac{1}{3}.$$

The system is given by $A = 1$, $B = 0$, and $H = 1$. The Kalman filter equations give

$$\hat{x}(1|0) = A\,\hat{x}(0|0) = 0$$
$$P(1|0) = AP(0|0)A^T + Q(1) = 2/3$$
$$K(1) = P(1|0)H^T\left(HP(1|0)H^T + R(1)\right)^{-1} = 2/3$$

and, finally,

$$\hat{x}(1|1) = \hat{x}(1|0) + K(1)\left(z(1) - H\hat{x}(1|0)\right) = 2/3.$$

c) Calculating the expected value of $(x(1)|z(1))$ with $f(x(1)|z(1))$ from part a) gives

$$\mathrm{E}\left[x(1)|z(1)\right] = -\frac{1}{2}\int_0^2 \left(x(1)-2\right)x(1)\,dx(1) = 2/3.$$

That is, for this problem, the Kalman filter estimate $\hat{x}(1|1)$ is the expected value of $(x(1)|z(1))$ as determined from the probability density function that was found in part a).