ETH
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 10, 2013**

Recursive Estimation (151-0566-00) **Dr. Sebastian Trimpe**

Example Solutions

Exam Duration: **150 minutes**

Number of Problems: **6**

Total Points: **125**

Permitted aids: Two one-sided A4 pages.

Use only the provided sheets for your solutions.

Problem 1 **20 Points**

Consider the joint probability density function (PDF)

$$f(x, y) = x + 2y + 2xy$$

for the continuous random variables x and y, where

$$x \in \mathcal{X} = [0, 1], \quad \text{and} \quad y \in \mathcal{Y} = [0, a].$$

a) Determine $a > 0$ such that $f(x, y)$ is a valid PDF.

b) Compute the probability that $y \leq 1/3$ given $x = 1/2$, i.e. calculate $\Pr(y \leq 1/3 | x = 1/2)$.

c) Are x and y independent? Provide reasons for your answer.

d) Compute the minimum mean squared error estimate of x, defined by

$$\hat{x}^{\text{MMSE}} := \arg\min_{\hat{x} \in \mathcal{X}} \operatorname*{E}_{x}[(x - \hat{x})^2].$$

Example Solution 1

a) In order for $f(x,y)$ to be a valid PDF, we require that

$$\int_{\mathcal{X}} \int_{\mathcal{Y}} f(x,y)\ dy\ dx = 1.$$

We find

$$\int_{\mathcal{Y}} f(x,y)\ dy = \int_0^a x + 2y + 2xy\ dy = \left[xy + y^2 + xy^2\right]_0^a = ax + a^2 + a^2x$$

and

$$\int_0^1 ax + a^2 + a^2x\ dx = \left[\frac{ax^2}{2} + a^2x + \frac{a^2x^2}{2}\right]_0^1 = \frac{3}{2}a^2 + \frac{1}{2}a.$$

Solving

$$\frac{3}{2}a^2 + \frac{1}{2}a = 1$$

for a, we find the two solutions

$$a^+ = -\frac{1}{6} + \frac{5}{6} \quad \text{and} \quad a^- = -\frac{1}{6} - \frac{5}{6}.$$

Since $a > 0$, we find $a = a^+ = 2/3$. Note that another requirement for a valid PDF is $f(x,y) \geq 0$ for all $x \in \mathcal{X},\ y \in \mathcal{Y}$. This holds for the given PDF and the sets $\mathcal{X}$ and $\mathcal{Y}$.

b) We compute the probability $\Pr(y \leq 1/3 | x = 1/2)$ with

$$\Pr(y \leq 1/3 | x = 1/2) = \int_{-\infty}^{\frac{1}{3}} f(y|x = 1/2)\ dy.$$

The conditional PDF is

$$f(y|x = 1/2) = \frac{f(x = 1/2, y)}{f(x = 1/2)}.$$

The marginal PDF of x is

$$f(x) = \int_0^{\frac{2}{3}} f(x,y)\ dy = \left[xy + y^2 + xy^2\right]_0^{\frac{2}{3}} = \frac{1}{9}(10x + 4).$$

For the given value of x, $f(x = 1/2) = 1$. Therefore,

$$\int_{-\infty}^{\frac{1}{3}} f(y|x = 1/2)\ dy = \int_0^{\frac{1}{3}} \frac{\frac{1}{2} + 2y + y}{1}\ dy = \left[\frac{1}{2}y + \frac{3}{2}y^2\right]_0^{\frac{1}{3}} = \frac{1}{3} = \Pr(y \leq 1/3 | x = 1/2).$$

c) The random variables are not independent since $f(x,y) = f(x)f(y)$ does not hold for all $x \in \mathcal{X},\ y \in \mathcal{Y}$. We show this by finding a counter-example. First, we find

$$f(y) = \int_0^1 f(x,y)\ dx = \left[\frac{1}{2}x^2 + 2xy + yx^2\right]_0^1 = 3y + \frac{1}{2}.$$

Then, for $x = 0$ and $y = 1/2$ we find

$$f(x = 0) \cdot f(y = 1/2) = \frac{4}{9} \cdot 2 \neq f(x = 0, y = 1/2) = 1.$$

d) In the following, we show that the MMSE estimate is equal to the expected value: $\hat{x}^{\text{MMSE}} = \text{E}[x]$. First, we expand the function to be minimized:

$$\text{E}[(x - \hat{x})^2] = \text{E}[x^2 - 2x\hat{x} + \hat{x}^2] = \text{E}[x^2] - 2\hat{x}\text{E}[x] + \hat{x}^2.$$

The derivative with respect to the scalar estimate $\hat{x}$ is then

$$\frac{d}{d\hat{x}} \left(\text{E}[x^2] - 2\hat{x}\text{E}[x] + \hat{x}^2 \right) = -2\text{E}[x] + 2\hat{x}.$$

Solving for the $\hat{x}$ where the derivative is zero

$$0 = -2\text{E}[x] + 2\hat{x}$$

yields the result

$$\hat{x} = \text{E}[x].$$

The second derivative

$$\frac{d^2}{d\hat{x}^2} \left(\text{E}[x^2] - 2\hat{x}\text{E}[x] + \hat{x}^2 \right) = 2$$

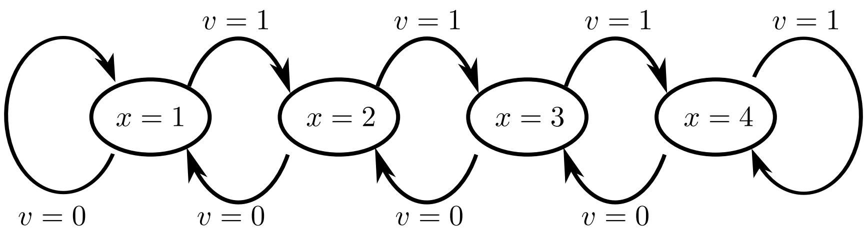
is positive, which shows that $\text{E}[x]$ is a minimum.

Therefore, we find the solution:

$$\hat{x}^{\text{MMSE}} = \text{E}[x] = \int_0^1 x f(x)\, dx = \frac{1}{9} \int_0^1 10x^2 + 4x\, dx = \frac{1}{9} \left[\frac{10}{3} x^3 + 2x^2 \right]_0^1 = \frac{1}{9} \left(\frac{10}{3} + 2 \right) = \frac{16}{27}.$$

Problem 2 **20 Points**

A state machine consists of four discrete states $x(k) \in \{1, 2, 3, 4\}$, where k is the discrete-time index. The discrete-time state dynamics of the state machine are given by the following diagram:



The diagram describes how the state $x(k)$ evolves depending on the previous state $x(k-1)$ and the noise signal $v(k-1)$, where $v(k-1)$ can take the values 0 and 1. For clarity, the time indices are omitted in the diagram. The arrows indicate the transitions from $x(k-1)$ to $x(k)$ depending on the noise signal: For example, if $x(k-1) = 2$ and $v(k-1) = 1$, then $x(k) = 3$, and if $x(k-1) = 1$ and $v(k-1) = 0$, then $x(k) = 1$. The noise $v(k-1)$ takes the value 1 with probability 0.3 for all k, i.e. $f_{v(k-1)}(1) = 0.3$ for all k.

At each time step, the measurement $z(k)$ provides information about whether the state is in the left half of the state machine ($x(k) = 1$ or $x(k) = 2$) or the right half of the state machine ($x(k) = 3$ or $x(k) = 4$). The measurement can take the values 0 and 1, and the probability of measuring 1 given the state $x(k)$ is given by

$$f_{z(k)|x(k)}(1|x(k)) = \begin{cases} 0.9 & \text{if } x(k) = 3 \text{ or } x(k) = 4, \\ 0.2 & \text{otherwise.} \end{cases}$$

a) *Prior update:* Assume that all values of $x(0)$ are equally likely. Compute the probability density function (PDF) $f(x(1))$.

b) *Posterior update:* The conditional PDF $f(x(5)|z(1:4))$ is given as follows:

i	1	2	3	4
$f_{x(5)\|z(1:4)}(i\|z(1:4))$	0.1	0.3	0.4	0.2

You now obtain the measurement $z(5) = 0$. Compute the conditional PDF $f(x(5)|z(1:5))$.

Example Solution 2

a) The problem states that at $k = 0$, all states are equally likely, i.e. $f(x(0)) = 1/4$ for all $x(0)$. The conditional PDF $f(x(1)|x(0))$ can be derived from the diagram and $f(v(k-1))$ to be the following:

$x(0)$ \ $x(1)$	1	2	3	4
1	0.7	0.3	0	0
2	0.7	0	0.3	0
3	0	0.7	0	0.3
4	0	0	0.7	0.3

By the total probability theorem, we can then compute the PDF

$$f(x(1)) = \sum_{i=1}^{4} f(x(1)|x(0) = i) f(x(0) = i)$$
$$f(x(1)) = \frac{1}{4} \sum_{i=1}^{4} f(x(1)|x(0) = i)$$

which results in the following values:

i	1	2	3	4
$f(x(1) = i)$	$\frac{7}{20}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{3}{20}$

b) By Bayes' rule,

$$f(x(5)|z(1\!:\!5)) = \frac{f(z(5)|x(5), z(1\!:\!4)) f(x(5)|z(1\!:\!4))}{f(z(5)|z(1\!:\!4))}\,. \tag{1}$$

Note that $f(z(5)|x(5), z(1 : 4)) = f(z(5)|x(5))$ because the current observation depends only on the current state. Furthermore, since the measurement only takes the values 0 and 1

$$f(z(5) = 0|x(5)) = 1 - f(z(5) = 1|x(5)) = \begin{cases} 0.8 & \text{if } x(5) = 1 \text{ or } x(5) = 2 \\ 0.1 & \text{if } x(5) = 3 \text{ or } x(5) = 4\,. \end{cases}$$

Using the given distribution $f(x(5)|z(1\!:\!4))$, the numerator in (1) can be evaluated to

$$f(z(5) = 0|x(5)) f(x(5)|z(1\!:\!4)) = \begin{cases} 0.08 & \text{for } x(5) = 1 \\ 0.24 & \text{for } x(5) = 2 \\ 0.04 & \text{for } x(5) = 3 \\ 0.02 & \text{for } x(5) = 4 \end{cases}$$

and the denominator is computed using the Total Probability Theorem:

$$\begin{aligned} f(z(5) = 0|z(1\!:\!4)) &= \sum_{x(5)=1}^{4} f(z(5) = 0|x(5)) f(x(5)|z(1\!:\!4)) \\ &= 0.08 + 0.24 + 0.04 + 0.02 \\ &= 0.38\,. \end{aligned}$$

The resulting PDF is

$$f(x(5)|z(1\!:\!5)) = \begin{cases} \frac{4}{19} & \text{for } x(5) = 1 \\ \frac{12}{19} & \text{for } x(5) = 2 \\ \frac{2}{19} & \text{for } x(5) = 3 \\ \frac{1}{19} & \text{for } x(5) = 4\,. \end{cases}$$

Problem 3 — 20 Points

Consider the following discrete-time system:

$$x(k) = \underbrace{\begin{bmatrix} 3 & 1 \\ 0 & 0.5 \end{bmatrix}}_{A} x(k-1) + \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{B} u(k-1) + v(k-1), \quad v(k-1) \sim \mathcal{N}\Big(0, \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}}_{Q}\Big) \tag{2}$$

$$z(k) = \underbrace{\begin{bmatrix} 0.5 & 0.5 \end{bmatrix}}_{H} x(k) + w(k), \quad w(k) \sim \mathcal{N}\big(0, \underbrace{0.5}_{R}\big) \tag{3}$$

where $\mathcal{N}(0, \Gamma)$ denotes a zero-mean Gaussian distribution with variance Γ, and $k = 1, 2, \ldots$ is the discrete-time index.

a) Does the Discrete Algebraic Riccati Equation (DARE)

$$P_\infty = AP_\infty A^T + Q - AP_\infty H^T \left(HP_\infty H^T + R\right)^{-1} HP_\infty A^T$$

have a unique positive semidefinite solution P_∞ for the given system? Justify your answer.

b) Consider the estimation error $e(k) := x(k) - \hat{x}(k)$ of the steady-state Kalman Filter given by

$$\hat{x}(k) = \big(I - K_\infty H\big) A\, \hat{x}(k-1) + \big(I - K_\infty H\big) B\, u(k-1) + K_\infty\, z(k) \tag{4}$$

with

$$K_\infty = P_\infty H^T (HP_\infty H^T + R)^{-1}.$$

Are the error dynamics stable, i.e. is $\lim_{k\to\infty} \mathrm{E}[e(k)] = 0$ for arbitrary $e(0)$? Justify your answer.

c) Now consider the feedback law

$$u(k-1) = \begin{bmatrix} -1 & 0 \end{bmatrix} \hat{x}(k-1). \tag{5}$$

Is the closed-loop system given by the system (2),(3); the estimator (4); and the feedback law (5) stable; i.e. given $\xi(k) := (x(k), e(k))$, is $\lim_{k\to\infty} \mathrm{E}[\xi(k)] = 0$ for arbitrary $\xi(0)$? Justify your answer.

d) Let the initial state $x(0)$ have a Gaussian distribution with mean x_0 and variance P_0: $x(0) \sim \mathcal{N}(x_0, P_0)$. You initialize the steady-state Kalman Filter given by (4) with the given mean x_0. After the first update with $u(0)$ and $z(1)$, the steady-state Kalman Filter state estimate is $\hat{x}(1) = (1.2087,\ 0.5)$.

You have also implemented a time-varying Kalman Filter for the system (2),(3). You initialize the time-varying Kalman Filter with the given mean x_0 and variance P_0. Using the same $u(0)$ and $z(1)$ as with the steady-state Kalman Filter, the first posterior state estimate of the time-varying Kalman Filter is $\hat{x}_m(1) = (1.4727,\ 0.1)$.

Is one of the two estimates equal to the true conditional expected value $\mathrm{E}\left[x(1)|z(1)\right]$? If this is not the case, which of the two is closer to $\mathrm{E}\left[x(1)|z(1)\right]$ (measured by the Euclidean norm)? Justify your answer.

Example Solution 3

a) A unique positive semidefinite solution P_∞ exists if and only if (A, H) is detectable and (A, G) is stabilizable, where G is any matrix such that $Q = GG^T$.

We begin by checking the pair (A, H) for observability, a sufficient condition for detectability. The observability matrix is

$$\begin{bmatrix} H \\ HA \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \\ 1.5 & 0.75 \end{bmatrix}$$

which is full rank. The pair (A, H) is therefore observable and thereby detectable.

Because the process noise variance matrix is positive definite, (A, G) is stabilizable and the two conditions for a unique positive-semidefinite solution to the Discrete Algebraic Riccati Equation (DARE) are satisfied. A unique positive-semidefinite steady-state prior variance P_∞ therefore exists.

b) Yes, the error dynamics are stable. This is guaranteed by the existence of a unique positive-semidefinite solution to the DARE, which was shown in part a).

c) Due to the separation principle, the closed-loop poles are given by the estimator and controller poles. The estimator is stable, as shown above, therefore its poles have magnitude smaller than one. The feedback law results in the closed-loop state dynamics

$$\begin{aligned} A + BF &= \begin{bmatrix} 3 & 1 \\ 0 & 0.5 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 2 & 1 \\ 0 & 0.5 \end{bmatrix}. \end{aligned}$$

The poles are at 2 and 0.5. The pole at 2 has a magnitude greater than one, hence the closed-loop state dynamics are unstable. The closed-loop system is therefore unstable.

d) The time-varying Kalman Filter is the exact solution to the Bayesian tracking problem for linear systems with Gaussian distributions. Because this is the case for the given problem data, $\mathrm{E}\left[x(1)|z(1)\right] = \hat{x}_m(1) = (1.4727,\ 0.1)$.

Problem 4 **30 Points**

Recall the basic particle filtering algorithm we derived in class:

0. *Initialization*: Draw N samples from the probability density function (PDF) of the initial state, $f_{x(0)}(x(0))$.
   Obtain $x_m^n(0),\ n = 1, 2, \ldots, N$. Set $k = 1$.
1. *Step 1*: Simulate the N particles via the process model.
   Obtain the prior particles $x_p^n(k)$.
2. *Step 2*: Given the measurement $z(k)$, scale each prior particle by its measurement likelihood, and obtain a corresponding normalized weight β_n for each particle.
   Resample to get the posterior particles $x_m^n(k)$ that have equal weights.
   Set k to $(k + 1)$. Go to Step 1.

Consider the discrete-time process and measurement model

$$\begin{aligned} x(k) &= -2u(k-1)x^2(k-1) + v(k-1) \\ z_1(k) &= 2x(k) + w_1(k) \\ z_2(k) &= 3x(k) + 2w_2(k) \end{aligned}$$

where $k = 1, 2, \ldots$ is the discrete-time index; $x(k)$ is the scalar system state; $u(k-1)$ is a known control input; $v(k-1)$ is process noise with the PDF

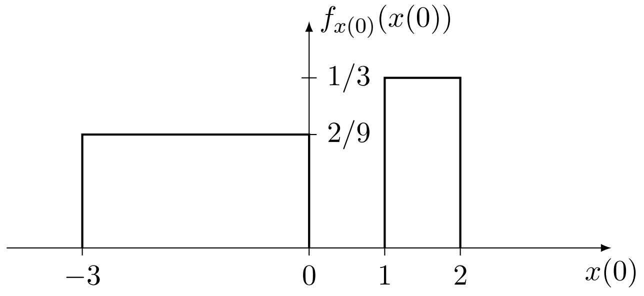
$$f_{v(k)}(v(k)) = \begin{cases} av^2(k) & v(k) \in [-2, 2] \\ 0 & \text{otherwise,} \end{cases} \qquad \text{for all } k$$

with normalization constant a; $z_1(k)$ and $z_2(k)$ are measurements; and $w_1(k)$ and $w_2(k)$ are measurement noise with the triangular PDFs

$$f_{w_1(k)}(w_1(k)) = \begin{cases} \frac{1}{3} + \frac{1}{9}w_1(k) & w_1(k) \in [-3, 0) \\ \frac{1}{3} - \frac{1}{9}w_1(k) & w_1(k) \in [0, 3] \\ 0 & \text{otherwise,} \end{cases} \qquad \text{for all } k$$

$$f_{w_2(k)}(w_2(k)) = \begin{cases} \frac{1}{2} + \frac{1}{4}w_2(k) & w_2(k) \in [-2, 0) \\ \frac{1}{2} - \frac{1}{4}w_2(k) & w_2(k) \in [0, 2] \\ 0 & \text{otherwise,} \end{cases} \qquad \text{for all } k.$$

The distribution of the initial state $x(0)$ is given by the PDF



The random variables $\{v(\cdot)\}$, $\{w_1(\cdot)\}$, $\{w_2(\cdot)\}$, and $x(0)$ are mutually independent.

a) *Initialization*: Initialize a particle filter with $N = 2$ particles. For this calculation, you have access to a random number generator that generates independent random samples r from a uniform distribution on the interval $[0, 1]$. For the calculation of $x_m^1(0)$, you obtained the sample $r_1 = 11/12$. For the calculation of $x_m^2(0)$, you obtained the sample $r_2 = 1/4$.

b) *Step 1*: At time $k = 4$, the posterior particles are $x_m^1(4) = 1/2$ and $x_m^2(4) = -1$. The control input is $u(4) = 2$. Calculate the prior particles $x_p^1(5)$ and $x_p^2(5)$. You have access to the same random number generator as in part a). For the calculation of $x_p^1(5)$, you obtained the sample $r_1 = 7/16$. For the calculation of $x_p^2(5)$, you obtained the sample $r_2 = 9/16$.

c) *Step 2*: At time $k = 7$, the prior particles are $x_p^1(7) = -1/3$ and $x_p^2(7) = -1$. The measurements are $z_1(7) = -1$ and $z_2(7) = -3$.

1. For the two particles, calculate the measurement likelihoods

$$f_{z_1(7),z_2(7)|x(7)}(-1, -3|x_p^1(7)) \quad \text{and} \quad f_{z_1(7),z_2(7)|x(7)}(-1, -3|x_p^2(7)).$$

2. Calculate the normalized particle weights β_1 for $x_p^1(7)$ and β_2 for $x_p^2(7)$.
3. Calculate the *a posteriori* particles $x_m^1(7)$ and $x_m^2(7)$. Do not apply any roughening. You have access to the same random number generator as in part a). For the calculation of $x_m^1(7)$, you obtained the sample $r_1 = 2/3$. For the calculation of $x_m^2(7)$, you obtained the sample $r_2 = 1/3$.

Example Solution 4

a) We sample the given state distribution $f_{x(0)}(x(0))$ using the algorithm discussed in class and the given random numbers. The cumulative distribution function (CDF) of $x(0)$ is

$$F_{x(0)}(\xi) = \int_{-\infty}^{\xi} f_{x(0)}(x(0))\ dx(0) = \begin{cases} 0 & \xi < -3 \\ \frac{2}{9}(\xi+3) & \xi \in [-3,0) \\ \frac{2}{3} & \xi \in [0,1) \\ \frac{2}{3}+\frac{1}{3}(\xi-1) & \xi \in [1,2) \\ 1 & \xi \geq 2. \end{cases}$$

We now solve $F_{x(0)}(x_m^1(0)) = r_1$ for $x_m^1(0)$. Since $r_1 = 11/12 > F_{x(0)}(1) = 2/3$ and $r_1 = 11/12 < 1$, we find that $x_m^1(0) \in [1,2]$. Therefore, we solve

$$\frac{11}{12} = \frac{2}{3} + \frac{1}{3}\left(x_m^1(0) - 1\right)$$

for $x_m^1(0)$ and find $x_m^1(0) = 7/4$. Analogously, we obtain $x_m^2(0) = -15/8$ from $r_2 = 1/4$.

b) We use the process model to simulate the two particles with random samples from the process noise distribution $f_{v(k)}(v(k))$. First, we calculate the normalization constant a of the process noise distribution by making sure that the integral of the PDF evaluates to one:

$$1 = \int_{-\infty}^{\infty} f_{v(k)}(v(k))\ dv(k) = \int_{-2}^{2} av^2(k)\ dv(k) = \left[\frac{av^3(k)}{3}\right]_{-2}^{2} = a\frac{16}{3}.$$

It follows that $a = 3/16$. We can then proceed to sampling $f_{v(k)}(v(k))$ analogously to Part a). The CDF of $v(k)$ follows straightforwardly from the above calculation of the constant a and is

$$F_{v(k)}(\eta) = \begin{cases} 0 & \eta < -2 \\ \frac{1}{2}+\frac{\eta^3}{16} & \eta \in [-2,2] \\ 1 & \eta > 2. \end{cases}$$

Solving

$$r_1 = F_{v(k)}(v^1(4)) \quad \Leftrightarrow \quad \frac{7}{16} = \frac{1}{2} + \frac{(v^1(4))^3}{16}$$

for $v^1(4)$, we find $v^1(4) = -1$. Analogously, we obtain $v^2(4) = 1$ from $r_2 = 9/16$. Then, we apply the process model with $u(4) = 2$ and find

$$x_p^1(5) = -2u(4)(x_m^1(4))^2 + v^1(4) = -2\cdot 2\cdot\left(\frac{1}{2}\right)^2 - 1 = -2$$
$$x_p^2(5) = -2u(4)(x_m^2(4))^2 + v^2(4) = -2\cdot 2\cdot(-1)^2 + 1 = -3.$$

c) 1. We first calculate the measurement likelihood for $z_1(7) = -1$ and $x_p^1(7) = -1/3$. By change of variables, we find

$$f_{z_1(7)|x(7)}(z_1(7)|x_p(7)) = f_{w_1(7)}\left(z_1(7) - 2x_p(7)\right).$$

Therefore,

$$f_{z_1(7)|x(7)}(-1|x_p^1(7) = -1/3) = f_{w_1(7)}\left(-1 + 2\cdot 1/3\right) = f_{w_1(7)}(-1/3) = \frac{1}{3} - \frac{1}{9}\cdot\frac{1}{3} = \frac{8}{27}.$$

Analogously for the second particle $x_p^2(7) = -1$, we find

$$f_{z_1(7)|x(7)}(-1|x_p^2(7)) = \frac{2}{9}.$$

Second, we calculate the measurement likelihood for $z_2(7) = -3$ and $x_p^1(7) = -1/3$. In order to apply the change of variables formula, we first define

$$\begin{aligned} z_2(7) &= g(w_2(7), x_p(7)) := 3x_p(7) + 2w_2(7) \\ w_2(7) &= h(z_2(7), x_p(7)) := \frac{1}{2}(z_2(7) - 3x_p(7)). \end{aligned}$$

Then, we adapt the change of variables formula to conditional PDFs and find

$$f_{z_2(7)|x(7)}(z_2(7)|x_p(7)) = \frac{f_{w_2(7)}\left(h(z_2(7), x_p(7))|x_p(7)\right)}{\left|\frac{dg}{dw_2(7)}\left(h(z_2(7), x_p(7)), x_p(7)\right)\right|}.$$

The derivative is straightforward:

$$\frac{dg}{dw_2(7)}\left(h(z_2(7), x_p(7)), x_p(7)\right) = 2, \quad \text{for all } z_2(7), x_p(7).$$

Therefore, we find

$$f_{z_2(7)|x(7)}(z_2(7)|x_p(7)) = \frac{f_{w_2(7)}(h(z_2(7), x_p(7))|x_p(7))}{\left|\frac{dg}{dw_2(7)}\left(h(z_2(7), x_p(7)), x_p(7)\right)\right|} = \frac{1}{2} f_{w_2(7)}\left(\frac{1}{2}(z_2(7) - 3x_p(7))\right)$$

where we used the fact that $w_2(7)$ and $x(7)$ are independent. Finally, we obtain

$$\begin{aligned} f_{z_2(7)|x(7)}(-3|x_p^1(7) = -1/3) &= \frac{1}{2} f_{w_2(7)}\left(\frac{1}{2}\left(-3 + 3 \cdot \frac{1}{3}\right)\right) \\ &= \frac{1}{2} f_{w_2(7)}(-1) = \frac{1}{2}\left(\frac{1}{2} - \frac{1}{4}\right) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}. \end{aligned}$$

Analogously, for the second particle $x_p^2(7) = -1$, we find

$$f_{z_2(7)|x(7)}(-3|x_p^2(7)) = \frac{1}{4}.$$

From the above change of variables and due to the independence of $w_1(7)$ and $w_2(7)$, it follows that $z_1(7)$ and $z_2(7)$ are conditionally independent when conditioned on $x(7)$. The calculation of the two measurement likelihoods is then straightforward:

$$\begin{aligned} f_{z_1(7),z_2(7)|x(7)}(-1, -3|x_p^1(7)) &= f_{z_1(7)|x(7)}(-1|x_p^1(7)) f_{z_2(7)|x(7)}(-3|x_p^1(7)) = \frac{8}{27} \cdot \frac{1}{8} = \frac{1}{27} \\ f_{z_1(7),z_2(7)|x(7)}(-1, -3|x_p^2(7)) &= f_{z_1(7)|x(7)}(-1|x_p^2(7)) f_{z_2(7)|x(7)}(-3|x_p^2(7)) = \frac{2}{9} \cdot \frac{1}{4} = \frac{1}{18}. \end{aligned}$$

2. The sum of the measurement likelihoods is $1/27 + 1/18 = 5/54 =: \alpha^{-1}$. Finally, we obtain the particle weights

$$\begin{aligned} \beta_1 &= \alpha f_{z_1(7),z_2(7)|x(7)}(-1, -3|x_p^1(7)) = \frac{54}{5} \cdot \frac{1}{27} = \frac{2}{5} \\ \beta_2 &= \alpha f_{z_1(7),z_2(7)|x(7)}(-1, -3|x_p^2(7)) = \frac{54}{5} \cdot \frac{1}{18} = \frac{3}{5}. \end{aligned}$$

3. We resample the prior particles based on the particle weights β_1, β_2 using the procedure presented in class. Since $\beta_1 < r_1 = 2/3$, the second particle $x_p^2(7)$ is sampled:

$$x_m^1(7) = x_p^2(7) = -1.$$

Analogously, since $\beta_1 > r_2 = 1/3$, the first particle $x_p^1(7)$ is sampled:

$$x_m^2(7) = x_p^1(7) = -\frac{1}{3}.$$

Problem 5 **15 Points**

Consider the measurement vector $z(k)$ at time k of a system with state vector $x(k)$, where the (known) input vector $u(k-1)$ also appears linearly in the measurement:

$$z(k) = Hx(k) + Gu(k-1) + w(k).$$

The measurement noise $w(k)$ is normally distributed with zero mean and variance R: $w(k) \sim \mathcal{N}(0, R)$. Furthermore, the measurement noise $w(k)$ is independent of the system state $x(k)$.

In this problem, you will adapt the Kalman Filter measurement update equations to accommodate the input term $Gu(k-1)$.

a) Determine $f_{z(k)|x(k)}(z(k)|x(k))$ as a function of $f_{w(k)}(\cdot)$.

b) Compute the conditional mean $\mathrm{E}[x(k)|z(k)]$ and variance $\mathrm{Var}[x(k)|z(k)]$ given that $x(k) \sim \mathcal{N}(\mu, \Sigma)$.

Hint: One can show that $f_{x(k)|z(k)}(x(k)|z(k))$ is a normal distribution, which you may assume for the solution of this problem.

Example Solution 5

a) Because $Gu(k-1)$ is a deterministic signal, it is straightforward to apply the change of variables and find

$$f_{z(k)|x(k)}(z(k)|x(k)) = f_{w(k)}\big(z(k) - Gu(k-1) - Hx(k)\big)\,.$$

b) In the following, we use the shorthand notation $f_{x(k)|z(k)}(x(k)|z(k)) = f(x(k)|z(k))$ for clarity purposes. By Bayes' rule,

$$f(x(k)|z(k)) = \frac{f(z(k)|x(k))f(x(k))}{f(z(k))}\,. \tag{6}$$

The distribution of $f(x(k))$ is given to be normal:

$$f(x(k)) \propto \exp\Big(-\frac{1}{2}\big(x(k)-\mu\big)^T\Sigma^{-1}(k)\big(x(k)-\mu\big)\Big)\,.$$

As shown in part **a)**, the PDF of $f(z(k)|x(k))$ is also Gaussian:

$$f(z(k)|x(k)) \propto \exp\Big(-\frac{1}{2}\big(z(k)-Gu(k-1)-Hx(k)\big)^T R^{-1}\big(z(k)-Gu(k-1)-Hx(k)\big)\Big)\,.$$

The PDF $f(z(k))$ does not depend on $x(k)$, and is therefore merely a constant in Equation (6). Therefore

$$\begin{aligned} f(x(k)|z(k)) \propto \exp\Big(-\frac{1}{2}\big((x(k)-\mu)^T\Sigma^{-1}(x(k)-\mu) \\ + \big(z(k)-Gu(k-1)-Hx(k)\big)^T R^{-1}\big(z(k)-Gu(k-1)-Hx(k)\big)\Big)\,. \end{aligned}$$

The mean and variance of $f(x(k)|z(k))$ follow from a comparison of coefficients, using the general form of a GRV for mean and variance:

$$f(x(k)|z(k)) \propto \exp\Big(-\frac{1}{2}\big(x(k)-\hat{x}_m(k)\big)^T P_m^{-1}(k)\big(x(k)-\hat{x}_m(k)\big)\Big)\,.$$

By comparing quadratic terms, it follows that

$$P_m^{-1}(k) = \Sigma^{-1} \;+ H^T R^{-1} H\,,$$

i.e. the variance of $f(x(k)|z(k))$ is $\big(\Sigma^{-1}+H^T R^{-1}H\big)^{-1}$. Comparison of linear terms yields

$$\begin{aligned} P_m^{-1}(k)\hat{x}_m(k) &= \Sigma^{-1}\mu + H^T R^{-1} z(k) - H^T R^{-1} Gu(k-1) \\ \hat{x}_m(k) &= P_m(k)\Big(\Sigma^{-1}\mu + H^T R^{-1}\big(z(k)-Gu(k-1)\big)\Big) \\ &= P_m(k)\Big(\Sigma^{-1}\mu + H^T R^{-1}\big(z(k)-Gu(k-1)+H\mu-H\mu\big)\Big) \\ &= P_m(k)\Big(P_m^{-1}(k)\mu + H^T R^{-1}\big(z(k)-Gu(k-1)-H\mu\big)\Big) \\ \hat{x}_m(k) &= \mu + P_m(k)H^T R^{-1}\big(z(k)-Gu(k-1)-H\mu\big) \end{aligned}$$

i.e. the mean of $f(x(k)|z(k))$ is $\mu + P_m(k)H^T R^{-1}\big(z(k)-Gu(k-1)-H\mu\big)$.

Alternative Solution:

As an alternative approach, we introduce the pseudo-measurement $\tilde{z}(k)$ as follows:

$$\tilde{z}(k) = z(k) - Gu(k-1) = Hx(k) + w(k)\,. \tag{7}$$

Using the pseudo-measurement, the problem description is in the form of the standard Kalman Filter as it was presented in the lecture. The required measurement update step for variance and mean then reads as follows:

$$\begin{aligned} P_m(k) &= \left(\Sigma^{-1} + H^T R^{-1} H\right)^{-1} \\ \hat{x}_m(k) &= \mu + P_m(k) H^T R^{-1} \left(\tilde{z}(k) - H\mu\right)\,. \end{aligned} \tag{8}$$

Substituting the pseudo-measurement (7) into the mean update equation (8) yields

$$\hat{x}_m(k) = \mu + P_m(k) H^T R^{-1} \left(z(k) - Gu(k-1) - H\mu\right)$$

and the distribution of $x(k)|z(k)$ is given by

$$x(k)|z(k) \sim \mathcal{N}(\hat{x}_m(k), P_m(k))\,.$$

Problem 6 **20 Points**

A time-varying Kalman Filter (KF) is used to estimate the state $x(k)$ of the process

$$x(k) = \begin{bmatrix} 0.5 & 0 \\ 0 & 2 \end{bmatrix} x(k-1) + v(k-1), \qquad v(k-1) \sim \mathcal{N}\left(0, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

from measurements $z(1), z(2), \ldots$, given by the measurement equation

$$z(k) = \begin{bmatrix} 0 & 1 \end{bmatrix} x(k) + w(k), \qquad w(k) \sim \mathcal{N}(0, 1)$$

where $k = 1, 2, \ldots$ is the discrete-time index. The initial state distribution is Gaussian with mean x_0 and variance P_0: $x(0) \sim \mathcal{N}(x_0, P_0)$. While the mean x_0 of the initial state distribution is known and used to initialize the KF, the variance P_0 is unknown.

Let $\hat{x}_m(k)$ and $P_m(k)$ denote, respectively, the posterior mean and variance of the KF at time k (that is, the mean and variance of the state $x(k)$, conditioned on the measurements $z(1), z(2), \ldots, z(k)$). The KF mean $\hat{x}_m(k)$ is initialized with the true mean of the initial state distribution: $\hat{x}_m(0) = x_0 = \mathrm{E}[x(0)]$. However, since the variance P_0 is unknown, the KF variance $P_m(k)$ is initialized with an arbitrary symmetric, positive semidefinite matrix $\bar{P}$: $P_m(0) = \bar{P}$.

The estimation error of the KF is defined as $e(k) := x(k) - \hat{x}_m(k)$. You shall analyze the long-term behavior of the estimation error $e(k)$ for $k \to \infty$.

a) What is $\lim_{k\to\infty} \mathrm{E}[e(k)]$? Justify your answer.

b) 1. What is $\lim_{k\to\infty} \mathrm{Var}[e(k)]$? State every entry of the matrix $\lim_{k\to\infty} \mathrm{Var}[e(k)]$.

2. Does the limit $\lim_{k\to\infty} \mathrm{Var}[e(k)]$ depend on $\bar{P}$? Provide a brief reasoning for your answer.

Now consider the system

$$x(k) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} x(k-1) + v(k-1), \qquad v(k-1) \sim \mathcal{N}\left(0, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

with the same measurement equation as above. For this part, you can assume that the KF variance $P_m(k)$ is initialized with an arbitrary *diagonal*, positive semidefinite matrix $\bar{P}$.

c) What is $\lim_{k\to\infty} \mathrm{Var}[e(k)]$? State every entry of the matrix $\lim_{k\to\infty} \mathrm{Var}[e(k)]$.

Example Solution 6

We use the typical symbols for the system matrices in the solution below:

$$A = \begin{bmatrix} 0.5 & 0 \\ 0 & 2 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I, \quad H = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad R = 1. \tag{9}$$

Furthermore, we use $P_p(k)$ to denote the KF prior variance at time k.

a) The KF estimation error obeys the following difference equation (derived in class):

$$e(k) = \big(I - K(k)H\big)A\,e(k-1) + \big(I - K(k)H\big)\,v(k-1) - K(k)\,w(k),$$

where $K(k)$ is the (time-varying) KF gain. Taking the expected value on both sides yields

$$\mathrm{E}[e(k)] = \big(I - K(k)H\big)A\,\mathrm{E}[e(k-1)]. \tag{10}$$

Since we initialize the KF with the mean of the initial state distribution ($\hat{x}_m(0) = x_0$), we have $\mathrm{E}[e(0)] = \mathrm{E}[x(0)] - \hat{x}_m(0) = x_0 - x_0 = 0$, and it follows from (10) that $\mathrm{E}[e(k)] = 0$ for all k, which implies $\lim_{k\to\infty} \mathrm{E}[e(k)] = 0$.

b) 1. We first verify that (A, H) is detectable. To see this, we note that the only unstable mode (corresponding to the eigenvalue 2 of the matrix A) is observable through the output (measurement). More formally, we can use the PBH-Test and check whether the rank condition is satisfied for all unstable eigenvalues of A: for $\lambda = 2$, we find that

$$\begin{bmatrix} A - \lambda I \\ H \end{bmatrix} = \begin{bmatrix} -1.5 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

has full column rank, which implies the detectability of (A, H).

Second, we verify that (A, G) is stabilizable, where G is any matrix such that $Q = GG^T$. Since Q is positive definite, it immediately follows that (A, G) is stabilizable. Using the convergence theorem for the Discrete Algebraic Riccati Equation (DARE)

$$P = APA^T + Q - APH^T\big(HPH^T + R\big)^{-1}HPA^T \tag{11}$$

discussed in class, we can thus conclude that the DARE (11) has a unique positive-semidefinite solution $P \geq 0$. Furthermore, $\lim_{k\to\infty} P_p(k) = P$ for any initial $P_p(1) \geq 0$ (and thus any initial $P_m(0) = \bar{P} \geq 0$).

Next, we compute P. Applying the problem data to (11) yields

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} p_{11}/4 + 1 & p_{12} \\ p_{12} & 4p_{22} + 1 \end{bmatrix} - \frac{1}{p_{22} + 1} \begin{bmatrix} p_{12}^2/4 & p_{12}p_{22} \\ p_{12}p_{22} & 4p_{22}^2 \end{bmatrix}.$$

From this, we obtain three scalar equations for p_{11}, p_{12}, and p_{22}:

$$p_{11} = p_{11}/4 + 1 - \frac{p_{12}^2/4}{p_{22} + 1} \tag{12}$$

$$p_{12} = p_{12} - \frac{p_{12}p_{22}}{p_{22} + 1} \tag{13}$$

$$p_{22} = 4p_{22} + 1 - \frac{4p_{22}^2}{p_{22} + 1}. \tag{14}$$

From equation (14), we find two possible solutions for p_{22}:

$$p_{22}^{-} = 2 - \sqrt{5} \quad \text{and} \quad p_{22}^{+} = 2 + \sqrt{5}. \tag{15}$$

Manipulating equation (13), we find

$$0 = -\frac{p_{12}p_{22}}{p_{22}+1}.$$

Since from (15), it is evident that $p_{22} \neq 0$, the above equation only holds if $p_{12} = 0$. Using this result, we can solve (12) for p_{11} and find $p_{11} = 4/3$. Since P must be positive-semidefinite, it follows that $p_{22} \geq 0$ (all eigenvalues must be nonnegative). Therefore, we find $p_{22} = 2 + \sqrt{5}$. Finally, we have

$$P = \begin{bmatrix} \frac{4}{3} & 0 \\ 0 & 2+\sqrt{5} \end{bmatrix}.$$

From the convergence of $P_p(k)$ for $k \to \infty$, it follows that also $P_m(k) = \text{Var}[e(k)]$ converges for $k \to \infty$. Let $\lim_{k\to\infty} P_m(k) = \tilde{P}$, where

$$\tilde{P} = \begin{bmatrix} \tilde{p}_{11} & \tilde{p}_{12} \\ \tilde{p}_{12} & \tilde{p}_{22} \end{bmatrix}.$$

From the KF prior update equation, we have

$$P = A\tilde{P}A^T + Q$$

$$\begin{bmatrix} 4/3 & 0 \\ 0 & 2+\sqrt{5} \end{bmatrix} = \begin{bmatrix} \tilde{p}_{11}/4+1 & \tilde{p}_{12} \\ \tilde{p}_{12} & 4\tilde{p}_{22}+1 \end{bmatrix}.$$

Solving for $\tilde{p}_{11}$, $\tilde{p}_{12}$, and $\tilde{p}_{22}$, we find

$$\lim_{k\to\infty} \text{Var}[e(k)] = \tilde{P} = \begin{bmatrix} \frac{4}{3} & 0 \\ 0 & \frac{1+\sqrt{5}}{4} \end{bmatrix}.$$

2. From the convergence theorem for the DARE, we know that the limit calculated in 1 is independent of the initial $\bar{P} \geq 0$.

Alternative approach. An alternative approach to this problem is to realize from the problem structure that the system can be decoupled into two separate systems:

$$x_1(k) = 0.5\, x_1(k-1) + v_1(k-1), \qquad v_1(k-1) \sim \mathcal{N}(0,1)$$

and

$$\begin{aligned} x_2(k) &= 2\, x_2(k-1) + v_2(k-1), & v_2(k-1) &\sim \mathcal{N}(0,1) \\ z(k) &= x_2(k) + w(k), & w(k) &\sim \mathcal{N}(0,1). \end{aligned}$$

Notice that rewriting the system as two separate systems is possible without loss of generality because there is no coupling of the dynamics (no off-diagonal entries in A), no cross correlation of the process noise (no off-diagonal entries in Q), and the measurement equation depends only on $x_2(k)$.

One can then apply the KF equations separately to each of the two systems and compute the steady-state variance accordingly.

c) For this part, $A = 2I$ and the other matrices are as in (9). The modified system is no longer detectable because the first element of the state vector is both unobservable and unstable. This can be checked analogously to Part b). Hence, we expect that the KF variance will diverge. We compute the individual entries of $\lim_{k\to\infty} P_m(k)$ next.

The update equation for the KF prior variance $P_p(k)$ reads

$$P_p(k{+}1) = AP(k)A^T + Q - AP(k)H^T\big(HP(k)H^T + R\big)^{-1}HP(k)A^T$$

which results in

$$\begin{bmatrix} p_{11}(k{+}1) & p_{12}(k{+}1) \\ p_{12}(k{+}1) & p_{22}(k{+}1) \end{bmatrix} = \begin{bmatrix} 4p_{11}(k)+1 & 4p_{12}(k) \\ 4p_{12}(k) & 4p_{22}(k)+1 \end{bmatrix} - \frac{1}{p_{22}(k)+1}\begin{bmatrix} 4p_{12}^2(k) & 4p_{12}(k)p_{22}(k) \\ 4p_{12}(k)p_{22}(k) & 4p_{22}^2(k) \end{bmatrix}$$

where $p_{11}(k)$, $p_{12}(k)$, and $p_{22}(k)$ are the entries of $P_p(k)$.

First, we notice that $p_{12}(k) = 0$ implies $p_{12}(k{+}1) = 0$. Furthermore, it is straightforward to show that $p_{12}(1) = 0$ from the assumption that $P_m(0)$ is diagonal and the fact that A and Q are also diagonal. Hence, $p_{12}(k) = 0$ for all k.

With this result, we obtain the following update equations for $p_{11}(k)$ and $p_{22}(k)$:

$$p_{11}(k{+}1) = 4p_{11}(k) + 1 \tag{16}$$

$$p_{22}(k{+}1) = 4p_{22}(k) + 1 - \frac{4p_{22}^2(k)}{p_{22}(k)+1}. \tag{17}$$

Notice that $p_{22}(k)$ obeys the same update equation as in Part b) (compare (17) to (14) augmented with appropriate time indices). Therefore, $\lim_{k\to\infty} p_{22}(k) = 2 + \sqrt{5}$ as in Part b). Furthermore, $P_p(1) \geq 0$ and diagonal implies that $p_{11}(1) \geq 0$, from which, together with (16), follows that $\lim_{k\to\infty} p_{11}(k) = \infty$.

Finally, we find from $P_p(k) = AP_m(k{-}1)A^T + Q = 4P_m(k{-}1) + I$ that

$$\lim_{k\to\infty} \mathrm{Var}[e(k)] = \lim_{k\to\infty} P_m(k) = \lim_{k\to\infty} P_m(k{-}1) = \frac{1}{4}\left(\lim_{k\to\infty} P_p(k) - I\right) = \begin{bmatrix} \infty & 0 \\ 0 & \frac{1+\sqrt{5}}{4} \end{bmatrix}.$$

Alternative approach. Since the structure of the problem is the same as in Part b), one can decouple the system into two separate systems as before:

$$x_1(k) = 2\,x_1(k{-}1) + v_1(k{-}1), \qquad v_1(k{-}1) \sim \mathcal{N}(0,1)$$

and

$$\begin{aligned} x_2(k) &= 2\,x_2(k{-}1) + v_2(k{-}1), & v_2(k{-}1) &\sim \mathcal{N}(0,1) \\ z(k) &= x_2(k) + w(k), & w(k) &\sim \mathcal{N}(0,1). \end{aligned}$$

The second system is the the same as in Part b), hence the (2,2)-entry of $\lim_{k\to\infty} P_m(k)$ is the same as in Part b). The first system is not detectable, hence, the (1,1)-entry of $\lim_{k\to\infty} P_m(k)$ tends to infinity as $k \to \infty$. The coupling terms of $\lim_{k\to\infty} P_m(k)$ are zero, since there is no coupling between the two systems.