A Solution Manual and Notes for:
Kalman Filtering: Theory and Practice using MATLAB
by Mohinder S. Grewal and Angus P. Andrews.

John L. Weatherwax*

April 30, 2012

Introduction

Here you'll find some notes that I wrote up as I worked through this excellent book. There is also quite a complete set of solutions to the various end of chapter problems. I've worked hard to make these notes as good as I can, but I have no illusions that they are perfect. If you feel that that there is a better way to accomplish or explain an exercise or derivation presented in these notes; or that one or more of the explanations is unclear, incomplete, or misleading, please tell me. If you find an error of any kind – technical, grammatical, typographical, whatever – please tell me that, too. I'll gladly add to the acknowledgments in later printings the name of the first person to bring each problem to my attention. I hope you enjoy this book as much as I have and that these notes might help the further development of your skills in Kalman filtering.

Acknowledgments

Special thanks to (most recent comments are listed first): Bobby Motwani and Shantanu Sultan for finding various typos from the text. All comments (no matter how small) are much appreciated. In fact, if you find these notes useful I would appreciate a contribution in the form of a solution to a problem that is not yet worked in these notes. Sort of a "take a penny, leave a penny" type of approach. Remember: pay it forward.

*wax@alum.mit.edu

Chapter 2: Linear Dynamic Systems

Notes On The Text

Notes on Example 2.5

We are told that the fundamental solution $\Phi(t)$ to the differential equation $\frac{d^n y}{dt} = 0$ when written in companion form as the matrix $\frac{d\mathbf{x}}{dt} = F\mathbf{x}$ or in components

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & & & & \\ 0 & 0 & 1 & & & & \\ 0 & 0 & 0 & \ddots & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & 0 & 1 & 0 \\ & & & & 0 & 0 & 1 \\ & & & & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-2} \\ x_{n-1} \\ x_n \end{bmatrix},$$

is

$$\Phi(t) = \begin{bmatrix} 1 & t & \frac{1}{2}t^2 & \frac{1}{3!}t^3 & \cdots & \frac{1}{(n-1)!}t^{n-1} \\ 0 & 1 & t & \frac{1}{2}t^2 & \cdots & \frac{1}{(n-2)!}t^{n-2} \\ 0 & 0 & 1 & t & \cdots & \frac{1}{(n-3)!}t^{n-3} \\ 0 & 0 & 0 & 1 & \cdots & \frac{1}{(n-4)!}t^{n-4} \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Note here the only nonzero values in the matrix F are the ones on its first superdiagonal. We can verify this by showing that the given $\Phi(t)$ satisfies the differential equation and has the correct initial conditions, that is $\frac{\Phi(t)}{dt} = F\Phi(t)$ and $\Phi(0) = I$. That $\Phi(t)$ has the correct initial conditions $\Phi(0) = I$ is easy to see. For the t derivative of $\Phi(t)$ we find

$$\Phi'(t) = \begin{bmatrix} 0 & 1 & t & \frac{1}{2!}t^2 & \cdots & \frac{1}{(n-2)!}t^{n-2} \\ 0 & 0 & 1 & t & \cdots & \frac{1}{(n-3)!}t^{n-3} \\ 0 & 0 & 0 & 1 & \cdots & \frac{1}{(n-4)!}t^{n-4} \\ 0 & 0 & 0 & 0 & \cdots & \frac{1}{(n-5)!}t^{n-5} \\ \vdots & & & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

From the above expressions for $\Phi(t)$ and F by considering the given product $F\Phi(t)$ we see that it is equal to $\Phi'(t)$ derived above as we wanted to show. As a simple modification of the above example consider what the fundamental solution would be if we were given the following companion form for a vector of unknowns $\mathbf{x}$

$$\frac{d}{dt}\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \vdots \\ \hat{x}_{n-2} \\ \hat{x}_{n-1} \\ \hat{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & & & & \\ 1 & 0 & 0 & & & & \\ 0 & 1 & 0 & \ddots & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & 0 & 0 & 0 \\ & & & & 1 & 0 & 0 \\ & & & & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \vdots \\ \hat{x}_{n-2} \\ \hat{x}_{n-1} \\ \hat{x}_n \end{bmatrix} = \hat{F} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \vdots \\ \hat{x}_{n-2} \\ \hat{x}_{n-1} \\ \hat{x}_n \end{bmatrix}.$$

Note in this example the only nonzero values in $\hat{F}$ are the ones on its first subdiagonal. To determine $\Phi(t)$ we note that since this coefficient matrix $\hat{F}$ in this case is the transpose of the first system considered above $\hat{F} = F^T$ the system we are asking to solve is $\frac{d}{dt}\hat{\mathbf{x}} = F^T\hat{\mathbf{x}}$. Thus the fundamental solution to this new problem is

$$\hat{\Phi}(t) = e^{F^T t} = (e^{Ft})^T = \Phi(t)^T ,$$

and that this later matrix looks like

$$\hat{\Phi}(t) = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ t & 1 & 0 & 0 & \cdots & 0 \\ \frac{1}{2}t^2 & t & 1 & 0 & \cdots & 0 \\ \frac{1}{3!}t^3 & \frac{1}{2}t^2 & t & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{(n-1)!}t^{n-1} & \frac{1}{(n-2)!}t^{n-2} & \frac{1}{(n-3)!}t^{n-3} & \frac{1}{(n-4)!}t^{n-4} & \cdots & 1 \end{bmatrix}.$$

Verification of the Solution to the Continuous Linear System

We are told that a solution to the continuous linear system with a time dependent companion matrix $F(t)$ is given by

$$x(t) = \Phi(t)\Phi(t_0)^{-1}x(t_0) + \Phi(t)\int_{t_0}^{t} \Phi^{-1}(\tau)C(\tau)u(\tau)d\tau \,. \qquad (1)$$

To verify this take the derivative of $x(t)$ with respect to time. We find

$$\begin{aligned} x'(t) &= \Phi'(t)\Phi^{-1}(t_0) + \Phi'(t)\int_{t_0}^{t} \Phi^{-1}(\tau)C(\tau)u(\tau)d\tau + \Phi(t)\Phi^{-1}(t)C(t)u(t) \\ &= \Phi'(t)\Phi^{-1}(t)x(t) + C(t)u(t) \\ &= F(t)\Phi(t)\Phi^{-1}(t)x(t) + C(t)u(t) \\ &= F(t)x(t) + C(t)u(t) \,. \end{aligned}$$

showing that the expression given in Equation 1 is indeed a solution. Note that in the above we have used the fact that for a fundamental solution $\Phi(t)$ we have $\Phi'(t) = F(t)\Phi(t)$.

Problem Solutions

Problem 2.2 (the companion matrix for $\frac{d^n y}{dt^n} = 0$)

We begin by defining the following functions $x_i(t)$

$$
\begin{aligned}
x_1(t) &= y(t) \\
x_2(t) &= \dot{x}_1(t) = \dot{y}(t) \\
x_3(t) &= \dot{x}_2(t) = \ddot{x}_1(t) = \ddot{y}(t) \\
&\vdots \\
x_n(t) &= \dot{x}_{n-1}(t) = \cdots = \frac{d^{n-1}y(t)}{dt^{n-1}} ,
\end{aligned}
$$

as the components of a state vector $\mathbf{x}$. Then the companion form for this system is given by

$$
\frac{d}{dt}\mathbf{x}(t) = \frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} = \begin{bmatrix} x_2(t) \\ x_3(t) \\ \vdots \\ x_n(t) \\ \frac{d^n y(t)}{dt^n} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ & & & \ddots & & 1 \\ 0 & & & \ldots & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_{n-1}(t) \\ x_n(t) \end{bmatrix} = Fx(t)
$$

With F the companion matrix given by

$$
F = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ & & & \ddots & & 1 \\ 0 & & & \ldots & 0 & 0 \end{bmatrix} .
$$

Which is of dimensions of $n \times n$.

Problem 2.3 (the companion matrix for $\frac{dy}{dt} = 0$ and $\frac{d^2y}{dt^2} = 0$)

If $n = 1$ the above specifies to the differential equation $\frac{dy}{dt} = 0$ and the companion matrix F is the zero matrix i.e. $F = [0]$. When $n = 2$ we are solving the differential equation given by $\frac{d^2y}{dt^2} = 0$, and a companion matrix F given by

$$
F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} .
$$

Problem 2.4 (the fundamental solution matrix for $\frac{dy}{dt} = 0$ and $\frac{d^2y}{dt^2} = 0$)

The fundamental solution matrix $\Phi(t)$ satisfies

$$
\frac{d\Phi}{dt} = F(t)\Phi(t) ,
$$

with an initial condition $\Phi(0) = I$. When $n = 1$, we have $F = [0]$, so $\frac{d\Phi}{dt} = 0$ giving that $\Phi(t)$ is a constant, say C. To have the initial condition hold $\Phi(0) = 1$, we must have $C = 1$, so that

$$\Phi(t) = 1 \, . \tag{2}$$

When $n = 2$, we have $F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, so that the equation satisfied by Φ is

$$\frac{d\Phi}{dt} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \Phi(t) \, .$$

If we denote the matrix $\Phi(t)$ into its components $\Phi_{ij}(t)$ we have that

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \Phi(t) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} = \begin{bmatrix} \Phi_{21} & \Phi_{22} \\ 0 & 0 \end{bmatrix} ,$$

so the differential equations for the components of Φ_{ij} satisfy

$$\begin{bmatrix} \frac{d\Phi_{11}}{dt} & \frac{d\Phi_{12}}{dt} \\ \frac{d\Phi_{21}}{dt} & \frac{d\Phi_{22}}{dt} \end{bmatrix} = \begin{bmatrix} \Phi_{21} & \Phi_{22} \\ 0 & 0 \end{bmatrix} .$$

Solving the scalar differential equations above for Φ_{21} and Φ_{22} using the known initial conditions for them we have $\Phi_{21} = 0$ and $\Phi_{22} = 1$. With these results the differential equations for Φ_{11} and Φ_{12} become

$$\frac{d\Phi_{11}}{dt} = 0 \quad \text{and} \quad \frac{d\Phi_{12}}{dt} = 1 \, ,$$

so that

$$\Phi_{11} = 1 \quad \text{and} \quad \Phi_{21}(t) = t \, .$$

Thus the fundamental solution matrix $\Phi(t)$ in the case when $n = 2$ is

$$\Phi(t) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} . \tag{3}$$

Problem 2.5 (the state transition matrix for $\frac{dy}{dt} = 0$ and $\frac{d^2y}{dt^2} = 0$)

Given the fundamental solution matrix $\Phi(t)$ for a linear system $\frac{dx}{dt} = F(t)x$ the state transition matrix $\Phi(\tau, t)$ is given by $\Phi(\tau)\Phi(t)^{-1}$. When $n = 1$ since $\Phi(t) = 1$ the state transition matrix in this case is $\Phi(\tau, t) = 1$ also. When $n = 2$ since $\Phi(t) = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$ we have

$$\Phi(t)^{-1} = \begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix} ,$$

so that

$$\Phi(\tau)\Phi(t)^{-1} = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -t \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -t + \tau \\ 0 & 1 \end{bmatrix} .$$

Problem 2.6 (an example in computing the fundamental solution)

We are asked to find the fundamental solution $\Phi(t)$ for the system

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

To find the fundamental solution for the given system we first consider the homogeneous system

$$\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}.$$

To solve this system we need to find the eigenvalues of $\begin{bmatrix} 0 & 0 \\ -1 & -2 \end{bmatrix}$. We solve for λ in the following

$$\begin{vmatrix} -\lambda & 0 \\ -1 & -2-\lambda \end{vmatrix} = 0\,,$$

or $\lambda^2 + 2\lambda = 0$. This equation has roots given by $\lambda = 0$ and $\lambda = -2$. The eigenvector of this matrix for the eigenvalue $\lambda = 0$ is given by solving for the vector with components v_1 and v_2 that satisfies

$$\begin{bmatrix} 0 & 0 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0\,,$$

so $-v_1 - 2v_2 = 0$ so $v_1 = -2v_2$. Which can be made true if we take $v_2 = -1$ and $v_1 = 2$, giving the eigenvector of $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$. When $\lambda = -2$ we have to find the vector $\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ such that

$$\begin{bmatrix} 2 & 0 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 0\,,$$

is satisfied. If we take $v_1 = 0$ and $v_2 = 1$ we find an eigenvector of $v = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Thus with these eigensystem the general solution for $x(t)$ is then given by

$$x(t) = c_1\begin{bmatrix} 2 \\ -1 \end{bmatrix} + c_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}e^{-2t} = \begin{bmatrix} 2 & 0 \\ -1 & e^{-2t} \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}, \tag{4}$$

for two constants c_1 and c_2. The initial condition requires that $x(0)$ be related to c_1 and c_2 by

$$x(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}.$$

Solving for c_1 and c_2 we find

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 1/2 & 1 \end{bmatrix}\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}. \tag{5}$$

Using Equation 4 and 5 $x(t)$ is given by

$$\begin{aligned} x(t) &= \begin{bmatrix} 2 & 0 \\ -1 & e^{-2t} \end{bmatrix}\begin{bmatrix} 1/2 & 0 \\ 1/2 & 1 \end{bmatrix}\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ \frac{1}{2}(-1+e^{-2t}) & e^{-2t} \end{bmatrix}\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}. \end{aligned}$$

From this expression we see that our fundamental solution matrix $\Phi(t)$ for this problem is given by

$$\Phi(t) = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2}(1-e^{-2t}) & e^{-2t} \end{bmatrix} . \tag{6}$$

We can verify this result by checking that this matrix has the required properties that $\Phi(t)$ should have. One property is $\Phi(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, which can be seen true from the above expression. A second property is that $\Phi'(t) = F(t)\Phi(t)$. Taking the derivative of $\Phi(t)$ we find

$$\Phi'(t) = \begin{bmatrix} 0 & 0 \\ -\frac{1}{2}(2e^{-2t}) & -2e^{-2t} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -e^{-2t} & -2e^{-2t} \end{bmatrix} ,$$

while the product $F(t)\Phi(t)$ is given by

$$\begin{bmatrix} 0 & 0 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{2}(1-e^{-2t}) & e^{-2t} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -e^{-2t} & -2e^{-2t} \end{bmatrix} , \tag{7}$$

showing that indeed $\Phi'(t) = F(t)\Phi(t)$ as required for $\Phi(t)$ to be a fundamental solution. Recall that the full solution for $x(t)$ is given by Equation 1 above. From this we see that we still need to calculate the second term above involving the fundamental solution $\Phi(t)$, the input coupling matrix $C(t)$, and the input $u(t)$ given by

$$\Phi(t) \int_{t_0}^{t} \Phi^{-1}(\tau) C(\tau) u(\tau) d\tau . \tag{8}$$

Now we can compute the inverse of our fundamental solution matrix $\Phi(t)^{-1}$ as

$$\Phi(t)^{-1} = \frac{1}{e^{-2t}} \begin{bmatrix} e^{-2t} & 0 \\ \frac{1}{2}(1-e^{-2t}) & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2}(e^{2t}-1) & e^{2t} \end{bmatrix} .$$

Then this term is given by

$$\begin{aligned}
&= \begin{bmatrix} 1 & 0 \\ -\frac{1}{2}(1-e^{-2t}) & e^{-2t} \end{bmatrix} \int_0^t \begin{bmatrix} 1 & 0 \\ \frac{1}{2}(e^{2\tau}-1) & e^{2\tau} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} d\tau \\
&= \begin{bmatrix} 1 & 0 \\ -\frac{1}{2}(1-e^{-2t}) & e^{-2t} \end{bmatrix} \int_0^t \begin{bmatrix} 1 \\ \frac{1}{2}e^{2\tau} - \frac{1}{2} + e^{2\tau} \end{bmatrix} d\tau \\
&= \begin{bmatrix} 1 & 0 \\ -\frac{1}{2}(1-e^{-2t}) & e^{-2t} \end{bmatrix} \begin{bmatrix} t \\ \frac{3}{4}(e^{2t}-1) - \frac{t}{2} \end{bmatrix} d\tau \\
&= \begin{bmatrix} t \\ -\frac{t}{2} + \frac{3}{4}(1-e^{-2t}) \end{bmatrix} .
\end{aligned}$$

Thus the entire solution for $x(t)$ is given by

$$x(t) = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2}(1-e^{-2t}) & e^{-2t} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + \begin{bmatrix} t \\ -\frac{t}{2} + \frac{3}{4}(1-e^{-2t}) \end{bmatrix} . \tag{9}$$

We can verify that this is indeed a solution by showing that it satisfies the original differential equation. We find $x'(t)$ given by

$$\begin{aligned}
x'(t) &= \begin{bmatrix} 0 & 0 \\ -e^{-2t} & -2e^{-2t} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + \begin{bmatrix} 1 \\ -\frac{1}{2} + \frac{3}{2}e^{-2t} \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{2}(1-e^{-2t}) & e^{-2t} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + \begin{bmatrix} 1 \\ -\frac{1}{2} + \frac{3}{2}e^{-2t} \end{bmatrix} ,
\end{aligned}$$

where we have used the factorization given in Equation 7. Inserting the the needed term to complete an expression for $x(t)$ (as seen in Equation 9) we find

$$\begin{aligned} x'(t) &= \begin{bmatrix} 0 & 0 \\ -1 & -2 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 \\ -\frac{1}{2}(1-e^{-2t}) & e^{-2t} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + \begin{bmatrix} t \\ -\frac{t}{2} + \frac{3}{4}(1-e^{-2t}) \end{bmatrix} \right) \\ &- \begin{bmatrix} 0 & 0 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} t \\ -\frac{t}{2} + \frac{3}{4}(1-e^{-2t}) \end{bmatrix} + \begin{bmatrix} 1 \\ -\frac{1}{2} + \frac{3}{2}e^{-2t} \end{bmatrix} . \end{aligned}$$

or

$$\begin{aligned} x'(t) &= \begin{bmatrix} 0 & 0 \\ -1 & -2 \end{bmatrix} x(t) - \begin{bmatrix} 0 \\ -\frac{3}{2}(1-e^{-2t}) \end{bmatrix} + \begin{bmatrix} 1 \\ -\frac{1}{2} + \frac{3}{2}e^{-2t} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ -1 & -2 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} , \end{aligned}$$

showing that indeed we do have a solution.

Problem 2.7 (solving a dynamic linear system)

Studying the homogeneous problem in this case we have

$$\frac{d}{dt} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} .$$

which has solution by inspection given by $x_1(t) = x_1(0)e^{-t}$ and $x_2(t) = x_2(0)e^{-t}$. Thus as a vector we have $x(t)$ given by

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-t} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} .$$

Thus the fundamental solution matrix $\Phi(t)$ for this problem is seen to be

$$\Phi(t) = e^{-t} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{so that} \quad \Phi^{-1}(t) = e^{t} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} .$$

Using Equation 8 we can calculate the inhomogeneous solution as

$$\begin{aligned} \Phi(t) \int_{t_0}^{t} \Phi^{-1}(\tau) C(\tau) u(\tau) d\tau &= e^{-t} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \int_0^t e^{\tau} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \end{bmatrix} d\tau \\ &= e^{-t}(e^t - 1) \begin{bmatrix} 5 \\ 1 \end{bmatrix} . \end{aligned}$$

Thus the total solution is given by

$$x(t) = e^{-t} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} + (1 - e^{-t}) \begin{bmatrix} 5 \\ 1 \end{bmatrix} .$$

Problem 2.8 (the reverse problem)

Warning: I was not really sure how to answer this question. There seem to be multiple possible continuous time systems for a given discrete time system and so multiple solutions are possible. If anyone has an suggestions improvements on this please let me know.

From the discussion in Section 2.4 in the book we can study our continuous system at only the discrete times t_k by considering

$$x(t_k) = \Phi(t_k, t_{k-1})x(t_{k-1}) + \int_{t_{k-1}}^{t_k} \Phi(t_k, \sigma)C(\sigma)u(\sigma)d\sigma\,. \tag{10}$$

Thus for the discrete time dynamic system given in this problem we could associate

$$\Phi(t_k, t_{k-1}) = \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix},$$

to be the state transition matrix which also happens to be a constant matrix. To complete our specification of the continuous problem we still need to find functions $C(\cdot)$ and $u(\cdot)$ such that they satisfy

$$\int_{t_{k-1}}^{t_k} \Phi(t_k, \sigma)C(\sigma)u(\sigma)d\sigma = \int_{t_{k-1}}^{t_k} \begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} C(\sigma)u(\sigma)d\sigma = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

There are many way to satisfy this equation. One simple method is to take $C(\sigma)$, the input coupling matrix, to be the identity matrix which then requires the input $u(\sigma)$ satisfy the following

$$\begin{bmatrix} 0 & 1 \\ -1 & 2 \end{bmatrix} \int_{t_{k-1}}^{t_k} u(\sigma)d\sigma = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

On inverting the matrix on the left-hand-side we obtain

$$\int_{t_{k-1}}^{t_k} u(\sigma)d\sigma = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

If we take $u(\sigma)$ as a constant say $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, then this equation will be satisfied if $u_2 = 0$, and $u_1 = -\frac{1}{\Delta t}$ with $\Delta t = t_k - t_{k-1}$ assuming a constant sampling step size Δt.

Problem 2.9 (conditions for observability and controllability)

Since the dynamic system we are given is continuous, with a dynamic coefficient matrix F given by $F = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, an input coupling matrix $C(t)$ given by $C = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, and a measurement sensitivity matrix $H(t)$ given by $H(t) = \begin{bmatrix} h_1 & h_2 \end{bmatrix}$, all of which are independent of time. The condition for observability is that the matrix M defined as

$$M = \begin{bmatrix} H^T & F^T H^T & (F^T)^2 H^T & \cdots & (F^T)^{n-1} H^T \end{bmatrix}, \tag{11}$$

has rank $n = 2$. We find with the specific H and F for this problem that

$$M = \left[\begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} \right] = \begin{bmatrix} h_1 & h_1 \\ h_2 & h_1 + h_2 \end{bmatrix},$$

needs to have rank 2. By reducing M to row reduced echelon form (assuming $h_1 \neq 0$) as

$$M \Rightarrow \begin{bmatrix} h_1 & h_1 \\ 0 & h_1 + h_2 - h_2 \end{bmatrix} \Rightarrow \begin{bmatrix} h_1 & h_1 \\ 0 & h_1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Thus we see that M will have rank 2 and our system will be observable as long as $h_1 \neq 0$. To be controllable we need to consider the matrix S given by

$$S = \begin{bmatrix} C & FC & F^2C & \cdots & F^{n-1}C \end{bmatrix}, \tag{12}$$

or in this case

$$S = \begin{bmatrix} c_1 & c_1 + c_2 \\ c_2 & c_2 \end{bmatrix}.$$

This matrix is the *same* as that in M except for the rows of S are exchanged from that of M. Thus for the condition needed for S to have a rank $n = 2$ requires $c_2 \neq 0$.

Problem 2.10 (controllability and observability of a dynamic system)

For this continuous time system the dynamic coefficient matrix $F(t)$ is given by $F(t) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, the input coupling matrix $C(t)$ is given by $C(t) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, and the measurement sensitivity matrix $H(t)$ is given by $H(t) = \begin{bmatrix} 0 & 1 \end{bmatrix}$. The observability of this system is determined by the rank of M defined in Equation 11, which in this case is given by

$$M = \left[\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Since this matrix M is of rank two, this system is observable. The controllability of this system is determined by the rank of the matrix S defined by Equation 12, which in this case since $FC = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ becomes

$$S = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{bmatrix}.$$

Since this matrix has a rank of two this system is controllable.

Problem 2.11 (the state transition matrix for a time-varying system)

For this problem the dynamic coefficient matrix is given by $F(t) = t \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. In terms of the components of the solution $\mathbf{x}(t)$ of we see that each $x_i(t)$ satisfies

$$\frac{dx_i(t)}{dt} = t x_i(t) \quad \text{for} \quad i = 1, 2.$$

Then solving this differential equation we have $x_i(t) = c_i e^{\frac{t^2}{2}}$ for $i = 1, 2$. As a vector $\mathbf{x}(t)$ can be written as

$$\mathbf{x}(t) = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} e^{\frac{t^2}{2}} = \begin{bmatrix} e^{\frac{t^2}{2}} & 0 \\ 0 & e^{\frac{t^2}{2}} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} .$$

Thus we find

$$\Phi(t) = e^{\frac{t^2}{2}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ,$$

is the fundamental solution and the state transition matrix $\Phi(\tau, t)$ is given by

$$\Phi(\tau, t) = \Phi(\tau)\Phi(t)^{-1} = e^{-\frac{1}{2}(t^2-\tau^2)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} .$$

Problem 2.12 (an example at finding the state transformation matrix)

We desire to find the state transition matrix for a continuous time system with a dynamic coefficient matrix given by

$$F = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} .$$

We will do this by finding the fundamental solution matrix $\Phi(t)$ that satisfies $\Phi'(t) = F\Phi(t)$, with an initial condition of $\Phi(0) = I$. We find the eigenvalues of F to be given by

$$\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = 0 \Rightarrow \lambda^2 - 1 = 0 \Rightarrow \lambda = \pm 1 .$$

The eigenvalue $\lambda_1 = -1$ has an eigenvector given by $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, while the eigenvalue $\lambda_2 = 1$ has an eigenvalue of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Thus the general solution to this linear time invariant system is given by

$$\mathbf{x}(t) = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} e^{-t} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{t} = \begin{bmatrix} e^{-t} & e^{t} \\ -e^{-t} & e^{t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} .$$

To satisfy the required initial conditions $\mathbf{x}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix}$, the coefficients c_1 and c_2 must equal

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} .$$

Thus the entire solution for $x(t)$ in terms of its two components $x_1(t)$ and $x_2(t)$ is given by

$$\begin{aligned} x(t) &= \frac{1}{2} \begin{bmatrix} e^{-t} & e^{t} \\ -e^{t} & e^{t} \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} e^{-t} + e^{t} & -e^{-t} + e^{t} \\ -e^{t} + e^{t} & e^{-t} + e^{t} \end{bmatrix} \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} . \end{aligned}$$

From which we see that the fundamental solution matrix $\Phi(t)$ for this system is given by

$$\Phi(t) = \frac{1}{2}\left[\begin{array}{cc} e^{-t}+e^{t} & -e^{-t}+e^{t} \\ -e^{t}+e^{t} & e^{-t}+e^{t} \end{array}\right].$$

The state transition matrix $\Phi(\tau,t) = \Phi(\tau)\Phi^{-1}(t)$. To get this we first compute Φ^{-1}. We find

$$\begin{aligned}
\Phi^{-1}(t) &= \frac{2}{(e^{-t}+e^{t})^2-(e^{-t}-e^{t})^2}\left[\begin{array}{cc} e^{-t}+e^{t} & e^{-t}-e^{t} \\ e^{-t}-e^{t} & e^{-t}+e^{t} \end{array}\right] \\
&= \frac{2}{((e^{-t}+e^{t})-(e^{-t}-e^{t}))((e^{-t}+e^{t})+(e^{-t}-e^{t}))}\left[\begin{array}{cc} e^{-t}+e^{t} & e^{-t}-e^{t} \\ e^{-t}-e^{t} & e^{-t}+e^{t} \end{array}\right] \\
&= \frac{1}{(2e^{t})(e^{-t})}\left[\begin{array}{cc} e^{-t}+e^{t} & e^{-t}-e^{t} \\ e^{-t}-e^{t} & e^{-t}+e^{t} \end{array}\right] \\
&= \frac{1}{2}\left[\begin{array}{cc} e^{-t}+e^{t} & e^{-t}-e^{t} \\ e^{-t}-e^{t} & e^{-t}+e^{t} \end{array}\right] = \Phi(t)\,.
\end{aligned}$$

Thus we have $\Phi(\tau,t)$ given by

$$\Phi(\tau,t) = \frac{1}{4}\left[\begin{array}{cc} e^{-\tau}+e^{\tau} & e^{-\tau}-e^{\tau} \\ e^{-\tau}-e^{\tau} & e^{-\tau}+e^{\tau} \end{array}\right]\left[\begin{array}{cc} e^{-t}+e^{t} & e^{-t}-e^{t} \\ e^{-t}-e^{t} & e^{-t}+e^{t} \end{array}\right].$$

Problem 2.13 (recognizing the companion form for $\frac{d^3y}{dt^3}$)

Part (a): Writing this system in the vector form with $\mathbf{x} = \left[\begin{array}{c} x_1(t) \\ x_2(t) \\ x_3(t) \end{array}\right]$, we have

$$\dot{\mathbf{x}}(t) = \left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right]\left[\begin{array}{c} x_1(t) \\ x_2(t) \\ x_3(t) \end{array}\right],$$

so we see the system companion matrix, F, is given by $F = \left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array}\right]$.

Part (b): For the F given above we recognize it as the companion matrix for the system $\frac{d^3y}{dt^3} = 0$, (see the section on Fundamental solutions of Homogeneous equations), and as such has a fundamental solution matrix $\Phi(t)$ given as in Example 2.5 of the appropriate dimension. That is

$$\Phi(t) = \left[\begin{array}{ccc} 1 & t & \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{array}\right].$$

Problem 2.14 (matrix exponentials of antisymmetric matrices are orthogonal)

If M is an antisymmetric matrix then $M^T = -M$. Consider the matrix A defined as the matrix exponential of M i.e. $A \equiv e^M$. Then since $A^T = e^{M^T} = e^{-M}$, is the inverse of e^M (equivalently A) we see that $A^T = A^{-1}$ so A is orthogonal.

Problem 2.15 (a derivation of the condition for continuous observability)

We wish to derive equation 2.32 which states that the observability of a continuous dynamic system is given by the singularity of the matrix $\mathcal{O}$ where

$$\mathcal{O} = \mathcal{O}(H, F, t_0, t_f) = \int_{t_0}^{t_f} \Phi^T(t) H^T(t) H(t) \Phi(t) dt \,,$$

in that if $\mathcal{O}$ is singular the Storm is not observable and if it is non-singular the system is observable. As in example 1.2 we measure $z(t)$ where $z(t)$ is obtained from $x(t)$ using the measurement sensitivity matrix $H(t)$ as $z(t) = H(t)x(t)$. Using our general solution for $x(t)$ from Equation 1 we have

$$z(t) = H(t)\Phi(t)\Phi(t_0)^{-1}x(t_0) + H(t)\Phi(t)\int_{t_0}^{t} \Phi^{-1}(\tau)C(\tau)u(\tau)d\tau \,, \tag{13}$$

observability is whether we can compute $x(t_0)$ given its inputs $u(\tau)$ and its outputs $z(t)$, over the real interval $t_0 < t < t_f$. Setting up an error criterion to estimate how well we estimate $\hat{x}_0$, assume that we have measured $z(t)$ out instantaneous error will then be

$$\begin{aligned}
\epsilon(t)^2 &= |z(t) - H(t)x(t)|^2 \\
&= x^T(t)H^T(t)H(t)x(t) - 2x^T(t)H^T(t)z(t) + |z(t)|^2 \,.
\end{aligned}$$

Since we are studying a linear continuous time system, the solution $x(t)$ in terms of the state transition matrix $\Phi(t, \tau)$, the input coupling matrix $C(t)$, the input $u(t)$, and the initial state $x(t_0)$ is given by Equation 1 above. Defining $\tilde{c}$ as the vector

$$\tilde{c} = \int_{t_0}^{t_f} \Phi^{-1}(\tau)C(\tau)u(\tau)d\tau \,,$$

we then have $x(t)$ given by $x(t) = \Phi(t)\Phi^{-1}(t_0)x(t_0) + \Phi(t)\tilde{c}$, thus the expression for $\epsilon(t)^2$ in terms of $x(t_0)$ is given by

$$\begin{aligned}
\epsilon^2(t) &= (x^T(t_0)\Phi^{-T}(t_0)\Phi^T(t) + \tilde{c}^T\Phi^T(t))H^T(t)H(t)(\Phi(t)\Phi^{-1}(t_0)x(t_0) + \Phi(t)\tilde{c}) \\
&- 2(x^T(t_0)\Phi^{-T}(t_0)\Phi^T(t) + \tilde{c}^T\Phi^T(t))H^T(t)z(t) + |z(t)|^2 \\
&= x^T(t_0)\Phi^{-T}(t_0)\Phi^T(t)H^T(t)H(t)\Phi(t)\Phi^{-1}(t_0)x(t_0) && (14) \\
&+ x^T(t_0)\Phi^{-T}(t_0)\Phi^T(t)H^T(t)H(t)\Phi(t)\tilde{c} && (15) \\
&+ \tilde{c}^T\Phi^T(t)H^T(t)H(t)\Phi(t)\Phi^{-1}(t_0)x(t_0) && (16) \\
&+ \tilde{c}^T\Phi^T(t)H^T(t)H(t)\Phi(t)\tilde{c} && (17) \\
&- 2x^T(t_0)\Phi^{-T}(t_0)\Phi^T(t)H^T(t)z(t) && (18) \\
&- 2\tilde{c}^T\Phi^T(t)H^T(t)z(t) && (19) \\
&+ |z(t)|^2 \,. && (20)
\end{aligned}$$

Since the terms corresponding to Equations 15, 16, and 18 are inner products they are equal to their transposes so the above is equal to

$$
\begin{aligned}
\epsilon^2(t) &= x^T(t_0)\Phi^{-T}(t_0)\Phi^T(t)H^T(t)H(t)\Phi(t)\Phi^{-1}(t_0)x(t_0) \\
&+ \left(2\tilde{c}\Phi^T(t)H^T(t)H(t)\Phi(t)\Phi^{-1}(t_0) - 2z^T(t)H(t)\Phi(t)\Phi^{-1}(t_0)\right)x(t_0) \\
&+ \tilde{c}^T\Phi^T(t)H^T(t)H(t)\Phi(t)\tilde{c} - 2\tilde{c}^T\Phi^T(t)H^T(t)z(t) + |z(t)|^2 .
\end{aligned}
$$

Now computing $||\epsilon||^2$ by integrating the above expression with respect to t over the interval $t_0 < t < t_f$ we have

$$
\begin{aligned}
||\epsilon||^2 &= x^T(t_0)\Phi^{-T}(t_0)\left(\int_{t_0}^{t_f}\Phi^T(t)H^T(t)H(t)\Phi(t)dt\right)\Phi^{-1}(t_0)x(t_0) \\
&+ \left(2\tilde{c}^T\left(\int_{t_0}^{t_f}\Phi^T(t)H^T(t)H(t)\Phi(t)dt\right)\Phi^{-1}(t_0) - 2\left(\int_{t_0}^{t_f}z^T(t)H(t)\Phi(t)dt\right)\Phi^{-1}(t_0)\right)x(t_0) \\
&+ \tilde{c}^T\left(\int_{t_0}^{t_f}\Phi^T(t)H^T(t)H(t)\Phi(t)dt\right)\tilde{c} - 2\tilde{c}^T\left(\int_{t_0}^{t_f}\Phi^T(t)H^T(t)z(t)dt\right) + \int_{t_0}^{t_f}|z(t)|^2dt .
\end{aligned}
$$

Defining $\mathcal{O}$ and $\tilde{z}$ as

$$
\begin{aligned}
\mathcal{O} \equiv \mathcal{O}(H, F, t_0, t_f) &= \int_{t_0}^{t_f}\Phi^T(t)H^T(t)H(t)\Phi(t)dt \qquad (21) \\
\tilde{z} &= \int_{t_0}^{t_f}\Phi^T(t)H^T(t)z(t)dt ,
\end{aligned}
$$

we see that the above expression for $||\epsilon||^2$ becomes

$$
\begin{aligned}
||\epsilon||^2 &= x^T(t_0)\Phi^{-T}(t_0)\mathcal{O}\Phi^{-1}(t_0)x(t_0) \\
&+ \left(2\tilde{c}^T\mathcal{O}\Phi^{-1}(t_0) - 2\tilde{z}^T\Phi^{-1}(t_0)\right)x(t_0) \\
&+ \tilde{c}^T\mathcal{O}\tilde{c} - 2\tilde{c}^T\tilde{z} + \int_{t_0}^{t_f}|z(t)|^2dt .
\end{aligned}
$$

Then by taking the derivative of $||\epsilon||^2$ with respect to the components of $x(t_0)$ and equating these to zero as done in Example 1.2, we can obtain an estimate for $x(t_0)$ by minimizing the above functional with respect to it. We find

$$
\hat{x}(t_0) = \left[\Phi^{-T}(t_0)\mathcal{O}\Phi^{-1}(t_0)\right]^{-1}\left[\Phi^{-T}(t_0)\mathcal{O}^T\tilde{c} - \Phi^{-T}(t_0)\tilde{z}\right] = \Phi^{-T}(t_0)\mathcal{O}^{-1}\left(\mathcal{O}^T\tilde{c} - \tilde{z}\right) .
$$

We can estimate $x(t_0)$ in this way using the equation above provided that $\mathcal{O}$, defined as Equation 21 is invertible, which was the condition we were to show.

Problem 2.16 (a derivation of the condition for discrete observability)

For this problem we assume that we are given the discrete time linear system and measurement equations in the standard form

$$
\begin{aligned}
x_k &= \Phi_{k-1}x_{k-1} + \Gamma_{k-1}u_{k-1} \qquad (22) \\
z_k &= H_kx_k + D_ku_k \quad \text{for} \quad k \geq 1 , \qquad (23)
\end{aligned}
$$

and that we wish to estimate the initial state x_0 from the received measurements z_k for a range of k say $1 \le k \le k_f$. To do this we will solve Equation 22 and 23 for x_k directly in terms of x_0 by induction. To get an idea of what the solution for x_k and z_k should look like a a function of k we begin by computing x_k and z_k for a few values of k. To begin with lets take $k = 1$ in Equation 22 and Equation 23 to find

$$\begin{aligned} x_1 &= \Phi_0 x_0 + \Gamma_0 u_0 \\ z_1 &= H_1 x_1 + D_1 u_1 = H_1 \Phi_0 x_0 + H_1 \Gamma_0 u_0 + D_1 u_1 \,. \end{aligned}$$

Where we have substituted x_1 into the second equation for z_1. Letting $k = 2$ in Equation 22 and Equation 23 we obtain

$$\begin{aligned} x_2 &= \Phi_1 x_1 + \Gamma_1 u_1 = \Phi_1(\Phi_0 x_0 + \Gamma_0 u_0) + \Gamma_1 u_1 \\ &= \Phi_1 \Phi_0 x_0 + \Phi_1 \Gamma_0 u_0 + \Gamma_1 u_1 \\ z_2 &= H_2 \Phi_1 \Phi_0 x_0 + H_2 \Phi_1 \Gamma_0 u_0 + H_2 \Gamma_1 u_1 \,. \end{aligned}$$

Observing one more value of x_k and z_k let $k = 3$ in Equation 22 and Equation 23 to obtain

$$\begin{aligned} x_3 &= \Phi_2 \Phi_1 \Phi_0 x_0 + \Phi_2 \Phi_1 \Gamma_0 u_0 + \Phi_2 \Gamma_1 u_1 + \Gamma_2 u_2 \\ z_3 &= H_3 \Phi_2 \Phi_1 \Phi_0 x_0 + H_3 \Phi_2 \Phi_1 \Gamma_0 u_0 + H_3 \Phi_2 \Gamma_1 u_1 + H_3 \Gamma_2 u_2 \,. \end{aligned}$$

From these specific cases we hypothesis that that the general expression for x_k in terms of x_0 is be given by the following specific expression

$$x_k = \left(\prod_{i=0}^{k-1} \Phi_i \right) x_0 + \sum_{l=0}^{k-1} \left(\prod_{i=0}^{k-1-l} \Phi_i \right) \Gamma_l u_l \tag{24}$$

Lets define some of these matrices. Define P_{k-1} as

$$P_{k-1} \equiv \prod_{i=0}^{k-1} \Phi_i = \Phi_{k-1} \Phi_{k-2} \cdots \Phi_1 \Phi_0 \,, \tag{25}$$

where since Φ_k are matrices the order of the factors in the product matters. Our expression for x_k in terms of x_0 becomes

$$x_k = P_{k-1} x_0 + \sum_{l=0}^{k-1} P_{k-1-l} \Gamma_l u_l \,.$$

From this expression for x_k we see that z_k is given by (in terms of x_0)

$$z_k = H_k P_{k-1} x_0 + H_k \sum_{l=0}^{k-1} P_{k-1-l} \Gamma_l u_l + D_k u_k \quad \text{for} \quad k \ge 1 \,. \tag{26}$$

We now set up a least squares problem aimed at the estimation of x_0. We assume we have k_f measurements of z_k and form the L_2 error functional $\epsilon(x_0)$ of all received measurements as

$$\epsilon^2(x_0) = \sum_{i=1}^{k_f} |H_i P_{i-1} x_0 + H_i \sum_{l=0}^{i-1} P_{i-1-l} \Gamma_l u_l + D_i u_i - z_i|^2$$

As in Example 1.1 in the book we can minimize $\epsilon(x_0)^2$ as a function of x_0 by taking the partial derivatives of the above with respect to x_0, setting the resulting expressions equal to zero and solving for x_0. To do this we simply things by writing $\epsilon(x_0)^2$ as

$$\epsilon^2(x_0) = \sum_{i=1}^{k_f} |H_i P_{i-1} x_0 - \tilde{z}_i|^2 \,,$$

where $\tilde{z}_i$ is defined as

$$\tilde{z}_i = z_i - H_i \sum_{l=0}^{i-1} P_{i-1-l} \Gamma_l u_l - D_i u_i \,. \tag{27}$$

With this definition the expression for $\epsilon^2(x_0)$ can be simplified by expanding the quadratic to get

$$\begin{aligned} \epsilon^2(x_0) &= \sum_{i=1}^{k_f} \left(x_0^T P_{i-1}^T H_i^T H_i P_{i-1} x_0 - 2 x_0^T P_{i-1}^T H_i^T \tilde{z}_i + \tilde{z}_i^T \tilde{z}_i \right) \\ &= x_0^T \left(\sum_{i=1}^{k_f} P_{i-1}^T H_i^T P_{i-1} \right) x_0 - 2 x_0^T \left(\sum_{i=1}^{k_f} P_{i-1}^T H_i^T \tilde{z}_i \right) + \sum_{i=1}^{k_f} \tilde{z}_i^T \tilde{z}_i \,. \end{aligned}$$

Taking the derivative of this expression and setting it equal to zero (so that we can solve for x_0) our least squares solution is given by solving

$$2 \mathcal{O} x_0 - 2 \left(\sum_{i=1}^{k_f} P_{i-1}^T H_i^T \tilde{z}_i \right) = 0 \,,$$

where we have defined the matrix $\mathcal{O}$ as

$$\mathcal{O} = \sum_{k=1}^{t_f} P_{i-1}^T H_i^T H_i P_{i-1} = \sum_{k=1}^{k_f} \left(\left[\prod_{i=0}^{k-1} \Phi_i \right]^T H_k^T H_k \left[\prod_{i=0}^{k-1} \Phi_i \right] \right) . \tag{28}$$

Where its important to take the products of the matrices Φ_k as in the order expressed in Equation 25. An estimate of x_0 can then be obtain as

$$\hat{x}_0 = \mathcal{O}^{-1} \sum_{i=1}^{k_f} P_{i-1}^T H_i^T \tilde{z}_i \,,$$

provided that the inverse of $\mathcal{O}$ exists, which is the desired discrete condition for observability.

Chapter 3: Random Processes and Stochastic Systems

Problem Solutions

Problem 3.1 (each pile contains one ace)

We can solve this problem by thinking about placing the aces individually ignoring the placement of the other cards. Then once the first ace is placed on a pile we have a probability of 3/4 to place the next ace in a untouched pile. Once this second ace is placed we have 2/4 of a probability of placing a new ace in another untouched pile. Finally, after the third ace is placed we have a probability of 1/4 of placing the final ace on the one pile that does not yet have an ace on it. Thus the probability that each pile contains an ace to be

$$\left(\frac{3}{4}\right)\left(\frac{2}{4}\right)\left(\frac{1}{4}\right) = \frac{3}{32}.$$

Problem 3.2 (a combinatorial identity)

We can show the requested identity by recalling that $\left(\begin{array}{c} n \\ k \end{array}\right)$ represents the number of ways to select k object from n where the order of the k selected objects does not matter. Using this representation we will derive an expression for $\left(\begin{array}{c} n \\ k \end{array}\right)$ as follows. We begin by considering the group of n objects with one object specified as distinguished or "special". Then the number of ways to select k objects from n can be decomposed into two distinct occurrences. The times when this "special" object *is* selected in the subset of size k and the times when its *not*. When it is *not* selected in the subset of size k we are specifying our k subset elements from the $n-1$ remaining elements giving $\left(\begin{array}{c} n-1 \\ k \end{array}\right)$ total subsets in this case. When it *is* selected into the subset of size k we have to select $k-1$ other elements from the $n-1$ remaining elements, giving $\left(\begin{array}{c} n-1 \\ k-1 \end{array}\right)$ additional subsets in this case. Summing the counts from these two occurrences we have that factorization can be written as the following

$$\left(\begin{array}{c} n \\ k \end{array}\right) = \left(\begin{array}{c} n-1 \\ k \end{array}\right) + \left(\begin{array}{c} n-1 \\ k-1 \end{array}\right).$$

Problem 3.3 (dividing the deck into four piles of cards)

We have $\left(\begin{array}{c} 52 \\ 13 \end{array}\right)$ ways of selecting the first hand of thirteen cards. After this hand is selected we have $\left(\begin{array}{c} 52-13 \\ 13 \end{array}\right) = \left(\begin{array}{c} 48 \\ 13 \end{array}\right)$ ways to select the second hand of cards. After these first two hands are selected we have $\left(\begin{array}{c} 52-2*13 \\ 13 \end{array}\right) = \left(\begin{array}{c} 26 \\ 13 \end{array}\right)$ ways to select the third hand after which the fourth hand becomes whatever cards are left. Thus the total number of ways to divide up a deck of 52 cards into four hands is given by the product of each of these expressions or

$$\left(\begin{array}{c} 52 \\ 13 \end{array}\right)\left(\begin{array}{c} 48 \\ 13 \end{array}\right)\left(\begin{array}{c} 26 \\ 13 \end{array}\right).$$

Problem 3.4 (exactly three spades in a hand)

We have $\left(\begin{array}{c} 52 \\ 13 \end{array}\right)$ ways to draw a random hand of cards. To draw a hand of cards with explicitly three spades, the spades can be drawn in $\left(\begin{array}{c} 13 \\ 3 \end{array}\right)$ ways, and the remaining nine other cards can be drawn in $\left(\begin{array}{c} 52-13 \\ 9 \end{array}\right) = \left(\begin{array}{c} 39 \\ 9 \end{array}\right)$ ways. The probability we have the hand requested is then

$$\frac{\left(\begin{array}{c} 13 \\ 3 \end{array}\right)\left(\begin{array}{c} 39 \\ 9 \end{array}\right)}{\left(\begin{array}{c} 52 \\ 13 \end{array}\right)}.$$

Problem 3.5 (south has three spades when north has three spades)

Since we are told that North has exactly three spades from the thirteen possible spade cards the players at the West, East, and South locations must have the remaining spade cards. Since they are assumed to be dealt randomly among these three players the probability South has exactly three of them is

$$\left(\begin{array}{c} 10 \\ 3 \end{array}\right)\left(\frac{2}{3}\right)^7\left(\frac{1}{3}\right)^3,$$

This is the same as a binomial distribution with probability of success of 1/3 (i.e. a success is when a spade goes to the player South) and 10 trials. In general, the probability South has k spade cards is given by

$$\left(\begin{array}{c} 10 \\ k \end{array}\right)\left(\frac{1}{3}\right)^k\left(\frac{2}{3}\right)^{10-k} \quad k = 0, 1, \cdots, 10.$$

Problem 3.6 (having 7 hearts)

Part (a): The number of ways we can select thirteen random cards from 52 total cards is $\left(\begin{array}{c} 52 \\ 13 \end{array}\right)$. The number of hands that contain seven hearts can be derived by first selecting the seven hearts to be in that hand in $\binom{13}{7}$ ways and then selecting the remaining $13-7=6$ cards in $\binom{52-13}{6} = \binom{39}{6}$ ways. Thus the probability for this hand of seven hearts is given by

$$\frac{\binom{13}{7}\binom{39}{6}}{\binom{52}{13}} = 0.0088\,.$$

Part (b): Let E_i be the event that our hand has i hearts where $0 \le i \le 13$. Then $P(E_7)$ is given in Part (a) above. Let F be the event that we observe a one card from our hand and it is a heart card. Then we want to calculate $P(E_7|F)$. From Bayes' rule this is given by

$$P(E_7|F) = \frac{P(F|E_7)P(E_7)}{P(F)}\,.$$

Now $P(F|E_i)$ is the probability one card observed as hearts given that we have i hearts in the hand. So $P(F|E_i) = \frac{i}{13}$ for $0 \le i \le 13$ and the denominator $P(F)$ can be computed as

$$P(F) = \sum_{i=0}^{13} P(F|E_i)P(E_i) = \sum_{i=0}^{13} \left(\frac{i}{13}\right) \frac{\binom{13}{i}\binom{52-13}{13-i}}{\binom{52}{13}}$$

Using this information and Bayes' rule above we can compute $P(E_7|F)$. Performing the above summation that $P(F) = 0.25$, see the MATLAB script `prob_3_6.m`. After computing this numerically we recognize that it is the probability we randomly draw a heart card and given that there are 13 cards from 52 this probability $P(F) = \frac{13}{52} = 0.25$. Then computing the desired probability $P(E_7|F)$ we find

$$P(E_7|F) = 0.0190\,.$$

As a sanity check note that $P(E_7|F)$ is greater than $P(E_7)$ as it should be, since once we have seen a heart in the hand there is a greater chance we will have seven hearts in that hand.

Problem 3.7 (the correlation coefficient between sums)

The correlation of a vector valued process $x(t)$ has components given by

$$E\langle x_i(t_1), x_j(t_2)\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i(t_1)x_j(t_2)p(x_i(t_1)x_j(t_2))dx_i(t_1)dx_j(t_2)\,.$$

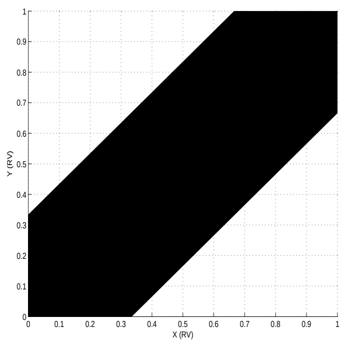

Figure 1: The integration region for Problem 3.8.

Using this definition lets begin by computing $E\langle Y_{n-1}, Y_n\rangle$. We find

$$\begin{aligned} E\langle Y_{n-1}, Y_n\rangle &= E\langle \sum_{j=1}^{n-1} X_j, \sum_{k=1}^{n} X_k\rangle \\ &= \sum_{j=1}^{n-1}\sum_{k=1}^{n} E\langle X_j X_k\rangle \,. \end{aligned}$$

Since the random variables X_i are zero mean and independent with individual variance of σ_X^2, we have that $E\langle X_j, X_k\rangle = \sigma_X^2 \delta_{kj}$ with δ_{kj} the Kronecker delta and the above double sum becomes a single sum given by

$$\sum_{j=1}^{n-1} E\langle X_j, X_j\rangle = (n-1)\sigma_X^2 \,.$$

Then the correlation coefficient is obtained by dividing the above expression by

$$\sqrt{E\langle Y_{n-1}^2\rangle E\langle Y_n^2\rangle}\,,$$

To compute $E\langle Y_n^2\rangle$, we have

$$E\langle Y_n^2\rangle = \sum_{j=1}^{n}\sum_{k=1}^{n} E\langle X_j X_k\rangle = n\sigma_X^2\,,$$

using the same logic as before. Thus our correlation coefficient r is

$$r = \frac{E\langle Y_{n-1}, Y_n\rangle}{\sqrt{E\langle Y_{n-1}^2\rangle E\langle Y_n^2\rangle}} = \frac{(n-1)\sigma_X^2}{\sqrt{(n-1)\sigma_X^2\, n\sigma_X^2}} = \left(\frac{n-1}{n}\right)^{1/2}\,.$$

Problem 3.8 (the density for $Z = |X - Y|$)

To derive the probability distribution for Z defined as $Z = |X - Y|$, we begin by considering the cumulative distribution function for the random variable Z defined as

$$F_Z(z) = Pr\{Z \le z\} = Pr\{|X - Y| \le z\}\,.$$

The region in the $X-Y$ plane where $|X-Y| \leq z$ is bounded by a strip around the line $X = Y$ given by

$$X - Y = \pm z \quad \text{or} \quad Y = X \pm z\,,$$

see Figure 1. Thus we can evaluate this probability $Pr\{|X-Y| \leq z\}$ as follows

$$F_Z(z) = \iint_{\Omega_{XY}} p(x,y)dxdy = \iint_{|X-Y|\leq z} dxdy\,.$$

This later integral can be evaluated by recognizing that the geometric representation of an integral is equivalent to the *area* in the $X-Y$ plane. From Figure 1 this is given by the sum of two trapezoids (the one above and the one below the line $X = Y$). Thus we can use the formula for the area of a trapezoid to evaluate the above integral. The area of a trapezoid requires knowledge of the lengths of the two trapezoid "bases" and its height. Both of these trapezoids have a larger base of $\sqrt{2}$ units long (the length of the diagonal line $X = Y$). For the trapezoid above the line $X = Y$ the other base has a length that can be derived by computing the distance between its two endpoints of $(0, z)$ and $(1-z, 1)$ or

$$b^2 = (0-(1-z))^2 + (z-1)^2 = 2(z-1)^2 \quad \text{for} \quad 0 \leq z \leq 1\,,$$

where b is this upper base length. Finally, the height of each trapezoid is z. Thus each trapezoid has an area given by

$$A = \frac{1}{2}z(\sqrt{2} + \sqrt{2(z-1)^2}) = \frac{z}{\sqrt{2}}(1+|z-1|) = \frac{z}{\sqrt{2}}(1+1-z) = \frac{1}{\sqrt{2}}z(2-z)\,.$$

Thus we find $Pr\{Z \leq z\}$ given by (remembering to double the above expression)

$$F_Z(z) = \frac{2}{\sqrt{2}}z(2-z) = \sqrt{2}(2z-z^2)\,.$$

Thus the probability density function for Z is then given by $F_Z'(z)$ or

$$f_Z(z) = 2\sqrt{2}(1-z)\,.$$

Problem 3.11 (an example autocorrelation functions)

Part (a): To be a valid autocorrelation function, $\psi_x(\tau)$ one must have the following properties

- it must be even
- it must have its maximum at the origin
- it must have a non-negative Fourier transform

For the given proposed autocorrelation function, $\psi_x(\tau)$, we see that it is even, has its maximum at the origin, and has a Fourier transform given by

$$\int_{-\infty}^{\infty} \frac{1}{1+\tau^2} e^{-j\omega\tau} d\tau = \pi e^{-|\omega|} \,, \tag{29}$$

which is certainly non-negative. Thus $\psi_x(\tau)$ is a valid autocorrelation function.

Part (b): We want to calculate the power spectral density (PSD) of $y(t)$ given that it is related to the stochastic process $x(t)$ by

$$y(t) = (1 + mx(t)) \cos(\Omega t + \lambda) \,.$$

The direct method of computing the power spectral density of $y(t)$ would be to first compute the autocorrelation function of $y(t)$ in terms of the autocorrelation function of $x(t)$ and then from this, compute the PSD of $y(t)$ in terms of the known PSD of $x(t)$. To first evaluate the autocorrelation of $y(t)$ we have

$$\begin{aligned}
\psi_y(\tau) &= E\langle y(t) y(t+\tau) \rangle \\
&= E\langle (1 + mx(t)) \cos(\Omega t + \lambda)(1 + mx(t+\tau)) \cos(\Omega(t+\tau) + \lambda) \rangle \\
&= E\langle (1 + mx(t))(1 + mx(t+\tau)) \rangle E\langle \cos(\Omega t + \lambda) \cos(\Omega(t+\tau) + \lambda) \rangle \,,
\end{aligned}$$

since we are told that the random variable λ is independent of $x(t)$. Continuing we can expand the products involving $x(t)$ to find

$$\begin{aligned}
\psi_y(\tau) &= \left(1 + mE\langle x(t) \rangle + mE\langle x(t+\tau) \rangle + m^2 E\langle x(t) x(t+\tau) \rangle\right) E\langle \cos(\Omega t + \lambda) \cos(\Omega(t+\tau) + \lambda) \rangle \\
&= (1 + m^2 \psi_x(\tau)) E\langle \cos(\Omega t + \lambda) \cos(\Omega(t+\tau) + \lambda) \rangle \,,
\end{aligned}$$

using the fact that $E\langle x(t) \rangle = 0$. Continuing to evaluate $\psi_y(\tau)$ we use the product of cosigns identity

$$\cos(\theta_1) \cos(\theta_2) = \frac{1}{2}(\cos(\theta_1 + \theta_2) + \cos(\theta_1 - \theta_2)) \,, \tag{30}$$

to find

$$\begin{aligned}
E\langle \cos(\Omega t + \lambda) \cos(\Omega(t+\tau) + \lambda) \rangle &= \frac{1}{2} E\langle \cos(2\Omega t + \Omega\tau + 2\lambda) + \cos(\Omega\tau) \rangle \\
&= \frac{1}{2} \cos(\Omega\tau) \,,
\end{aligned}$$

since the expectation of the first term is zero. Thus we find for $\psi_y(\tau)$ the following

$$\psi_y(\tau) = \frac{1}{2}(1 + m^2 \psi_x(\tau)) \cos(\Omega\tau) = \frac{1}{2}\left(1 + \frac{m^2}{\tau^2 + 1}\right) \cos(\Omega\tau) \,.$$

To continue we will now take this expression for $\psi_y(\tau)$ and compute its PSD function. Recalling the product of convolution identity for Fourier transforms of

$$f(\tau) g(\tau) \Leftrightarrow (\hat{f} \star \hat{g})(\omega) \,,$$

and the fact that the Fourier Transform (FT) of $\cos(a\tau)$ given by

$$\int_{-\infty}^{\infty} \cos(a\tau) e^{-j\omega\tau} d\tau = \pi(\delta(\omega - a) + \delta(\omega + a)) \,. \tag{31}$$

We begin with the Fourier transform of the expression $\frac{\cos(\Omega\tau)}{1+\tau^2}$. We find

$$\begin{aligned}\int_{-\infty}^{\infty}\left(\frac{\cos(\Omega\tau)}{1+\tau^2}\right)e^{-j\omega\tau}d\tau &= \pi(\delta(\tau-\Omega)+\delta(\tau+\Omega))\star\pi e^{-|\tau|} \\ &= \pi^2\int_{-\infty}^{\infty}e^{-|\tau-\omega|}(\delta(\tau-\Omega)+\delta(\tau+\Omega))d\tau \\ &= \pi^2\left(e^{-|\Omega-\omega|}+e^{-|\Omega+\omega|}\right)\,,\end{aligned}$$

Thus the total PSD of $y(t)$ is then given by

$$\Psi_y(\tau)=\frac{\pi}{2}(\delta(\omega-\Omega)+\delta(\omega+\Omega))+\frac{\pi^2m^2}{2}\left(e^{-|\Omega-\omega|}+e^{-|\Omega+\omega|}\right)\,,$$

which shows that the combination of a fixed frequency term and an exponential decaying component.

Problem 3.12 (do PSD functions always decay to zero)

The answer to the proposed question is no and an example where $\lim_{|\omega|\to\infty}\Psi_x(\omega)\neq 0$ is if $x(t)$ is the *white noise* process. This process has an autocorrelation function that is a delta function

$$\psi_x(\tau)=\sigma^2\delta(\tau)\,, \tag{32}$$

which has a Fourier transform $\Psi_x(\omega)$ that is a constant

$$\Psi_x(\omega)=\sigma^2\,. \tag{33}$$

This functional form does not have limits that decay to zero as $|\omega|\to\infty$. This assumes that the white noise process is mean square continuous.

Problem 3.13 (the Dryden turbulence model)

The Dryden turbulence model a type of exponentially correlated autocorrelation model under which when $\psi_x(\tau)=\hat{\sigma}^2e^{-\alpha|\tau|}$ has a power spectral density (PSD) given by

$$\Psi_x(\omega)=\frac{2\hat{\sigma}^2\alpha}{\omega^2+\alpha^2}\,. \tag{34}$$

From the given functional form for the Dryden turbulence PSD given in the text we can write it as

$$\Psi(\omega)=\frac{2\left(\frac{\sigma^2}{\pi}\right)\left(\frac{V}{L}\right)}{\omega^2+\left(\frac{V}{L}\right)^2} \tag{35}$$

To match this to the exponential decaying model requires $\alpha = \frac{V}{L}$, $\hat{\sigma}^2 = \frac{\sigma^2}{\pi}$, and the continuous state space formulation of this problem is given by

$$\begin{aligned}
\dot{x}(t) &= -\alpha x(t) + \hat{\sigma}\sqrt{2\alpha} w(t) \\
&= -\left(\frac{V}{L}\right) x(t) + \left(\frac{\sigma}{\sqrt{\pi}}\right)\sqrt{2\frac{V}{L}} w(t) \\
&= -\left(\frac{V}{L}\right) x(t) + \sigma\sqrt{\frac{2V}{\pi L}} w(t) \,.
\end{aligned}$$

The different models given in this problem simply specify different constants to use in the above formulation.

Problem 3.14 (computing $\psi_x(\tau)$ and $\Psi_x(\omega)$ for a product of cosigns)

Part (a): Note that for the given stochastic process $x(t)$ we have $E\langle x(t)\rangle = 0$, due to the randomness of the variables θ_i for $i = 1, 2$. To derive the autocorrelation function for $x(t)$ consider $E\langle x(t)x(t+\tau)\rangle$ as

$$\begin{aligned}
E\langle x(t)x(t+\tau)\rangle &= E\langle \cos(\omega_0 t + \theta_1)\cos(\omega_0 t + \theta_2)\cos(\omega_0(t+\tau) + \theta_1)\cos(\omega_0(t+\tau) + \theta_2)\rangle \\
&= E\langle \cos(\omega_0 t + \theta_1)\cos(\omega_0(t+\tau) + \theta_1)\rangle E\langle \cos(\omega_0 t + \theta_2)\cos(\omega_0(t+\tau) + \theta_2)\rangle \,,
\end{aligned}$$

by the independence of the random variables θ_1 and θ_2. Recalling the product of cosign identity given in Equation 30 we have that

$$\begin{aligned}
E\langle \cos(\omega_0 t + \theta_1)\cos(\omega_0(t+\tau) + \theta_1)\rangle &= \frac{1}{2}E\langle \cos(2\omega_0 t + \omega_0\tau + 2\theta_1)\rangle + \frac{1}{2}E\langle \cos(\omega_0\tau)\rangle \\
&= \frac{1}{2}\cos(\omega_0\tau) \,.
\end{aligned}$$

So the autocorrelation function for $x(t)$ (denoted $\psi_x(\tau)$) then becomes, since we have *two* products of the above expression for $E\langle x(t)x(t+\tau)\rangle$, the following

$$\psi_x(\tau) = \frac{1}{4}\cos(\omega_0\tau)^2 \,.$$

Since this is a function of only τ, the stochastic process $x(t)$ is wide-sense stationary.

Part (b): To calculate $\Psi_x(\omega)$ we again use the product of cosign identity to write $\psi_x(\tau)$ as

$$\psi_x(\tau) = \frac{1}{4}\left(\frac{1}{2}(\cos(2\omega_0\tau) + 1)\right) \,.$$

Then to take the Fourier transform (FT) of $\psi_x(\tau)$ we need the Fourier transform of $\cos(\cdot)$ and the Fourier transform of the constant 1. The Fourier transform of $\cos(\cdot)$ is given in Equation 31 while the Fourier transform of 1 is given by

$$\int_{-\infty}^{\infty} 1 e^{-j\omega\tau} d\tau = 2\pi\delta(\omega) \,. \tag{36}$$

Thus the power spectral density of $x(t)$ is found to be

$$\Psi_x(\omega) = \frac{\pi}{8}(\delta(\omega - 2\omega_0) + \delta(\omega + 2\omega_0) + \frac{\pi}{4}\delta(\omega) .$$

Part (c): Ergodicity of $x(t)$ means that all of this process's statistical parameters, mean, variance etc. can be determined from an observation of its historical time series. That is its time-averaged statistics are equivalent to the ensemble average statistics. For this process again using the product of cosign identity we can write it as

$$x(t) = \frac{1}{2}\cos(2\omega_0 t + \theta_1 + \theta_2) + \frac{1}{2}\cos(\theta_1 + \theta_2) .$$

Then for every realization of this process θ_1 and θ_2 are specified fixed constants. Taking the *time* average of $x(t)$ as apposed to the parameter (θ_1 and θ_2) averages we then obtain

$$E_t\langle x(t)\rangle = \frac{1}{2}\cos(\theta_1 + \theta_2) ,$$

which is not zero in general. Averaging over the ensemble of signals $x(t)$ (for all parameters θ_1 and θ_2) we do obtain an expectation of zero. The fact that the time average of $x(t)$ does not equal the parameter average implies that $x(t)$ is *not* ergodic.

Problem 3.15 (the real part of an autocorrelation function)

From the discussion in the book if $x(t)$ is assumed to be a real valued stochastic process then it will have a real autocorrelation function $\psi(\tau)$, so its real part will the same as itself and by definition will again be an autocorrelation function. In the case where the stochastic process $x(t)$ is complex the common definition of the autocorrelation function is

$$\psi(\tau) = E\langle x(t)x^*(t + \tau)\rangle , \tag{37}$$

which may or may not be real and depends on the values taken by $x(t)$. To see if the real part of $\psi(\tau)$ is an autocorrelation function recall that for any complex number z the real part of z can be obtained by

$$\text{Re}(z) = \frac{1}{2}(z + z^*) , \tag{38}$$

so that if we define the real part of $\psi(\tau)$ to be $\psi_r(\tau)$ we have that

$$\begin{aligned}
\psi_r(\tau) &= E\langle \text{Re}(x(t)x^*(t + \tau))\rangle \\
&= \frac{1}{2}E\langle (x(t)x^*(t + \tau) + x^*(t)x(t + \tau))\rangle \\
&= \frac{1}{2}E\langle (x(t)x^*(t + \tau)\rangle + \frac{1}{2}E\langle x^*(t)x(t + \tau))\rangle \\
&= \frac{1}{2}\psi(\tau) + \frac{1}{2}\psi^*(\tau) .
\end{aligned}$$

From which we can see that $\psi_r(\tau)$ is a symmetric function since $\psi(\tau)$ is. Now both $\psi(\tau)$ and $\psi^*(\tau)$ have their maximum at $\tau = 0$ so $\psi_r(\tau)$ will have its maximum there also. Finally, the Fourier transform (FT) of $\psi(\tau)$ is nonnegative and thus the FT of $\psi^*(\tau)$ must be nonnegative which implies that the FT of $\psi_r(\tau)$ is nonnegative. Since $\psi_r(\tau)$ satisfies all of the requirements on page 21 for an autocorrelation function, $\psi_r(\tau)$ is an autocorrelation function.

Problem 3.16 (the cross-correlation of a cosign modified signal)

We compute the cross-correlation $\psi_{xy}(\tau)$ directly

$$\begin{aligned}
\psi_{xy}(\tau) &= E\langle x(t)y(t+\tau)\rangle \\
&= E\langle x(t)x(t+\tau)\cos(\omega t + \omega\tau + \theta)\rangle \\
&= E\langle x(t)x(t+\tau)\rangle E\langle \cos(\omega t + \omega\tau + \theta)\rangle \,,
\end{aligned}$$

assuming that $x(t)$ and θ are independent. Now $E\langle x(t)x(t+\tau)\rangle = \psi_x(\tau)$ by definition. We next compute

$$\begin{aligned}
E\langle \cos(\omega t + \omega\tau + \theta)\rangle &= \frac{1}{2\pi}\int_0^{2\pi} \cos(\omega t + \omega\tau + \theta)d\theta \\
&= \frac{1}{2\pi}\left(\sin(\omega t + \omega\tau + \theta)\right|_0^{2\pi} = 0\,.
\end{aligned}$$

Thus $\psi_{xy}(\tau) = 0$.

Problem 3.17 (the autocorrelation function for the integral)

We are told the autocorrelation function for $x(t)$ is given by $\psi_x(\tau) = e^{-|\tau|}$ and we want to compute the autocorrelation function for $y(t) = \int_0^t x(u)du$. Computing this directly we have

$$\begin{aligned}
E\langle y(t)y(t+\tau)\rangle &= E\langle \left(\int_0^t x(u)du\right)\left(\int_0^{t+\tau} x(v)dv\right)\rangle \\
&= \int_0^t \int_0^{t+\tau} E\langle x(u)x(v)\rangle dvdu \\
&= \int_0^t \int_0^{t+\tau} e^{-|u-v|} dvdu\,,
\end{aligned}$$

Where we have used the fact that we know the autocorrelation function for $x(t)$ that is $E\langle x(u)x(v)\rangle = e^{-|u-v|}$. To perform this double integral in the (u, v) plane to evaluate $|u-v|$ we need to break the domain of integration up into two regions depending on whether $v < u$

or $v > u$. We find (assuming that $\tau > 0$)

$$
\begin{aligned}
&= \int_{u=0}^{t}\int_{v=0}^{u} e^{-|u-v|}dvdu + \int_{u=0}^{t}\int_{v=u}^{t+\tau} e^{-|u-v|}dvdu \\
&= \int_{u=0}^{t}\int_{v=0}^{u} e^{-(u-v)}dvdu + \int_{u=0}^{t}\int_{v=u}^{t+\tau} e^{-(v-u)}dvdu \\
&= \int_{u=0}^{t}\int_{v=0}^{u} e^{-u}e^{v}dvdu + \int_{u=0}^{t}\int_{v=u}^{t+\tau} e^{-v}e^{u}dvdu \\
&= \int_{u=0}^{t} e^{-u}(e^{u}-1)du - \int_{u=0}^{t} e^{u}\left(e^{-v}\big|_{u}^{t+\tau}\right)du \\
&= \int_{u=0}^{t} (1-e^{-u})du - \int_{u=0}^{t} e^{u}\left(e^{-(t+\tau)} - e^{-u}\right)du \\
&= t + e^{-t} - 1 - e^{-(t+\tau)}\int_{u=0}^{t} e^{u}du + t \\
&= 2t + e^{-t} - e^{-\tau} + e^{-(t+\tau)} - 1\,.
\end{aligned}
$$

As this is *not* a function of only τ the stochastic process $y(t)$ is not wide-sense stationary. The calculation when $\tau < 0$ would be similar.

Problem 3.18 (the power spectral density of a cosign modified signal)

When $y(t) = x(t)\cos(\Omega t + \theta)$ we find its autocorrelation function $\psi_y(\tau)$ given by

$$
\begin{aligned}
\psi_y(\tau) &= E\langle x(t+\tau)x(t)\cos(\Omega(t+\tau)+\theta)\cos(\Omega t+\theta)\rangle \\
&= \psi_x(\tau)E\langle cos(\Omega(t+\tau)+\theta)\cos(\Omega t+\theta)\rangle \\
&= \frac{1}{2}\psi_x(\tau)\cos(\Omega\tau)\,.
\end{aligned}
$$

Then using this expression, the power spectral density of the signal $y(t)$ where y's autocorrelation function $\psi_y(\tau)$ is a product like above is the convolution of the Fourier transform of $\psi_x(\tau)$ and that of $\frac{1}{2}\cos(\Omega\tau)$. The Fourier transform of $\psi_x(\tau)$ is given in the problem. The Fourier transform of $\frac{1}{2}\cos(\Omega\tau)$ is given by Equation 31 or

$$
\frac{\pi}{2}(\delta(\omega-\Omega) + \delta(\omega+\Omega))\,.
$$

Thus the power spectral density for $y(t)$ is given by

$$
\begin{aligned}
\Psi_y(\omega) &= \frac{\pi}{2}\int_{-\infty}^{\infty}\Psi_x(\xi-\omega)(\delta(\xi-\Omega)+\delta(\xi+\Omega))d\xi \\
&= \frac{\pi}{2}(\Psi_x(\Omega-\omega) + \Psi_x(-\Omega-\omega)) \\
&= \frac{\pi}{2}(\Psi_x(\omega-\Omega) + \Psi_x(\omega+\Omega))\,.
\end{aligned}
$$

The first term in the above expression is $\Psi_x(\omega)$ shifted to the right by Ω, while the second term is $\Psi_x(\omega)$ shifted to the left by Ω. Since we are told that $\Omega > a$ we have that these two shifts move the the functional form of $\Psi_x(\omega)$ to the point where there is no overlap between the support of the two terms.

Problem 3.19 (definitions of random processes)

Part (a): A stochastic process is wide-sense stationary (WSS) if it has a constant mean for all time i.e. $E\langle x(t)\rangle = c$ and its second order statistics are independent of the time origin. That is, its autocorrelation function defined by $E\langle x(t_1)x(t_2)^t\rangle$ is a function of the time difference $t_2 - t_1$, rather than an arbitrary function of two variables t_1 and t_2. In equations this is represented as

$$E\langle x(t_1)x(t_2)^t\rangle = Q(t_2 - t_1)\,, \tag{39}$$

where $Q(\cdot)$ is a arbitrary function.

Part (b): A stochastic process $x(t)$ is strict-sense stationary (SSS) if it has *all* of its pointwise sample statistics independent of the time origin. In terms of the density function of samples of $x(t)$ this becomes

$$p(x_1, x_2, \cdots, x_n, t_1, t_2, \cdots, t_n) = p(x_1, x_2, \cdots, x_n, t_1 + \epsilon, t_2 + \epsilon, \cdots, t_n + \epsilon)\,.$$

Part (c): A linear system is said to realizable if the time domain representation of the impulse response of the system $h(t)$ is zero for $t < 0$. This a representation of the fact that in the time domain representation of the output signal $y(t)$ cannot depend on values of the input signal $x(t)$ occurring after time t. That is if $h(t) = 0$, when $t < 0$ we see that our system output $y(t)$ is given by

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)x(\tau)d\tau = \int_{-\infty}^{t} h(t-\tau)x(\tau)d\tau\,,$$

and $y(t)$ can be computed only using values of $x(\tau)$ "in the past" i.e. when $\tau < t$.

Part (d): Considering the table of properties required for an autocorrelation function given on page 21 the only one that is not obviously true for the given expression $\psi(\tau)$ is that the Fourier transform of $\psi(\tau)$ be nonnegative. Using the fact that the Fourier transform of this function (called the **triangular function**) is given by

$$\int_{-\infty}^{\infty} \text{tri}(a\tau)e^{j\omega\tau}d\tau = \frac{1}{|a|}\text{sinc}^2(\frac{\omega}{2\pi a})\,, \tag{40}$$

where the functions $\text{tri}(\cdot)$ and $\text{sinc}(\cdot)$ are defined by

$$\text{tri}(\tau) = \max(1-|\tau|, 0) = \begin{cases} 1-|\tau| & |\tau| < 1 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \tag{41}$$

$$\text{sinc}(\tau) = \frac{\sin(\pi\tau)}{\pi\tau}\,. \tag{42}$$

This result is derived when $a = 1$ in Problem 3.20 below. We see that in fact the above Fourier transform *is* nonnegative and the given functional form for $\psi(\tau)$ is an autocorrelation function.

Problem 3.20 (the power spectral density of the product with a cosign)

The autocorrelation function for $y(t)$ is given by

$$\psi_y(\tau) = \psi_x(\tau)\frac{1}{2}\cos(\omega_0\tau)\,,$$

see Exercise 3.18 above where this expression is derived. Then the power spectral density, $\Psi_y(\omega)$, is the Fourier transform of the above product, which in tern is the convolution of the Fourier transforms of the individual terms in the product above. Since the Fourier transform of $\cos(\omega_0\tau)$ is given by Equation 31 we need to compute the Fourier transform of $\psi_x(\tau)$.

$$\begin{aligned}
\int_{-\infty}^{\infty}\psi_x(\tau)e^{-j\omega\tau}d\tau &= \int_{-1}^{0}(1+\tau)e^{-j\omega\tau}d\tau + \int_{0}^{1}(1-\tau)e^{-j\omega\tau}d\tau \\
&= \int_{-1}^{0}e^{-j\omega\tau}d\tau + \int_{-1}^{0}\tau e^{-j\omega\tau}d\tau + \int_{0}^{1}e^{-j\omega\tau}d\tau - \int_{0}^{1}\tau e^{-j\omega\tau}d\tau \\
&= \left.\frac{e^{-j\omega\tau}}{(-j\omega)}\right|_{-1}^{0} + \left.\frac{\tau e^{-j\omega\tau}}{(-j\omega)}\right|_{-1}^{0} - \int_{-1}^{0}\frac{e^{-j\omega\tau}}{(-j\omega)}d\tau \\
&+ \left.\frac{e^{-j\omega\tau}}{(-j\omega)}\right|_{0}^{1} - \left.\frac{\tau e^{-j\omega\tau}}{(-j\omega)}\right|_{0}^{1} + \int_{0}^{1}\frac{e^{-j\omega\tau}}{(-j\omega)}d\tau \\
&= \frac{1-e^{j\omega}}{(-j\omega)} + \frac{e^{j\omega}}{(-j\omega)} - \frac{1}{(-j\omega)^2}\left.e^{-j\omega\tau}\right|_{-1}^{0} \\
&+ \frac{e^{-j\omega}-1}{(-j\omega)} - \frac{e^{-j\omega}}{(-j\omega)} + \frac{1}{(-j\omega)^2}\left.e^{-j\omega\tau}\right|_{0}^{1} \\
&= \frac{2}{\omega^2} - \frac{e^{j\omega}}{\omega^2} - \frac{e^{-j\omega}}{\omega^2} = 2\left(\frac{1-\cos(\omega)}{\omega^2}\right) \\
&= 4\left(\frac{\sin^2(\omega/2)}{\omega^2}\right) = \frac{\sin^2(\omega/2)}{(\omega/2)^2} \\
&= \operatorname{sinc}^2\left(\frac{\omega}{2\pi}\right)\,,
\end{aligned}$$

providing a proof of Equation 40 when $a = 1$. With these two expressions we can compute the power spectral density of $y(t)$ as the convolution. We find

$$\begin{aligned}
\Psi_y(\omega) &= \frac{\pi}{2}\int_{-\infty}^{\infty}\Psi_x(\xi-\omega)\left(\delta(\xi-\omega_0)+\delta(\xi+\omega_0)\right)d\xi \\
&= \frac{\pi}{2}\left(\Psi_x(\omega-\omega_0)+\Psi_x(\omega+\omega_0)\right) \\
&= \frac{\pi}{2}\left(\operatorname{sinc}^2\left(\frac{\omega-\omega_0}{2\pi}\right)+\operatorname{sinc}^2\left(\frac{\omega+\omega_0}{2\pi}\right)\right)\,.
\end{aligned}$$

Problem 3.21 (the autocorrelation function for an integral of $\cos(\cdot)$)

When $x(t) = \cos(t + \theta)$ we find $\psi_y(t, s)$ from its definition the following

$$\begin{aligned}
\psi_y(t, s) &= E\langle y(t)y(s)\rangle \\
&= E\langle \int_0^t x(u)du \int_0^s x(v)dv\rangle \\
&= \int_0^t \int_0^s E\langle x(u)x(v)\rangle dvdu\,.
\end{aligned}$$

From the given definition of $x(t)$ (and the product of cosign identity Equation 30) we now see that the expectation in the integrand becomes

$$\begin{aligned}
E\langle x(u)x(v)\rangle &= E\langle \cos(u + \theta)\cos(v + \theta)\rangle \\
&= \frac{1}{2}E\langle \cos(u - v)\rangle + \frac{1}{2}E\langle \cos(u + v + 2\theta)\rangle \\
&= \frac{1}{2}\cos(u - v) + \frac{1}{2}\left(\frac{1}{2\pi}\int_0^{2\pi} \cos(u + v + 2\theta)d\theta\right) \\
&= \frac{1}{2}\cos(u - v) + \frac{1}{8\pi}\sin(u + v + 2\theta)|_0^{2\pi} \\
&= \frac{1}{2}\cos(u - v)\,.
\end{aligned}$$

Thus we see that $\psi_y(t, s)$ is given by

$$\begin{aligned}
\psi_y(t, s) &= \int_0^t \int_0^s \frac{1}{2}\cos(u - v)dvdu \\
&= \frac{1}{2}\int_0^t - \sin(u - v)|_{v=0}^s\, du \\
&= -\frac{1}{2}\int_0^t (\sin(u - s) - \sin(u))du \\
&= -\frac{1}{2}(-\cos(u - s) + \cos(u)|_0^t \\
&= \frac{1}{2}\cos(t - s) - \frac{1}{2}\cos(s) - \frac{1}{2}\cos(t) + \frac{1}{2}\,.
\end{aligned}$$

As an alternative way to work this problem, in addition to the above method, since we explicitly know the functional form form $x(t)$ we can directly integrate it to obtain the function $y(t)$. We find

$$\begin{aligned}
y(t) &= \int_0^t x(u)du = \int_0^t \cos(u + \theta)du \\
&= \sin(u + \theta)|_0^t \\
&= \sin(t + \theta) - \sin(\theta)\,.
\end{aligned}$$

Note that $y(t)$ is a zero mean sequence when averaging over all possible values of θ. Now to compute $\psi_y(t,s)$ we have

$$\begin{aligned}
\psi_y(t,s) &= E\langle y(t)y(s)\rangle \\
&= \frac{1}{2\pi}\int_0^{2\pi}(\sin(t+\theta)-\sin(\theta))(\sin(s+\theta)-\sin(\theta))d\theta \\
&+ \frac{1}{2\pi}\int_0^{2\pi}\sin(t+\theta)\sin(s+\theta)d\theta \\
&- \frac{1}{2\pi}\int_0^{2\pi}\sin(\theta)\sin(t+\theta)d\theta - \frac{1}{2\pi}\int_0^{2\pi}\sin(\theta)\sin(s+\theta)d\theta \\
&+ \frac{1}{2\pi}\int_0^{2\pi}\sin(\theta)^2 d\theta\,.
\end{aligned}$$

Using the product of sines identity given by

$$\sin(\theta_1)\sin(\theta_2) = \frac{1}{2}(\cos(\theta_1-\theta_2)-\sin(\theta_1+\theta_2))\,, \tag{43}$$

we can evaluate these integrals. Using Mathematical (see `prob_3_21.nb`) we find

$$\psi_y(t,s) = \frac{1}{2} + \frac{1}{2}\cos(s-t) - \frac{1}{2}\cos(s) - \frac{1}{2}\cos(t)\,,$$

the same expression as before.

Problem 3.22 (possible autocorrelation functions)

To study if the given expressions are autocorrelation functions we will simply consider the required properties of autocorrelation functions given on page 21. For the proposed autocorrelation functions given by $\psi_1\psi_2$, $\psi_1+\psi_2$, and $\psi_1 \star \psi_2$ the answer is yes since each has a maximum at the origin, is even, and has a nonnegative Fourier transform whenever the individual ψ_i functions do. For the expression $\psi_1 - \psi_2$ it is unclear whether this expression would have a nonnegative Fourier transform as the sign of the Fourier transform of this expression would depend on the magnitude of the Fourier transform of each individual autocorrelation functions.

Problem 3.23 (more possible autocorrelation functions)

Part (a): In a similar way as in Problem 3.22 all of the required autocorrelation properties hold for $f^2(t) + g(t)$ to be an autocorrelation function.

Part (b): In a similar way as in Problem 3.22 Part (c) this expression may or may not be an autocorrelation function.

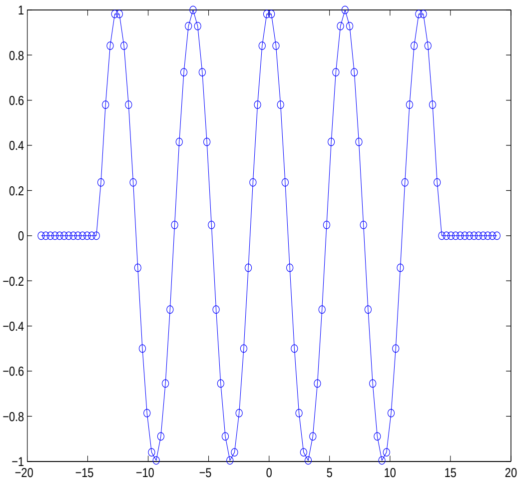

Figure 2: A plot of the function $w(\tau)$ given in Part (d) of Problem 3.23.

Part (c): If $x(t)$ is strictly stationary then *all* of its statistics are invariant of the time origin. As in the expression $x^2(t) + 2x(t-1)$ each term is strictly stationary then I would guess the entire expression is strictly stationary.

Part (d): The function $w(\tau)$ is symmetrical and has a positive (or zero Fourier transform) but $w(\tau)$ has multiple maximum, see Figure 2 and so it cannot be an autocorrelation function. This figure is plotted using the MATLAB script `prob_3_23_d.m`.

Part (e): Once the random value of α is drawn the functional form for $y(t)$ is simply a multiple of that of $x(t)$ and would also be ergodic.

Problem 3.24 (possible autocorrelation functions)

Part (a), (b): These are valid autocorrelation functions.

Part (c): The given function, $\Gamma(t)$, is related to the **rectangle function** defined by

$$\text{rect}(\tau) = \begin{cases} 0 & |\tau| > \frac{1}{2} \\ \frac{1}{2} & |\tau| = \frac{1}{2} \\ 1 & |\tau| < \frac{1}{2} \end{cases} , \tag{44}$$

as $\Gamma(t) = \text{rect}(\frac{1}{2}t)$. This rectangle function has a Fourier transform given by

$$\int_{-\infty}^{\infty} \text{rect}(a\tau)e^{j\omega\tau}d\tau = \frac{1}{|a|}\text{sinc}(\frac{\omega}{2\pi a}) . \tag{45}$$

this later expression is non-positive and therefore $\Gamma(t)$ cannot be an autocorrelation function.

Part (d): This function is not even and therefore cannot be an autocorrelation function.

Part (e): Recall that when the autocorrelation function $\psi_x(\tau) = \sigma^2 e^{-\alpha|\tau|}$, we have a power spectral density of $\Psi_x(\omega) = \frac{2\sigma^2\alpha}{\omega^2+\alpha^2}$, so that the Fourier transform of the proposed autocorrelation function in this case is

$$\frac{2(3/2)}{\omega^2+1} - \frac{2(1)(2)}{\omega^2+4} = \frac{3}{\omega^2+1} - \frac{4}{\omega^2+4}\,.$$

This expression is negative when $\omega = 0$, thus the proposed function $\frac{3}{2}e^{-|\tau|} - e^{-2|\tau|}$ cannot be an autocorrelation function.

Part (f): From Part (e) above this proposed autocorrelation function would have a Fourier transform that is given by

$$\frac{2(2)(2)}{\omega^2+4} - \frac{2(1)(1)}{\omega^2+1} = 2\left(\frac{3\omega^2}{(\omega^2+1)(\omega^2+4)}\right)\,,$$

which is nonnegative, so this expression *is* a valid autocorrelation function.

Problem 3.25 (some definitions)

Part (a): Wide-sense stationary is a less restrictive condition than full stationary in that it only requires the first *two* statistics of our process to be time independent (stationary requires *all* statistics to be time independent).

Problem 3.29 (the autocorrelation function for a driven differential equation)

Part (a): For the given linear dynamic system a fundamental solution $\Phi(t, t_0)$ is given explicitly by $\Phi(t, t_0) = e^{-(t-t_0)}$ so the full solution for the unknown $x(t)$ in terms of the random forcing $n(t)$ is given by using Equation 1 to get

$$x(t) = e^{-(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{-(t-\tau)}n(\tau)d\tau\,. \tag{46}$$

Letting our initial time be $t_0 = -\infty$ we obtain

$$x(t) = \int_{-\infty}^{t} e^{-(t-\tau)}n(\tau)d\tau = e^{-t}\int_{-\infty}^{t} e^{\tau}n(\tau)d\tau\,.$$

With this expression, the autocorrelation function $\psi_x(t_1, t_2)$ is given by

$$\begin{aligned}
\psi_x(t_1, t_2) &= E\left\langle \left(e^{-t_1}\int_{-\infty}^{t_1} e^{u}n(u)du\right)\left(e^{-t_2}\int_{-\infty}^{t_2} e^{v}n(v)dv\right)\right\rangle \\
&= e^{-(t_1+t_2)}\int_{-\infty}^{t_1}\int_{-\infty}^{t_2} e^{u+v}E\langle n(u)n(v)\rangle dvdu\,.
\end{aligned}$$

Since $E\langle n(u)n(v)\rangle = 2\pi\delta(u-v)$ if we assume $n(t)$ has a power spectral density of 2π. With this the above becomes

$$2\pi e^{-(t_1+t_2)} \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} e^{u+v} \delta(u-v) dv du \,.$$

Without loss of generality assume that $t_1 < t_2$ and the above becomes

$$\begin{aligned} 2\pi e^{-(t_1+t_2)} \int_{-\infty}^{t_1} e^{2u} du &= 2\pi e^{-(t_1+t_2)} \left. \frac{e^{2u}}{2} \right|_{-\infty}^{t_1} \\ &= \pi e^{-(t_1+t_2)} e^{2t_1} = \pi e^{-t_2+t_1} = \pi e^{-(t_2-t_1)} \,. \end{aligned}$$

If we had assumed that $t_1 > t_2$ we would have found that $\psi_x(t_1,t_2) = \pi e^{-(t_1-t_2)}$. Thus combining these two we have show that

$$\psi_x(t_1,t_2) = \pi e^{-|t_1-t_2|} \,, \tag{47}$$

and $x(t)$ *is* wide-sense stationary.

Part (b): If the functional form of the right hand side of our differential equation changes we will need to recompute the expression for $\psi_x(t_1,t_2)$. Taking $x(t_0) = 0$ and with the new right hand side Equation 46 now gives a solution for $x(t)$ of

$$x(t) = e^{-t} \int_0^t e^{\tau} n(\tau) d\tau \,,$$

note the lower limit of the integral of our noise term is now 0. From this expression the autocorrelation function then becomes

$$\begin{aligned} \psi_x(t_1,t_2) &= E\left\langle \left(e^{-t_1} \int_0^{t_1} e^u n(u) du \right) \left(e^{-t_2} \int_0^{t_2} e^v n(v) dv \right) \right\rangle \\ &= e^{-(t_1+t_2)} \int_0^{t_1} \int_0^{t_2} e^{u+v} E\langle n(u)n(v)\rangle dv du \\ &= e^{-(t_1+t_2)} \int_0^{t_1} \int_0^{t_2} e^{u+v} 2\pi\delta(u-v) dv du \,. \end{aligned}$$

Assume $t_1 < t_2$ and the above becomes

$$\begin{aligned} \psi_x(t_1,t_2) &= 2\pi e^{-(t_1+t_2)} \int_0^{t_1} e^{2u} du = 2\pi e^{-(t_1+t_2)} \left(\left. \frac{e^{2u}}{2} \right|_0^{t_1} \right. \\ &= \pi e^{-(t_1+t_2)} \left(e^{2t_1} - 1 \right) \\ &= \pi (e^{-(t_2-t_1)} - e^{-(t_2+t_1)}) \,. \end{aligned}$$

Considering the case when $t_1 > t_2$ we would find

$$\psi_x(t_1,t_2) = \pi (e^{-(t_1-t_2)} - e^{-(t_2+t_1)}) \,.$$

When we combine these two results we find

$$\psi_x(t_1,t_2) = \pi (e^{-|t_1-t_2|} - e^{-(t_2+t_1)}) \,.$$

Note that in this case $x(t)$ is *not* wide-sense stationary. This is a consequent of the fact that our forcing function (the right hand side) was "switched on" at $t = 0$ rather than having been operating from $t = -\infty$ until the present time t. The algebra for this problem is verified in the Mathematica file `prob_3_29.nb`.

Part (c): Note that in general when $y(t) = \int_0^t x(\tau)d\tau$ we can evaluate the cross-correlation function $\psi_{xy}(t_1, t_2)$ directly from the autocorrelation function, $\psi_x(t_1, t_2)$, for $x(t)$. Specifically we find

$$\begin{aligned}
\psi_{xy}(t_1, t_2) &= E\langle x(t_1)y(t_2)\rangle \\
&= E\left\langle x(t_1)\left(\int_0^{t_2} x(\tau)d\tau\right)\right\rangle \\
&= \int_0^{t_2} E\langle x(t_1)x(\tau)\rangle d\tau \\
&= \int_0^{t_2} \psi_x(t_1, \tau)d\tau\,.
\end{aligned}$$

Since we have calculated ψ_x for both of the systems above we can use these results and the identity above to evaluate ψ_{xy}. For the system in Part (a) we have when $t_1 < t_2$ that

$$\begin{aligned}
\psi_{xy}(t_1, t_2) &= \int_0^{t_2} \psi_x(t_1, \tau)d\tau = \int_0^{t_2} \pi e^{-|t_1-\tau|}d\tau \\
&= \pi\int_0^{t_1} e^{-(t_1-\tau)}d\tau + \pi\int_{t_1}^{t_2} e^{+(t_1-\tau)}d\tau \\
&= 2\pi - \pi e^{-t_1} - \pi e^{t_1-t_2}\,.
\end{aligned}$$

If $t_2 < t_1$ then we have

$$\begin{aligned}
\psi_{xy}(t_1, t_2) &= \pi\int_0^{t_2} e^{-|t_1-\tau|}d\tau \\
&= \pi\int_0^{t_2} e^{-(t_1-\tau)}d\tau = \pi e^{t_2-t_1} - \pi e^{-t_1}\,.
\end{aligned}$$

Thus combining these two results we find

$$\psi_{xy}(t_1, t_2) = \begin{cases} 2\pi - \pi e^{-t_1} - \pi e^{t_1-t_2} & t_1 < t_2 \\ \pi e^{t_2-t_1} - \pi e^{-t_1} & t_1 > t_2 \end{cases}\,.$$

While in the second case (Part (b)) since $\psi_x(t_1, t_2)$ has a term of the form $\pi e^{|t_1-t_2|}$ which is exactly the same as the first case in Part (a) we only need to evaluate

$$\begin{aligned}
-\pi\int_0^{t_2} e^{-(t_1+\tau)}d\tau &= \pi e^{-t_1}\left(e^{-\tau}\Big|_0^{t_2}\right. \\
&= \pi(e^{-(t_1+t_2)} - e^{-t_1})\,.
\end{aligned}$$

Thus we finally obtain for Part (b)

$$\psi_{xy}(t_1, t_2) = \begin{cases} 2\pi - 2\pi e^{-t_1} - \pi e^{t_1-t_2} + \pi e^{-(t_1+t_2)} & t_1 < t_2 \\ -2\pi e^{-t_1} + \pi e^{t_2-t_1} + \pi e^{-(t_1+t_2)} & t_1 > t_2 \end{cases}\,.$$

Part (d): To predict $x(t+\alpha)$ using $x(t)$ using an estimate $\hat{x}(t+\alpha) = ax(t)$ we will minimize the mean-square prediction error $E\langle[\hat{x}(t+\alpha) - x(t+\alpha)]^2\rangle$ as a function of a. For the given linear form for $\hat{x}(t+\alpha)$ the expression we will minimize for a is given by

$$\begin{aligned} F(a) &\equiv E\langle[ax(t) - x(t+\alpha)]^2\rangle \\ & E\langle a^2x^2(t) - 2ax(t)x(t+\alpha) + x^2(t+\alpha)\rangle \\ &= a^2\psi_x(t,t) - 2a\psi_x(t,t+\alpha) + \psi_x(t+\alpha,t+\alpha)\,. \end{aligned}$$

Since we are considering the functional form for $\psi_x(t_1,t_2)$ derived for in Part (a) above we know that $\psi_x(t_1,t_2) = \pi e^{-|t_1-t_2|}$ so

$$\psi_x(t,t) = \pi = \psi_x(t+\alpha,t+\alpha) \quad \text{and} \quad \psi_x(t,t+\alpha) = \pi e^{-|\alpha|} = \pi e^{-\alpha}\,,$$

since $\alpha > 0$. Thus the function $F(a)$ then becomes

$$F(a) = \pi a^2 - 2\pi a e^{-\alpha} + \pi\,.$$

To find the minimum of this expression we take the derivative of F with respect to a, set the resulting expression equal to zero and solve for a. We find

$$F'(a) = 2\pi a - 2\pi e^{-\alpha} = 0 \quad \text{so} \quad a = e^{-\alpha}\,.$$

Thus to optimally predict $x(t+\alpha)$ given $x(t)$ on should use the prediction $\hat{x}(t+\alpha)$ given by

$$\hat{x}(t+\alpha) = e^{-\alpha}x(t)\,. \tag{48}$$

Problem 3.30 (a random initial condition)

Part (a): This equation is similar to Problem 3.29 Part (b) but now $x(t_0) = x_0$ is non-zero and random rather than deterministic. For this given linear system we have a solution still given by Equation 46

$$x(t) = e^{-t}x_0 + e^{-t}\int_0^t e^{\tau}n(\tau)d\tau = e^{-t}x_0 + I(t)\,,$$

where we have defined the function $I(t) \equiv e^{-t}\int_0^t e^{\tau}n(\tau)d\tau$. To compute the autocorrelation function $\psi_x(t_1,t_2)$ we use its definition to find

$$\begin{aligned} \psi_x(t_1,t_2) &= E\langle x(t_1)x(t_2)\rangle \\ &= E\langle(e^{-t_1}x_0 + I(t_1))(e^{-t_2}x_0 + I(t_2))\rangle \\ &= e^{-(t_1+t_2)}E\langle x_0^2\rangle + e^{-t_1}E\langle x_0 I(t_2)\rangle + e^{-t_2}E\langle I(t_1)x_0\rangle + E\langle I(t_1)I(t_2)\rangle \\ &= \sigma^2 e^{-(t_1+t_2)} + E\langle I(t_1)I(t_2)\rangle\,, \end{aligned}$$

since the middle two terms are zero and we are told that x_0 is zero mean with a variance σ^2. The expression $E\langle I(t_1)I(t_2)\rangle$ was computed in Problem 3.29 b. Thus we find

$$\psi_x(t_1,t_2) = \sigma^2 e^{-(t_1+t_2)} + \pi(e^{-|t_1-t_2|} - e^{-(t_1+t_2)})\,.$$

Part (b): If we take $\sigma^2 = \sigma_0^2 = \pi$ then the autocorrelation function becomes

$$\psi_x(t_1, t_2) = \pi e^{-|t_1 - t_2|} \,,$$

so in this case $x(t)$ *is* wide-sense stationary (WSS).

Part (c): Now $x(t)$ will be wise-sense stationary since if the white noise is turned on at $t = -\infty$ because the initial condition x_0 will have no effect on the solution $x(t)$ at current times. This is because the effect of the initial condition at the time t_0 from Equation 46 is given by

$$x_0 e^{-t+t_0} \,,$$

and if $t_0 \to -\infty$ the contribution of this term vanishes no matter what the statistical properties of x_0 are.

Problem 3.31 (the mean and covariance for the given dynamical system)

From the given dynamical system

$$\dot{x}(t) = F(t)x(t) + w(t) \quad \text{with} \quad x(a) = x_a \,,$$

The full solution to this equation can be obtained symbolically given the fundamental solution matrix $\Phi(t, t_0)$ as

$$x(t) = \Phi(t, a)x(a) + \int_a^t \Phi(t, \tau)w(\tau)d\tau \,,$$

then taking the expectation of this expression gives an equation for the mean $m(t)$

$$m(t) = E\langle x(t)\rangle = \Phi(t, a)E\langle x(a)\rangle + \int_a^t \Phi(t, \tau)E\langle w(\tau)\rangle d\tau = 0 \,,$$

since $E\langle w(\tau)\rangle = 0$, and $E\langle x(a)\rangle = E\langle x_a\rangle = 0$ as we assume that x_a is zero mean.

The covariance matrix $P(t)$ for this system is computed as

$$\begin{aligned}
P(t) &= E\langle (x(t) - m(t))(x(t) - m(t))^T\rangle \\
&= E\left\langle \left(\Phi(t, a)x_a + \int_a^t \Phi(t, \tau)w(\tau)d\tau\right)\left(\Phi(t, a)x_a + \int_a^t \Phi(t, \tau)w(\tau)d\tau\right)^T \right\rangle \\
&= \Phi(t, a)E\langle x_a x_a^T\rangle \Phi(t, a)^T \\
&+ \Phi(t, a)E\left\langle x_a \left(\int_a^t \Phi(t, \tau)w(\tau)d\tau\right)^T \right\rangle + E\left\langle \left(\int_a^t \Phi(t, \tau)w(\tau)d\tau\right) x_a^T \right\rangle \Phi(t, a)^T \\
&+ E\left\langle \left(\int_a^t \Phi(t, \tau)w(\tau)d\tau\right)\left(\int_a^t \Phi(t, \tau)w(\tau)d\tau\right)^T \right\rangle \\
&= \Phi(t, a)P_a\Phi(t, a)^T \\
&+ \Phi(t, a)\int_a^t E\langle x_a w^T(\tau)\rangle \Phi(t, \tau)^T d\tau + \left(\int_a^t \Phi(t, \tau)E\langle w(\tau)x_a^T\rangle d\tau\right)\Phi(t, a)^T \\
&+ \int_{u=a}^t \int_{v=a}^t \Phi(t, u)E\langle w(u)w(v)^T\rangle \Phi(t, v)^T dv du \,.
\end{aligned}$$

Now as $E\langle x_a w^T \rangle = 0$ the middle two terms above vanish. Also $E\langle w(u)w(v)^T\rangle = Q(u)\delta(u-v)$ so the fourth term becomes

$$\int_{u=a}^{t} \Phi(t,u)Q(u)\Phi(t,u)^T du\,.$$

With these two simplifications the covariance $P(t)$ for $x(t)$ is given by

$$P(t) = \Phi(t,a)P_a\Phi(t,a)^T + \int_{u=a}^{t} \Phi(t,u)Q(u)\Phi(t,u)^T du\,.$$

Part (b): A differential equation for $P(t)$ is given by taking the derivative of the above expression for $P(t)$ with respect to t. We find

$$\begin{aligned}
\frac{dP}{dt} &= \frac{d\Phi(t,a)}{dt}P_a\Phi(t,a)^T + \Phi(t,a)P_a\frac{d\Phi(t,a)^T}{dt} \\
&+ \Phi(t,t)Q(t)\Phi(t,t)^T \\
&+ \int_{u=a}^{t} \frac{d\Phi(t,u)}{dt}Q(u)\Phi(t,u)^T du + \int_{u=a}^{t} \Phi(t,u)Q(u)\frac{d\Phi(t,u)^T}{dt}du\,.
\end{aligned}$$

Recall that the fundamental solution $\Phi(t,a)$ satisfies the following $\frac{d\Phi(t,a)}{dt} = F(t)\Phi(t,a)$ and that $\Phi(t,t) = I$ with I the identity matrix. With these expressions the right-hand-side of $\frac{dP}{dt}$ then becomes

$$\begin{aligned}
\frac{dP}{dt} &= F(t)\Phi(t,a)P_a\Phi(t,a)^T + \Phi(t,a)P_a\Phi(t,a)^T F^T(t) + Q(t) \\
&+ \int_{u=a}^{t} F(t)\Phi(t,u)Q(u)\Phi(t,u)^T du + \int_{u=a}^{t} \Phi(t,u)Q(u)\Phi(t,u)^T F(t)^T du \\
&= F(t)\left[\Phi(t,a)P_a\Phi(t,a)^T + \int_{u=a}^{t} \Phi(t,u)Q(u)\Phi(t,u)^T du\right] \\
&+ \left[\Phi(t,a)P_a\Phi(t,a)^T + \int_{u=a}^{t} \Phi(t,u)Q(u)\Phi(t,u)^T du\right] F(t)^T + Q(t) \\
&= F(t)P(t) + P(t)F(t)^T + Q(t)\,,
\end{aligned}$$

as a differential equation for $P(t)$.

Problem 3.32 (examples at computing the covariance matrix $P(t)$)

To find the steady state value for $P(t)$ i.e. $P(\infty)$ we can either compute the fundamental solutions, $\Phi(t,\tau)$, for the given systems and use the "direct formulation" for the time value of $P(t)$ i.e.

$$P(t) = \Phi(t,t_0)P(t_0)\Phi^T(t,t_0) + \int_{t_0}^{t} \Phi(t,\tau)G(\tau)QG^T(\tau)\Phi^T(t,\tau)d\tau\,. \tag{49}$$

or use the "differential equation formulation" for $P(t)$ given by

$$\frac{dP}{dt} = F(t)P(t) + P(t)F^T(t) + G(t)QG^T(t)\,. \tag{50}$$

Since this later equation involves only the expressions F, G, and Q which we are given directly from the continuous time state *definition* repeated here for convenience

$$\dot{x} = F(t)x(t) + G(t)w(t) \quad (51)$$
$$E\langle w(t)\rangle = 0 \quad (52)$$
$$E\langle w(t_1)w^T(t_2)\rangle = Q(t_1,t_2)\delta(t_1-t_2)\,. \quad (53)$$

Part (a): For this specific linear dynamic system we have Equation 50 given by

$$\begin{aligned}\dot{P}(t) &= \begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix} P(t) + P(t)\begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix}^T + \begin{bmatrix} 1 \\ 1 \end{bmatrix} 1 \begin{bmatrix} 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix} P(t) + P(t)\begin{bmatrix} -1 & -1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},\end{aligned}$$

In terms of components of the matrix $P(t)$ we would have the following system

$$\begin{bmatrix} \dot{p}_{11}(t) & \dot{p}_{21}(t) \\ \dot{p}_{21}(t) & \dot{p}_{22}(t) \end{bmatrix} = \begin{bmatrix} -p_{11} & -p_{12} \\ -p_{11} & -p_{12} \end{bmatrix} + \begin{bmatrix} -p_{11} & -p_{11} \\ -p_{21} & -p_{21} \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

or

$$\begin{bmatrix} \dot{p}_{11}(t) & \dot{p}_{21}(t) \\ \dot{p}_{21}(t) & \dot{p}_{22}(t) \end{bmatrix} = \begin{bmatrix} -2p_{11}+1 & -p_{21}-p_{11}+1 \\ -p_{11}-p_{21}+1 & -2p_{21}+1 \end{bmatrix}.$$

Note that we have enforced the symmetry of $P(t)$ by explicitly taking $p_{12} = p_{21}$. To solve the $(1,1)$ component in the matrix above we need to consider the differential equation given by

$$\dot{p}_{11}(t) = -2p_{11}(t) + 1 \quad \text{with} \quad p_{11}(0) = 1\,.$$

which has a solution

$$p_{11}(t) = \frac{1}{2}(1 + e^{-2t})\,.$$

Using this then $p_{21}(t)$ must satisfy

$$\begin{aligned}\dot{p}_{21}(t) &= -p_{21}(t) - p_{11} + 1 \\ &= -p_{21}(t) + \frac{1}{2} - \frac{1}{2}e^{-2t},\end{aligned}$$

with an initial condition of $p_{21}(0) = 0$. Solving this we find a solution given by

$$p_{21}(t) = \frac{1}{2} - e^{-t} + \frac{1}{2}e^{-2t}\,.$$

Finally the function $p_{22}(t)$ must solve

$$\begin{aligned}\dot{p}_{22}(t) &= -2p_{21}(t) + 1 \\ &= 2e^{-t} - e^{-2t},\end{aligned}$$

with the initial condition that $p_{22}(0) = 1$. Solving this we conclude that

$$p_{22}(t) = \frac{5}{2} - 2e^{-t} + \frac{1}{2}e^{-2t}\,.$$

The time-dependent matrix $P(t)$ is then given by placing all of these function in a matrix form. All of the functions considered above give

$$P(\infty) = \begin{bmatrix} p_{11}(\infty) & p_{21}(\infty) \\ p_{21}(\infty) & p_{22}(\infty) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{5}{2} \end{bmatrix} .$$

Part (b): For the given linear dynamic system, the differential equations satisfied by the covariance matrix $P(t)$ become (when we recognized that $F = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ and $G = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$)

$$\begin{aligned} \dot{P}(t) &= F(t)P(t) + P(t)F(t)^T + G(t)QG^T(t) \\ &= \begin{bmatrix} -p_{11} & -p_{21} \\ -p_{21} & -p_{22} \end{bmatrix} + \begin{bmatrix} -p_{11} & -p_{21} \\ -p_{21} & -p_{22} \end{bmatrix} + \begin{bmatrix} 25 & 5 \\ 5 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -2p_{11} + 25 & -2p_{21} + 5 \\ -2p_{21} + 5 & -2p_{22} + 1 \end{bmatrix} . \end{aligned}$$

Solving for the $(1,1)$ element we have the differential equation given by

$$\dot{p}_{11}(t) = -2p_{11}(t) + 25 \quad \text{with} \quad p_{11}(0) = 1 .$$

This has a solution given by

$$p_{11}(t) = \frac{1}{2}e^{-2t}(-23 + 25e^{2t}) .$$

Solving for the $(2,2)$ element we have the differential equation

$$\dot{p}_{22}(t) = -2p_{22}(t) + 1 \quad \text{with} \quad p_{22}(0) = 1 .$$

This has a solution given by

$$p_{22}(t) = \frac{1}{2}e^{-2t}(1 + e^{2t}) .$$

Finally, equation for the $(1,2)$ element (equivalently the $(2,1)$ element) when solved gives

$$p_{21}(t) = \frac{5}{2}e^{-2t}(-1 + e^{2t}) .$$

All of the functions considered above give

$$P(\infty) = \begin{bmatrix} p_{11}(\infty) & p_{21}(\infty) \\ p_{21}(\infty) & p_{22}(\infty) \end{bmatrix} = \begin{bmatrix} \frac{25}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} ,$$

for the steady-state covariance matrix. The algebra for solving these differential equation is given in the Mathematica file `prob_3_32.nb`.

Problem 3.33 (an example computing the discrete covariance matrix P_k)

The discrete covariance propagation equation is given by

$$P_k = \Phi_{k-1}P_{k-1}\Phi_{k-1}^T + G_{k-1}Q_{k-1}G_{k-1}^T , \tag{54}$$

which for this discrete linear system is given by

$$P_k = \begin{bmatrix} 0 & 1/2 \\ -1/2 & 2 \end{bmatrix} P_{k-1} \begin{bmatrix} 0 & -1/2 \\ 1/2 & 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} 1 \begin{bmatrix} 1 & 1 \end{bmatrix}$$

Define $P_k = \begin{bmatrix} p_{11}(k) & p_{12}(k) \\ p_{12}(k) & p_{22}(k) \end{bmatrix}$ and we obtain the set of matrix equations given by

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

As a linear system for the unknown functions $p_{11}(k)$, $p_{12}(k)$, and $p_{22}(k)$ we can write it as

$$\begin{bmatrix} p_{11}(k+1) \\ p_{12}(k+1) \\ p_{22}(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1/4 \\ 0 & -1/4 & 1 \\ 1/4 & -2 & 4 \end{bmatrix} \begin{bmatrix} p_{11}(k) \\ p_{12}(k) \\ p_{22}(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

This is a linear vector difference equation and can be solved by methods discussed in [1]. Using Rather than carry out these calculations by hand in the Mathematica file `prob_3_33.nb` their solution is obtained symbolically.

Problem 3.34 (the steady-state covariance matrix for the harmonic oscillator)

Example 3.4 is a linear dynamic system given by

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\zeta\omega_n \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} a \\ b - 2a\zeta\omega_n \end{bmatrix} w(t) .$$

Then the equation for the covariance of these state $x(t)$ or $P(t)$ is given by

$$\begin{aligned} \frac{dP}{dt} &= F(t)P(t) + P(t)F(t)^T + G(t)Q(t)G(t)^T \\ &= \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & -2\zeta\omega_n \end{bmatrix} \begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix} + \begin{bmatrix} p_{11}(t) & p_{12}(t) \\ p_{12}(t) & p_{22}(t) \end{bmatrix} \begin{bmatrix} 0 & -\omega_n^2 \\ 1 & -2\zeta\omega_n \end{bmatrix} \\ &+ \begin{bmatrix} a \\ b - 2a\zeta\omega_n \end{bmatrix} \begin{bmatrix} a & b - 2a\zeta\omega_n \end{bmatrix} . \end{aligned}$$

Since we are only looking for the steady-state value of P i.e. $P(\infty)$ let $t \to \infty$ in the above to get a linear system for the limiting values $p_{11}(\infty)$, $p_{12}(\infty)$, and $p_{22}(\infty)$. The remaining portions of this exercise are worked just like Example 3.9 from the book.

Problem 3.35 (a negative solution to the steady-state Ricatti equation)

Consider the scalar case suggested where $F = Q = G = 1$ and we find that the continuous-time steady state algebraic equation becomes

$$0 = 1P(+\infty) + P(+\infty) + 1 \Rightarrow P(\infty) = -\frac{1}{2},$$

which is a negative solution in contradiction to the definition of $P(\infty)$.

Problem 3.36 (no solution to the steady-state Ricatti equation)

Consider the given discrete-time steady-state algebraic equation, specified to the scalar case. Then assuming a solution for P_∞ exists this equation gives

$$P_\infty = P_\infty + 1\,,$$

which after canceling P_∞ on both sides implies given the contradiction $0 = 1$. This implies that no solution exists.

Problem 3.37 (computing the discrete-time covariance matrix)

From the given discrete time process model, by taking expectations of both sides we have $E\langle x_k\rangle = -2E\langle x_{k-1}\rangle$, which has a solution given by $E\langle x_k\rangle = E\langle x_0\rangle(-2)^k$, for some constant $E\langle x_0\rangle$. If $E\langle x_0\rangle = 0$, then the expectation of the state x_k is also zero. The discrete covariance of the state is given by solving the difference equation

$$P_k = \Phi_{k-1}P_{k-1}\Phi_{k-1}^T + Q_{k-1}\,,$$

for P_k. For the given discrete-time system this becomes

$$P_k = 4P_{k-1} + 1\,.$$

The solution to this difference equation is given by (see the Mathematica file `prob_3_37.nb`),

$$P_k = \frac{1}{3}(-1 + 4^k + 3\,P_0\,4^k)\,.$$

If we take $P_0 = 1$ then this equation becomes

$$P_k = \frac{1}{3}(-1 + 4^{k+1})\,.$$

The steady-state value of this covariance is $P_\infty = \infty$.

Problem 3.38 (computing the time-varying covariance matrix)

For a continuous linear system like this one the differential equation satisfied by the covariance of $x(t)$ or $P(t)$ is given by the solution of the following differential equation

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G(t)QG^T(t)\,.$$

For this scalar problem we have $F(t) = -2$, $G(t) = 1$, and $Q(t_1, t_2) = e^{-|t_2-t_1|}\delta(t_1 - t_2)$, becomes

$$\dot{P}(t) = -2P - 2P + 1 = -4P + 1\,.$$

Solving this equation for $P(t)$ gives (see the Mathematica file `prob_3_38.nb`)

$$P(t) = \frac{1}{4}e^{-4t}(-1 + 4P(0) + e^{4t}) .$$

If we assume that $P(0) = 1$ then the above becomes

$$P(t) = \frac{1}{4}e^{-4t}(3 + e^{4t}) = \frac{1}{4}(3e^{-4t} + 1) .$$

The steady-state value of the above expression is given by $P(\infty) = \frac{1}{4}$.

Problem 3.39 (linear prediction of $x(t + \alpha)$ using the values of $x(s)$ for $s < t$)

Part (a): We assume that our predictor in this case will have a mathematical form given by

$$\hat{x}(t + \alpha) = \int_{-\infty}^{t} a(v)x(v)dv ,$$

for some as yet undetermined function $a(v)$. With this expression we seek to minimize the prediction error when using this function $a(\cdot)$. That is we seek to minimize $F(a) \equiv E\langle|\hat{x}(t + \alpha) - x(t + \alpha)|^2\rangle$ which can be expressed as

$$F(a) = E\langle\left|\int_{-\infty}^{t} a(v)x(v)dv - x(t + \alpha)\right|^2\rangle ,$$

which when we expand out the arguments inside the expectation becomes

$$E\left\langle\left(\int_{u=-\infty}^{t}\int_{v=-\infty}^{t} a(u)a(v)x(u)x(v)dudv \;-\; 2\int_{-\infty}^{t} a(v)x(v)x(t + \alpha)ds + x^2(t + \alpha)\right)\right\rangle ,$$

or passing the expectation inside the integrals above we find $F(a)$ becomes

$$\begin{aligned} F(a) &= \int_{u=-\infty}^{t}\int_{v=-\infty}^{t} a(u)a(v)E\langle x(u)x(v)\rangle dudv \\ &- 2\int_{-\infty}^{t} a(v)E\langle x(v)x(t + \alpha)\rangle ds + E\langle x^2(t + \alpha)\rangle . \end{aligned}$$

Using the given autocorrelation function for $x(t)$ we see that these expectations take the values

$$\begin{aligned} E\langle x(u)x(v)\rangle &= e^{-c|u-v|} \\ E\langle x(v)x(t + \alpha)\rangle &= e^{-c|t+\alpha-v|} \\ E\langle x^2(t + \alpha)\rangle &= 1 , \end{aligned}$$

so that the above becomes

$$F(a) = \int_{u=-\infty}^{t}\int_{v=-\infty}^{t} a(u)a(v)e^{-c|u-v|}dudv - 2\int_{-\infty}^{t} a(v)e^{-c|t+\alpha-v|}dv + 1 .$$

To optimize $F(\cdot)$ as a function of the unknown function $a(\cdot)$ using the calculus of variations we compute $\delta F = F(a + \delta a) - F(a)$, where δa is a "small" functional perturbation of the function a. We find

$$
\begin{aligned}
F(a+\delta a) - F(a) &= \int_{u=-\infty}^{t} \int_{v=-\infty}^{t} (a(u) + \delta a(u))(a(v) + \delta a(v)) e^{-c|u-v|} du dv \\
&- 2 \int_{v=-\infty}^{t} (a(v) + \delta a(v)) e^{-c|t+\alpha-v|} dv \\
&- \int_{u=-\infty}^{t} \int_{v=-\infty}^{t} a(u) a(v) e^{-c|u-v|} du dv + 2 \int_{v=-\infty}^{t} a(v) e^{-c|t+\alpha-v|} dv \\
&= \int_{u=-\infty}^{t} \int_{v=-\infty}^{t} a(u) \delta a(v) e^{-c|u-v|} du dv \qquad (55) \\
&+ \int_{u=-\infty}^{t} \int_{v=-\infty}^{t} a(v) \delta a(u) e^{-c|u-v|} du dv \qquad (56) \\
&+ \int_{u=-\infty}^{t} \int_{v=-\infty}^{t} \delta a(u) \delta a(v) e^{-c|u-v|} du dv \\
&- 2 \int_{v=-\infty}^{t} \delta a(v) e^{-c|t+\alpha-v|} dv \,.
\end{aligned}
$$

Now the two integrals Equation 55 and 56 are equal and using this the above expression for δF becomes

$$
2 \int_{u=-\infty}^{t} \int_{v=-\infty}^{t} a(u) \delta a(v) e^{-c|u-v|} du dv - 2 \int_{v=-\infty}^{t} \delta a(v) e^{-c|t+\alpha-v|} dv + O(\delta a^2) \,.
$$

Recalling that $t + \alpha > v$ we can drop the absolute value in the exponential of the second term and if we assume that $O(\delta a^2)$ is much smaller than the other two terms, we can ignore it. Then by taking the v integration to the outside we obtain

$$
2 \int_{v=-\infty}^{t} \left[\int_{u=-\infty}^{t} a(u) e^{-c|u-v|} du - e^{-c(t+\alpha-v)} \right] \delta a(v) dv \,.
$$

Now the calculus of variations assumes that at the optimum value for a, the first variation vanishes or $\delta F = 0$. This implies that we must have in argument of the above integrand identically equal to zero or $a(\cdot)$ must satisfy

$$
\int_{u=-\infty}^{t} a(u) e^{-c|u-v|} du - e^{-c(t+\alpha-v)} = 0 \,.
$$

Taking the derivative of this expression with respect to t we then obtain (since $v < t$)

$$
a(t) e^{-c(t-v)} = e^{-c(t+\alpha-v)} \,.
$$

when we solve this for $a(t)$ we find that $a(t)$ is not actually a function of t but is given by

$$
a(t) = e^{-c\alpha} \,, \qquad (57)
$$

so that our estimator becomes

$$
\hat{x}(t + \alpha) = e^{-c\alpha} x(t) \,, \qquad (58)
$$

as we were to show.

Part (b): To find the mean-square error we want to evaluate $F(a)$ at the $a(\cdot)$ we calculated above. We find

$$
\begin{aligned}
F(a) &= E\langle [e^{-c\alpha}x(t) - x(t+\alpha)]^2 \rangle \\
&= E\langle e^{-2c\alpha}x^2(t) - 2e^{-c\alpha}x(t)x(t+\alpha) + x^2(t+\alpha) \rangle \\
&= e^{-2c\alpha} - 2e^{-c\alpha}e^{-c\alpha} + 1 \\
&= 1 - e^{-2c\alpha}\,.
\end{aligned}
$$

Chapter 4: Linear Optimal Filters and Predictors

Notes On The Text

Estimators in Linear Form

For this chapter we will consider an estimator of the unknown state x at the k-th time step to be denoted $\hat{x}_k(+)$, given the k-th measurement z_k, and our previous estimate of x before the measurement (denoted $\hat{x}_k(-)$) of the following linear form

$$\hat{x}_k(+) = K_k^1 \hat{x}_k(-) + \overline{K}_k z_k \,, \tag{59}$$

for some as yet undetermined coefficients K_k^1 and $\overline{K}_k$. The requiring the orthogonality condition that this estimate must satisfy is then that

$$E\langle [x_k - \hat{x}_k(+)] z_i^T \rangle = 0 \quad \text{for} \quad i = 1, 2, \cdots, k-1 \,. \tag{60}$$

Note that this orthogonality condition is stated for the posterior (after measurement) estimate $\hat{x}_k(+)$ but for a recursive filter we expect it to hold for the a-priori (before measurement) estimate $\hat{x}_k(-)$ also. These orthogonality conditions can be simplified to determine conditions on the unknown coefficients K_k^1 and $\overline{K}_k$. From our chosen form for $\hat{x}_k(+)$ from Equation 59 the orthogonality conditions imply

$$E\langle [x_k - K_k^1 \hat{x}_k(-) - \overline{K}_k z_k] z_i^T \rangle = 0 \,.$$

Since our measurement z_k in terms of the true state x_k is given by

$$z_k = H_k x_k + v_k \,, \tag{61}$$

the above expression becomes

$$E\langle [x_k - K_k^1 \hat{x}_k(-) - \overline{K}_k H_k x_k - \overline{K}_k v_k] z_i^T \rangle = 0 \,.$$

Recognizing that the measurement noise v_k is assumed uncorrelated with the measurement z_i we $E\langle v_k z_k^T \rangle = 0$ so this term drops from the orthogonality conditions and we obtain

$$E\langle [x_k - K_k^1 \hat{x}_k(-) - \overline{K}_k H_k x_k] z_i^T \rangle = 0 \,.$$

From this expression we now adding and subtracting $K_k^1 x_k$ to obtain

$$E\langle [x_k - \overline{K}_k H_k x_k - K_k^1 x_k - K_K^1 \hat{x}_k(-) + K_k^1 x_k] z_i^T \rangle = 0 \,,$$

so that by grouping the last two terms we find

$$E\langle [x_k - \overline{K}_k H_k x_k - K_k^1 x_k - K_K^1 (\hat{x}_k(-) - x_k)] z_i^T \rangle = 0 \,.$$

This last term $E\langle (\hat{x}_k(-) - x_k) z_i^T \rangle = 0$ due to the orthogonality condition satisfied by the previous estimate $\hat{x}_k(-)$. Factoring out x_k and applying the expectation to each individual term this becomes

$$(I - \overline{K}_k H_k - K_k^1) E\langle x_k z_i^T \rangle = 0 \,. \tag{62}$$

For this to be true in general the coefficient of $E\langle x_k z_i^T \rangle$ must vanish, thus we conclude that

$$K_k^1 = I - \overline{K}_k H_k \,, \tag{63}$$

which is the books equation 4.13.

Using the two orthogonality conditions $E\langle (x_k - \hat{x}_k(+)) z_k(-)^T \rangle = 0$ and $E\langle (x_k - \hat{x}_k(+)) z_k^T \rangle = 0$ we can subtract these two expressions and introduce the variable $\tilde{z}_k$ defined as the error in our measurement prediction $z_k(-)$ or

$$\tilde{z}_k = z_k(-) - z_k \,, \tag{64}$$

to get $E\langle (x_k - \hat{x}_k(+)) \tilde{z}_k^T \rangle = 0$. Now using the definition of $\tilde{z}_k$ written in terms of $\hat{x}_k$ of

$$\tilde{z}_k = H_k \hat{x}_k(-) - z_k \,, \tag{65}$$

we find the orthogonality $E\langle (x_k - \hat{x}_k(+)) \tilde{z}_k^T \rangle = 0$ condition becomes

$$E\langle [x_k - K_k^1 \hat{x}_k(-) - \overline{K}_k z_k](H_k \hat{x}_k(-) - z_k)^T \rangle = 0 \,.$$

using the expression we found for K_k^1 in Equation 63 and the measurement Equation 61 this becomes

$$E\langle [x_k - \hat{x}_k(-) - \overline{K}_k H_k \hat{x}_k(-) - \overline{K}_k H_k x_k - \overline{K}_k v_k](H_k \hat{x}_k(-) - H_k x_k - v_k)^T \rangle = 0 \,.$$

Group some terms to introduce the definition of $\tilde{x}_k(-)$

$$\tilde{x}_k = x_k - \hat{x}_k(-) \,, \tag{66}$$

we have

$$E\langle [-\tilde{x}_k(-) + \overline{K}_k H_k \tilde{x}_k(-) - \overline{K}_k v_k](H_k \tilde{x}_k(-) - v_k)^T \rangle = 0 \,.$$

If we define the value of $P_k(-)$ to be the prior covariance $P_k(-) \equiv E\langle \tilde{x}_k(-) \tilde{x}_k(-)^T \rangle$ the above becomes six product terms

$$\begin{aligned} 0 &= -E\langle \tilde{x}_k(-) \tilde{x}_k(-)^T \rangle H_k^T + E\langle \tilde{x}_k(-) v_k^T \rangle \\ &+ \overline{K}_k H_k E\langle \tilde{x}_k(-) \tilde{x}_k(-)^T \rangle H_k^T - \overline{K}_k H_k E\langle \tilde{x}_k(-) v_k^T \rangle \\ &- \overline{K}_k E\langle v_k \tilde{x}_k(-)^T \rangle H_k^T + \overline{K}_k E\langle v_k v_k^T \rangle \,. \end{aligned}$$

Since $E\langle \tilde{x}_k(-) v_k^T \rangle = 0$ several terms cancel and we obtain

$$-P_k(-) H_k^T + \overline{K}_k H_k P_k(-) H_k^T + \overline{K}_k R_k = 0 \,. \tag{67}$$

Which is a linear equation for the unknown $\overline{K}_k$. Solving it we find the gain or the multiplier of the measurement given by solving the above for $\overline{K}_k$ or

$$\overline{K}_k = P_k(-) H_k^T (H_k P_k(-) H_k^T + R_k)^{-1} \,. \tag{68}$$

Using the expressions just derived for K_k^1 and $\overline{K}_k$, we would like to derive an expression for the *posterior* covariance error. The posterior covariance error is defined in a similar manner to the a-priori error $P_k(-)$ namely

$$P_k(+) = E\langle \tilde{x}_k(+) \tilde{x}_k(+) \rangle \,, \tag{69}$$

Then with the value of K_k^1 given by $K_k^1 = I - \overline{K}_k H_k$ we have our posterior state estimate $\hat{x}_k(+)$ using Equation 59 in terms of our prior estimate $\hat{x}_k(-)$ and our measurement z_k of

$$\begin{aligned}\hat{x}_k(+) &= (I - \overline{K}_k H_k)\hat{x}_k(-) + \overline{K}_k z_k \\ &= \hat{x}_k(-) + \overline{K}_k(z_k - H_k \hat{x}_k(-)) \,.\end{aligned}$$

Subtracting the true state x_k from this and writing the measurement in terms of the state as $z_k = H_k x_k + v_k$ we have

$$\begin{aligned}\hat{x}_k(+) - x_k &= \hat{x}_k(-) - x_k + \overline{K}_k H_k x_k + \overline{K}_k v_k - \overline{K}_k H_k \hat{x}_k(-) \\ &= \tilde{x}_k(-) - \overline{K}_k H_k(\hat{x}_k(-) - x_k) + \overline{K}_k v_k \\ &= \tilde{x}_k(-) - \overline{K}_k H_k \tilde{x}_k(-) + \overline{K}_k v_k \,.\end{aligned}$$

Thus the update of $\tilde{x}_k(+)$ from $\tilde{x}_k(-)$ is given by

$$\tilde{x}_k(+) = (I - \overline{K}_k H_k)\tilde{x}_k(-) + \overline{K}_k v_k \,. \tag{70}$$

Using this expression we can derive $P_k(+)$ in terms of $P_k(-)$ as

$$\begin{aligned}P_k(+) &= E\langle \tilde{x}_k(+)\tilde{x}_k^T \rangle \\ &= E\langle [I - \overline{K}_k H_k)\tilde{x}_k(-) + \overline{K}_k v_k][\tilde{x}_k^T(-)(I - \overline{K}_k H_k)^T + v_k^T \overline{K}_k^T]\rangle \,.\end{aligned}$$

By expanding the terms on the right hand side and remembering that $E\langle v_k \tilde{x}_k^T(-)\rangle = 0$ gives

$$P_k(+) = (I - \overline{K}_k H_k)P_k(-)(I - \overline{K}_k H_k)^T + \overline{K}_k R_k \overline{K}_k^T \tag{71}$$

or the so called **Joseph form** of the covariance update equation.

Alternative forms for the state covariance update equation can also be obtained. Expanding the product on the right-hand-side of Equation 71 gives

$$P_k(+) = P_k(-) - P_k(-)(\overline{K}_k H_k)^T - \overline{K}_k H_k P_k(-) + \overline{K}_k H_k P_k(-)(\overline{K}_k H_k)^T + \overline{K}_k R_k \overline{K}_k^T \,.$$

Grouping the first and third term and the last two terms together in the expression in the right-hand-side we find

$$P_k(+) = (I - \overline{K}_k H_k)P_k(-) - P_k(-)H_k^T \overline{K}_k^T + \overline{K}_k(H_k P_k(-)H_k^T + R_k)\overline{K}_k^T \,.$$

Recognizing that since the expression $H_k P_k(-)H_k^T + R_k$ appears in the definition of the Kalman gain Equation 68 the product $\overline{K}_k(H_k P_k(-)H_k^T + R_k)$ is really equal to

$$\overline{K}_k(H_k P_k(-)H_k^T + R_k) = P_k(-)H_k^T \,,$$

and we find $P_k(+)$ takes the form

$$\begin{aligned}P_k(+) &= (I - \overline{K}_k H_k)P_k(-) - P_k(-)H_k^T \overline{K}_k^T + P_k(-)H_k^T \overline{K}_k^T \\ &= (I - \overline{K}_k H_k)P_k(-) \,.\end{aligned} \tag{72}$$

This later form is most often used in computation.

Given the estimate of the error covariance at the *previous* time step or $P_{k-1}(+)$ by using the discrete state-update equation

$$x_k = \Phi_{k-1} x_{k-1} + w_{k-1} \,, \tag{73}$$

the prior error covariance at the next time step k is given by the simple form

$$P_k(-) = \Phi_{k-1} P_{k-1}(+) \Phi_{k-1}^T + Q_{k-1} \,. \tag{74}$$

Notes on Treating Uncorrelated Measurement Vectors as Scalar Measurements

In this subsection of the book a very useful algorithm for dealing with uncorrelated measurement vectors is presented. The main idea is to treat the totality of vector measurement $\mathbf{z}$ as a *sequence* of scalar measurements z_k for $k = 1, 2, \cdots, l$. This can have several benefits. In addition to the two reasons stated in the text: reduced computational time and improved numerical accuracy, in practice this algorithm can be especially useful in situations where the individual measurements are known with different uncertainties where some maybe more informative and useful in predicting an estimate of the total state $\hat{x}_k(+)$ than others. In an ideal case one would like to use the information from all of the measurements but time may require estimates of $\hat{x}_k(+)$ quicker than the computation with all measurements could be done. If the measurements could be sorted based on some sort of priority (like uncertainty) then an approximation of $\hat{x}_k(+)$ could be obtained by applying on the most informative measurements z_k first and stopping before processing all of the measurements. This algorithm is also a very interesting way of thinking about how the Kalman filter is in general processing vector measurements. There is slight typo in the book's presented algorithm which we now fix. The algorithm is to begin with our initial estimate of the state and covariance $P_k^{[0]} = P_k(-)$ and $\hat{x}_k^{[0]} = \hat{x}_k(-)$ and then to iteratively apply the following equations

$$
\begin{aligned}
\overline{K}_k^{[i]} &= \frac{1}{H_k^{[i]} P_k^{[i-1]} H_{[i]}^T + R_k^{[i]}} (H_k^{[i]} P_k^{[i-1]})^T \\
P_k^{[i]} &= P_k^{[i-1]} - \overline{K}_k^{[i]} H_k^{[i]} P_k^{[i-1]} \\
\hat{x}_k^{[i]} &= \hat{x}_k^{[i-1]} + \overline{K}_k^{[i]} [\{z_k\}_i - H_k^{[i]} \hat{x}_k^{[i-1]}] \,,
\end{aligned}
$$

for $i = 1, 2, \cdots, l$. As shown above, a simplification over the normal Kalman update equations that comes from using this procedure is that now the expression $H_k^{[i]} P_k^{[i-1]} H_{[i]}^T + R_k^{[i]}$ is a *scalar* and inverting it is simply division. Once we have processed the l-th scalar measurement $\{z_k\}_l$, using this procedure the final state and uncertainty estimates are given by

$$
P_k(+) = P_k^{[l]} \quad \text{and} \quad \hat{x}_k(+) = \hat{x}_k^{[l]} \,.
$$

On Page 80 of these notes we derive the computational requirements for the normal Kalman formulation (where the measurements $\mathbf{z}$ are treated as a vector) and the above "scalar" procedure. In addition, we should note that theoretically the *order* in which we process each scalar measurement should not matter. In practice, however, it seems that it does matter and different ordering can give different state estimates. Ordering the measurements from most informative (the measurement with the smallest uncertainty is first) to least informative seems to be a good choice. This corresponds to a greedy like algorithm in that if we have to stop processing measurements at some point we would have processed the measurements with the largest amount of information.

Notes on the Section Entitled: The Kalman-Bucy filter

Warning: I was not able to get the algebra in this section to agree with the results presented in the book. If anyone sees an error in my reasoning or a method by which I should do these calculations differently please email me.

By putting the covariance update Equation 89 into the error covariance extrapolation Equation 74 we obtain a recursive equation for $P_k(-)$ given by

$$P_k(-) = \Phi_{k-1}(I - \overline{K}_{k-1}H_{k-1})P_{k-1}(-)\Phi_{k-1}^T + G_kQ_kG_k^T\,. \qquad (75)$$

Mapping from the discrete space to the continuous space we assume $F_{k-1} = F(t_{k-1})$, $G_k = G(t_k)$, $Q_k = Q(t_k)\Delta t$, and $\Phi_{k-1} \approx I + F_{k-1}\Delta t$ then the above discrete approximations to the continuous Kalman-Bucy system becomes

$$P_k(-) = (I + F_{k-1}\Delta t)(I - \overline{K}_{k-1}H_{k-1})P_{k-1}(-)(I + F_{k-1}\Delta t)^T + G_kQ_kG_k^T\Delta t\,.$$

On expanding the product in the right hand side (done in two steps) of the above we find

$$\begin{aligned}
P_k(-) &= (I + F_{k-1}\Delta t) \\
&\times\ (P_{k-1}(-) + \Delta t P_{k-1}(-)F_{k-1}^T - \overline{K}_{k-1}H_{k-1}P_{k-1}(-) - \Delta t\overline{K}_{k-1}H_{k-1}P_{k-1}(-)F_{k-1}^T) \\
&+\ G_kQ_kG_k^T\Delta t \\
&= P_{k-1}(-) + \Delta t P_{k-1}(-)F_{k-1}^T - \overline{K}_{k-1}H_{k-1}P_{k-1}(-) - \Delta t\overline{K}_{k-1}H_{k-1}P_{k-1}(-)F_{k-1}^T \\
&+\ \Delta t F_{k-1}P_{k-1}(-) + \Delta t^2 F_{k-1}P_{k-1}(-)F_{k-1}^T - \Delta t F_{k-1}\overline{K}_{k-1}H_{k-1}P_{k-1}(-) \\
&-\ \Delta t^2 F_{k-1}\overline{K}_{k-1}H_{k-1}P_{k-1}(-)F_{k-1}^T \\
&+\ G_kQ_kG_k^T\Delta t\,.
\end{aligned}$$

Now forming the first difference of $P_k(-)$ on the left hand side of the above and rearranging terms we find to

$$\begin{aligned}
\frac{P_k(-) - P_{k-1}(-)}{\Delta t} &= P_{k-1}(-)F_{k-1}^T - \frac{1}{\Delta t}\overline{K}_{k-1}H_{k-1}P_{k-1}(-) - \mathbf{\overline{K}_{k-1}H_{k-1}P_{k-1}}(-)\mathbf{F_{k-1}^T} \\
&+\ F_{k-1}P_{k-1}(-) + \Delta t F_{k-1}P_{k-1}(-)F_{k-1}^T - \mathbf{F_{k-1}\overline{K}_{k-1}H_{k-1}P_{k-1}}(-) \\
&-\ \Delta t F_{k-1}\overline{K}_{k-1}H_{k-1}P_{k-1}(-)F_{k-1}^T \\
&+\ G_kQ_tG_k^T\,.
\end{aligned}$$

Taking $\Delta t \to 0$ and using the fact that $\lim_{\Delta t\to 0}\frac{\overline{K}_{k-1}}{\Delta t} = PH^TR^{-1} = \overline{K}(t)$ should give the *continuous* matrix Riccati equation

$$\dot{P}(t) = P(t)F(t)^T + F(t)P(t) - P(t)H(t)^TR^{-1}(t)H(t)P(t) + G(t)Q(t)G(t)^T\,. \qquad (76)$$

Note: As mentioned above, I don't see how when the limit $\Delta t \to 0$ is taken to eliminate the terms in **bold** above: $-\overline{K}_{k-1}H_{k-1}P_{k-1}(-)F_{k-1}^T$ and $-F_{k-1}\overline{K}_{k-1}H_{k-1}P_{k-1}(-)$. If anyone can find an error in what I have done please email me.

Notes on the Section Entitled: Solving the Matrix Riccati Differential Equation

Consider a fractional decomposition of the covariance $P(t)$ as $P(t) = A(t)B(t)^{-1}$. Then the continuous Riccati differential equation

$$\dot{P}(t) = F(t)P(t) + P(t)F(t)^T - P(t)H(t)^TR^{-1}(t)H(t)P(t) + Q(t)\,,$$

under this substitution becomes

$$
\begin{aligned}
\frac{d}{dt}P(t) &= \frac{d}{dt}(A(t)B(t)^{-1}) = \dot{A}(t)B(t)^{-1} - A(t)B(t)^{-1}\dot{B}(t)B^{-1}(t) \\
&= F(t)A(t)B(t)^{-1} + A(t)B(t)^{-1}F(t)^T - A(t)B(t)^{-1}H(t)^T R(t)^{-1}H(t)A(t)B(t)^{-1} + Q(t)\,.
\end{aligned}
$$

Or multiplying by $B(t)$ on the left the above becomes

$$
\begin{aligned}
\dot{A}(t) - A(t)B(t)^{-1}\dot{B}(t) &= F(t)A(t) + A(t)B(t)^{-1}F(t)^T B(t) \\
&- A(t)B(t)^{-1}H(t)^T R(t)^{-1}H(t)A(t) + Q(t)B(t)\,.
\end{aligned}
$$

Now factor the expansion $A(t)B(t)^{-1}$ from the second and third terms as

$$
\begin{aligned}
\dot{A}(t) - A(t)B(t)^{-1}\dot{B}(t) &= F(t)A(t) + Q(t)B(t) \\
&+ A(t)B(t)^{-1}(F(t)^T B(t) - H(t)^T R(t)^{-1}H(t)A(t))\,.
\end{aligned}
$$

This equation will be satisfied if we can find matrices $A(t)$ and $B(t)$ such that the coefficients of $A(t)B(t)^{-1}$ are equal. Equating the *zeroth* power of $A(t)B(t)^{-1}$ gives an equation for $A(t)$ of

$$
\dot{A}(t) = F(t)A(t) + Q(t)B(t)\,.
$$

Equating the *first* powers of $A(t)B(t)^{-1}$ requires that $B(t)$ must satisfy

$$
\dot{B}(t) = H(t)^T R(t)^{-1}H(t)A(t) - F(t)^T B(t)\,.
$$

In matrix form these two equations can be expressed as

$$
\frac{d}{dt}\begin{bmatrix} A(t) \\ B(t) \end{bmatrix} = \begin{bmatrix} F(t) & Q(t) \\ H(t)^T R(t)^{-1}H(t) & -F(t)^T \end{bmatrix}\begin{bmatrix} A(t) \\ B(t) \end{bmatrix},
$$

which is the books equation 4.67.

Notes on: General Solution of the Scalar Time-Invariant Riccati Equation

Once we have solved for the scalar functions $A(t)$ and $B(t)$ we can explicitly evaluate the time varying scalar covariance $P(t)$ as $P(t) = \frac{A(t)}{B(t)}$. If we desire to consider the *steady-state* value of this expression we have (using some of the results from this section of the book) that

$$
\begin{aligned}
\lim_{t\to\infty} P(t) &= \frac{\lim_{t\to\infty}\mathcal{N}_P(t)}{\lim_{t\to\infty}\mathcal{D}_P(t)} \\
&= \frac{R\left[P(0)\left(\sqrt{F^2 + \frac{H^2 Q}{R}} + F\right) + Q\right]}{H^2 P(0) + R\left(\sqrt{F^2 + \frac{H^2 Q}{R}} - F\right)} \\
&= \left(\frac{R}{H^2}\right)\frac{\left(F + \sqrt{F^2 + \frac{H^2 Q}{R}}\right)\left[P(0) + Q\left(\sqrt{F^2 + \frac{H^2 Q}{R}} + F\right)^{-1}\right]}{\left[P(0) + \frac{R}{H^2}\left(\sqrt{F^2 + \frac{H^2 Q}{R}} + F\right)\right]}\,.
\end{aligned}
$$

Consider the expression in the upper right hand "corner" of the above expression or

$$\frac{Q}{\sqrt{F^2 + \frac{H^2 Q}{R}} + F} \,,$$

by multiplying top and bottom of this fraction by $\frac{\sqrt{F^2+\frac{H^2Q}{R}}-F}{\sqrt{F^2+\frac{H^2Q}{R}}-F}$ we get

$$\frac{Q\left(\sqrt{F^2+\frac{H^2Q}{R}}-F\right)}{F^2+\frac{H^2Q}{R}-F^2} = \frac{R}{H^2}\left(\sqrt{F^2+\frac{H^2Q}{R}}-F\right) \,,$$

and the terms in the brackets $[\cdot]$ cancel each from the numerator and denominator to give the expression

$$\lim_{t\to\infty} P(t) = \frac{R}{H^2}\left(F + \sqrt{F^2+\frac{H^2Q}{R}}\right) \,, \tag{77}$$

which is the books equation 4.72.

Notes on: The Steady-State Riccati equation using the Newton-Raphson Method

In the notation of this section, the identity that

$$\frac{\partial P}{\partial P_{kl}} = I_{\cdot k} I_{\cdot l}^T \,, \tag{78}$$

can be reasoned as correct by recognizing that I_l^T represents the row vector with a one in the l-th spot and $I_{\cdot k}$ represents a column vector with a one in the k-th spot, so the product of $I_{\cdot k} I_{\cdot l}^T$ represents a matrix of zeros with a single non-zero element (a 1) in the kl-th spot. This is the equivalent effect of taking the derivative of P with respect to its kl-th element or the expression $\frac{\partial P}{\partial P_{kl}}$.

From the given definition of $\mathcal{Z}$, the product rule, and Equation 78 we have

$$\begin{aligned}
\frac{\partial \mathcal{Z}}{\partial P_{kl}} &= \frac{\partial}{\partial P_{kl}}(FP + PF^T - PH^TR^{-1}HP + Q) \\
&= F\frac{\partial P}{\partial P_{kl}} + \frac{\partial P}{\partial P_{kl}}F^T - \frac{\partial P}{\partial P_{kl}}H^TR^{-1}HP - PH^TR^{-1}H\frac{\partial P}{\partial P_{kl}} \\
&= FI_{\cdot k}I_{\cdot l}^T + I_{\cdot k}I_{\cdot l}^TF^T - I_{\cdot k}I_{\cdot l}^TH^TR^{-1}HP - PH^TR^{-1}HI_{\cdot k}I_{\cdot l}^T \\
&= F_{\cdot k}I_{\cdot l}^T + I_{\cdot k}F_{\cdot l}^T - I_{\cdot k}I_{\cdot l}^T(PH^TR^{-1}H)^T - (PH^TR^{-1}H)I_{\cdot k}I_{\cdot l}^T \,.
\end{aligned}$$

In deriving the last line we have used the fact $I_{\cdot l}^T F^T = (FI_{\cdot l})^T = F_{\cdot l}^T$. Note that the last term above is

$$-(PH^TR^{-1}H)I_{\cdot k}I_{\cdot l}^T = -MI_{\cdot k}I_{\cdot l}^T = -M_{\cdot k}I_{\cdot l}^T \,,$$

where we have introduced the matrix $M \equiv PH^TR^{-1}H$, since $MI_{\cdot k}$ selects the kth column from the matrix M. This is the fourth term in the books equation 4.85. The product in the second to last term is given by

$$-I_{\cdot k}I_{\cdot l}^TH^TR^{-1}HP = -I_{\cdot k}(PH^TR^{-1}HI_{\cdot l})^T = -I_{\cdot k}M_{\cdot l}^T \,,$$

and is the third term in the books equation 4.85. Taken together we get the books equation 4.86. Rearranging the resulting terms and defining the matrix $\mathcal{S} \equiv F - M$ gives

$$\begin{aligned}
\frac{\partial \mathcal{Z}}{\partial P_{kl}} &= (F_{\cdot k} - M_{\cdot k})I_{\cdot l}^T + I_{\cdot k}(F_{\cdot l}^T - M_{\cdot l}^T) \\
&= (F - M)_{\cdot k}I_{\cdot l}^T + I_{\cdot k}((F - M)^T)_{\cdot l} \\
&= \mathcal{S}_{\cdot k}I_{\cdot l}^T + I_{\cdot k}(\mathcal{S}_{\cdot l}^T) \\
&= \mathcal{S}_{\cdot k}I_{\cdot l}^T + (\mathcal{S}_{\cdot l}I_{\cdot k})^T ,
\end{aligned}$$

this is the books equation 4.87.

Now recall that $I_{\cdot k}$ represents a column vector with one in the k-th spot, and $I_{\cdot l}^T$ is a row vector with a one in the l-th spot, so the product $\mathcal{S}_{\cdot k}I_{\cdot l}^T$ (which is the first term in the above expression) represents the k-th column of the matrix $\mathcal{S}$ times the row vector $I_{\cdot l}^T$ where only the l-th column element is non-zero and therefore equals a matrix of all zeros except in the the *l-th* column where the elements are equal to the k-th column of $\mathcal{S}$. In the same way the term in the above expression $(\mathcal{S}_{\cdot l}I_{\cdot k}^T)^T$ has the l-th column of $\mathcal{S}$ in the k-th *row* of the resulting matrix.

Now the expression $\frac{\partial \mathcal{Z}_{ij}}{\partial P_{kl}}$, represents taking the derivative of the ij-th element of the matrix $\mathcal{Z}$ with respect to the kl-th element of the matrix P. Since we have already calculated the *matrix* $\frac{\partial \mathcal{Z}}{\partial P_{kl}}$, to calculate

$$\mathcal{F}_{pq} \equiv \frac{\partial f_p}{\partial x_q} = \frac{\partial \mathcal{Z}_{ij}}{\partial P_{kl}} ,$$

we need to extract the ij-th element from this matrix. As discussed above, since $\mathcal{S}_{\cdot k}I_{\cdot l}^T$ has only a nonzero l-th column this derivative will be non-zero if and only if $j = l$, where its value will be $\mathcal{S}_{ik}$. Also since $I_{\cdot k}\mathcal{S}_{\cdot l}^T$ has only a nonzero k-th row, this derivative will be non-zero if and only if $i = k$ where its value will be $\mathcal{S}_{jl}$. Thus we finally obtain

$$\frac{\partial \mathcal{Z}_{ij}}{\partial P_{kl}} = \Delta_{jl}\mathcal{S}_{ik} + \Delta_{ik}\mathcal{S}_{jl} , \tag{79}$$

which is the books equation 4.80.

Notes on: MacFarlane-Potter-Fath Eigenstructure Method

From the given definition of the continuous-time system Hamiltonian matrix, Ψ_c, we can compute the product discussed in Lemma 1

$$\Psi_c \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} F & Q \\ H^T R^{-1} H & -F^T \end{bmatrix} \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} FA + QB \\ H^T R^{-1} HA - F^T B \end{bmatrix} = \begin{bmatrix} AD \\ BD \end{bmatrix} .$$

Looking at the individual equations we have the system of

$$AD = FA + QB \tag{80}$$
$$BD = H^T R^{-1} HA - F^T B \tag{81}$$

Multiply both equations by B^{-1} on the right to get

$$\begin{aligned} ADB^{-1} &= FAB^{-1} + Q && (82) \\ BDB^{-1} &= H^T R^{-1} H A B^{-1} - F^T && (83) \end{aligned}$$

No multiply Equation 83 on the left by AB^{-1} to get

$$ADB^{-1} = AB^{-1} H^T R^{-1} H A B^{-1} - AB^{-1} F^T \,. \quad (84)$$

Setting the expressions for ADB^{-1} in Equations 82 and 84 equal while recalling our fractional factorization of $P = AB^{-1}$ we obtain

$$0 = FP - PH^T R^{-1} HP + PF^T + Q \,,$$

the continuous steady-state Riccati equation.

Steady-State Solution of the Time-Invariant Discrete-Time Riccati Equation

For this section we need the following "Riccati" result which is the recursive representation of the a priori covariance matrix $P_k(-)$. Recall that the covariance extrapolation step in discrete Kalman filtering can be written recursively as

$$\begin{aligned} P_{k+1}(-) &= \Phi_k P_k(+) \Phi_k^T + Q_k \\ &= \Phi_k (I - \overline{K}_k H_k) P_k(-) \Phi_k^T + Q_k \\ &= \Phi_k \{ I - P_k(-) H_k^T (H_k P_k(-) H_k^T + R_k)^{-1} H_k \} P_k(-) \Phi_k^T + Q_k \,. && (85) \end{aligned}$$

As discussed in the book this equation has a solution given in the following factorization

$$P_k(-) = A_k B_k^{-1} \,,$$

where A_k and B_k satisfy the following recursion relationship

$$\begin{aligned} \begin{bmatrix} A_{k+1} \\ B_{k+1} \end{bmatrix} &= \begin{bmatrix} Q_k & I \\ I & 0 \end{bmatrix} \begin{bmatrix} \Phi_k^{-T} & 0 \\ 0 & \Phi_k \end{bmatrix} \begin{bmatrix} H_k^T R_k^{-1} H_k & I \\ I & 0 \end{bmatrix} \begin{bmatrix} A_k \\ B_k \end{bmatrix} \\ &= \begin{bmatrix} \Phi_k + Q_k \Phi_k^{-T} H_k^T R_k^{-1} H_k & Q_k \Phi_k^{-T} \\ \Phi_k^{-T} H_k^T R_k^{-1} H_k & \Phi_k^{-T} \end{bmatrix} \begin{bmatrix} A_k \\ B_k \end{bmatrix} . \end{aligned}$$

We define the coefficient matrix above as Ψ_d or

$$\Psi_d \equiv \begin{bmatrix} \Phi_k + Q_k \Phi_k^{-T} H_k^T R_k^{-1} H_k & Q_k \Phi_k^{-T} \\ \Phi_k^{-T} H_k^T R_k^{-1} H_k & \Phi_k^{-T} \end{bmatrix} . \quad (86)$$

If we restrict to the case where everything is a scalar and time-invariant the coefficient matrix Ψ_d in this case becomes

$$\Psi_d = \begin{bmatrix} Q & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \Phi^{-1} & 0 \\ 0 & \Phi \end{bmatrix} \begin{bmatrix} \frac{H^2}{R} & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{Q}{\Phi} & \Phi \\ \frac{1}{\Phi} & 0 \end{bmatrix} \begin{bmatrix} \frac{H^2}{R} & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \Phi + \frac{QH^2}{\Phi R} & \frac{Q}{\Phi} \\ \frac{H^2}{\Phi R} & \frac{1}{\Phi} \end{bmatrix} .$$

To solve for A_k and B_k for all k we then diagonalize Ψ_d as MDM^{-1} and begin from the initial condition on P translated into initial conditions on A and B. That is we want $P_0 = A_0 B_0^{-1}$ which we can obtain by taking $A_0 = P_0$ and $B_0 = I$.

If we assume that our system is time-invariant to study the steady-state filter performance we let $k \to \infty$ in Equation 85 and get

$$P_\infty = \Phi\{I - P_\infty H^T (H P_\infty H^T + R)^{-1} H\} P_\infty \Phi^T + Q\,. \tag{87}$$

Which is the equation we desire to solve via the eigenvalues of the block matrix Ψ_d. Specifically the steady state solution to Equation 87 can be represented as $P_\infty = AB^{-1}$ where A and B satisfy

$$\Psi_d \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix} D\,,$$

for a $n \times n$ nonsingular matrix D. In practice A and B are formed from the n characteristic vectors of Ψ_d corresponding to the *nonzero* characteristic values of Ψ_d.

Problem Solutions

Problem 4.1 (the non-recursive Bayes solution)

The way to view this problem is to recognize that since everything is linear and distributed as a Gaussian random variable the end result (i.e. the posteriori distribution of x_1 given z_0, z_1, z_2) must also be Gaussian. Thus if we can compute the *joint* distribution of the vector $\begin{bmatrix} x_1 \\ z_0 \\ z_1 \\ z_2 \end{bmatrix}$, say $p(x_1, z_0, z_1, z_2)$, then using this we can compute the optimal estimate of x_1 by computing the posterior-distribution of x_1 i.e. $p(x_1|z_0, z_1, z_2)$. Since everything is linear and Gaussian the joint distribution $p(x_1, z_0, z_1, z_2)$ will be Gaussian and the posterior-distribution $p(x_1|z_0, z_1, z_2)$ will *also* be Gaussian with a mean and a covariance given by classic formulas.

Thus as a first step we need to determine the probability density of the vector $\begin{bmatrix} x_1 \\ z_0 \\ z_1 \\ z_2 \end{bmatrix}$.

From the problem specified system dynamic and measurement equation we can compute the various sequential measurements and dynamic time steps starting from the first measurement

z_0 until the third measurement z_2 as

$$\begin{aligned}
z_0 &= x_0 + v_0 \\
x_1 &= \frac{1}{2}x_0 + w_0 \\
z_1 &= x_1 + v_1 = \frac{1}{2}x_0 + w_0 + v_1 \\
x_2 &= \frac{1}{2}x_1 + w_1 = \frac{1}{2}\left(\frac{1}{2}x_0 + w_0\right) + w_1 = \frac{1}{4}x_0 + \frac{1}{2}w_0 + w_1 \\
z_2 &= x_2 + v_2 = \frac{1}{4}x_0 + \frac{1}{2}w_0 + w_1 + v_2 \,.
\end{aligned}$$

In matrix notation these equations are given by

$$\begin{bmatrix} x_1 \\ z_0 \\ z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 1 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{2} & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ v_0 \\ w_0 \\ v_1 \\ w_1 \\ v_2 \end{bmatrix} .$$

Note these are written in such a way that the variables on the right-hand-side of the above expression: $x_0, v_0, w_0, v_1, w_1, v_1$ are independent and drawn from zero mean unit variance normal distributions. Because of this, the vector on the left-hand-side, $\begin{bmatrix} x_1 \\ z_0 \\ z_1 \\ z_2 \end{bmatrix}$, has a Gaussian distribution with a expectation given by the zero vector and a covariance given by

$$C \equiv \begin{bmatrix} \frac{1}{2} & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 1 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{2} & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 1 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{2} & 0 & 1 & 1 \end{bmatrix}^T = \frac{1}{16} \begin{bmatrix} 20 & 8 & 20 & 10 \\ 8 & 32 & 8 & 4 \\ 20 & 8 & 36 & 10 \\ 10 & 4 & 10 & 37 \end{bmatrix} ,$$

since the variance of the vector of variables $x_0, v_0, w_0, v_1, w_1, v_1$ is the six-by-six identity matrix. We will partition this covariance matrix in the following way

$$C = \begin{bmatrix} c_{x_1}^2 & b^T \\ b & \hat{C} \end{bmatrix} .$$

Here the upper left corner element $c_{x_1}^2$ is the variance of the random variable x_1 that we want to compute the expectation of. Thus we have defined

$$c_{x_1}^2 = 5/4 \,, \quad b^T = \begin{bmatrix} 1/2 & 5/4 & 5/8 \end{bmatrix} , \quad \text{and} \quad \hat{C} = \begin{bmatrix} 2 & 1/2 & 1/4 \\ 1/2 & 9/4 & 5/8 \\ 1/4 & 5/8 & 37/16 \end{bmatrix} .$$

Given the distribution of the *joint* we would like to compute the distribution of x_1 given the values of z_0, z_1, and z_2. To do this we will use the following theorem.

Given X, a multivariate Gaussian random variable of dimension n with vector mean μ and covariance matrix Σ. If we *partition* X, μ, and Σ into two parts of sizes q and $n-q$ as

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \quad \text{and} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{bmatrix}.$$

Then the conditional distribution of the first q random variables in X *given* the second $n-q$ of the random variables (say $X_2 = a$) is another multivariate normal with mean $\bar{\mu}$ and covariance $\bar{\Sigma}$ given by

$$\begin{aligned} \bar{\mu} &= \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(a - \mu_2) && (88) \\ \bar{\Sigma} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}. && (89) \end{aligned}$$

For this problem we have that $\Sigma_{11} = c_{x_1}^2$, $\Sigma_{12} = b^T$, and $\Sigma_{22} = \hat{C}$, so that we compute the matrix product $\Sigma_{12}\Sigma_{22}^{-1}$ of

$$\Sigma_{12}\Sigma_{22}^{-1} = \frac{1}{145}\begin{bmatrix} 16 & 72 & 18 \end{bmatrix}.$$

Thus if we are given the values of z_0, z_1, and z_2 for the components of X_2 from the above theorem the value of $E[x_1|z_0, z_1, z_2]$ is given by $\bar{\mu}$ which in this case since $\mu_1 = 0$ and $\mu_2 = 0$ becomes

$$E[x_1|z_0, z_1, z_2] = \frac{1}{145}\begin{bmatrix} 16 & 72 & 18 \end{bmatrix}\begin{bmatrix} z_0 \\ z_1 \\ z_2 \end{bmatrix} = \frac{1}{145}(16z_0 + 72z_1 + 18z_2).$$

The simple numerics for this problem are worked in the MATLAB script `prob_4_1.m`.

Problem 4.2 (solving Problem 4.1 using the discrete Kalman filter)

Part (a): For this problem we have $\Phi_{k-1} = \frac{1}{2}$, $H_k = 1$, $R_k = 1$, and $Q_k = 1$, then the discrete Kalman equations become

$$\begin{aligned} \hat{x}_k(-) &= \Phi_{k-1}\hat{x}_{k-1}(+) = \frac{1}{2}\hat{x}_{k-1}(+) \\ P_k(-) &= \Phi_{k-1}P_{k-1}(+)\Phi_{k-1}^T + Q_{k-1} = \frac{1}{4}P_{k-1}(+) + 1 \\ \overline{K}_k &= P_k(-)H_k^T(H_kP_k(-)H_k^T + R_k)^{-1} = \frac{P_k(-)}{P_k(-)+1} \\ \hat{x}_k(+) &= \hat{x}_k(-) + \overline{K}_k(z_k - H_k\hat{x}_k(-)) = \hat{x}_k(-) + \overline{K}_k(z_k - \hat{x}_k(-)) && (90) \\ P_k(+) &= (I - \overline{K}_kH_k)P_k(-) = (1 - \overline{K}_k)P_k(-). && (91) \end{aligned}$$

Part (b): If the measurement z_2 was not received we can skip the equations used to update the state and covariance after each measurement. Thus Equations 90 and 91 would instead become (since z_2 is not available)

$$\begin{aligned} \hat{x}_2(+) &= \hat{x}_2(-) \\ P_2(+) &= P_2(-), \end{aligned}$$

but this modification happens only for this one step.

Part (c): Now when we compute $\hat{x}_3(-)$ assuming we had the measurement z_2 we would have a contribution

$$
\begin{aligned}
\hat{x}_3(-) &= \frac{1}{2}\hat{x}_2(+) = \frac{1}{2}\left(\hat{x}_2(-) + \overline{K}_2(z_2 - \hat{x}_2(-)\right) \\
&= \frac{1}{2}\hat{x}_2(-) + \frac{1}{2}\overline{K}_2\left(z_2 - \hat{x}_2(-)\right) .
\end{aligned}
$$

The measured z_2 is *not* received the corresponding expression above won't have the term $\frac{1}{2}\overline{K}_2(z_2 - \hat{x}_2(-))$ which quantifies the loss of information in the estimate $\hat{x}_3(-)$.

Part (d): The iterative update equations for $P_k(+)$ are obtained as

$$
\begin{aligned}
P_k(+) &= \left(1 - \frac{P_k(-)}{P_k(-) + 1}\right) P_k(-) \\
&= \left(\frac{1}{P_k(-) + 1}\right) P_k(-) = \left(\frac{1}{\frac{1}{4}P_{k-1}(+) + 2}\right)\left(\frac{1}{4}P_{k-1}(+) + 1\right) .
\end{aligned}
$$

When $k \to \infty$ our steady state covariance $P_k(+) = P_\infty(+)$ which we could then solve. For $P_\infty(-)$ we have

$$
\begin{aligned}
P_k(-) &= \frac{1}{4}P_{k-1}(+) + 1 \\
&= \frac{1}{4}(1 - \overline{K}_{k-1})P_{k-1}(-) + 1 \\
&= \frac{1}{4}\left(1 - \frac{P_{k-1}(-)}{P_{k-1}(-) + 1}\right) P_{k-1}(-) + 1 \\
&= \frac{1}{4}\left(\frac{1}{P_{k-1}(-) + 1}\right) P_{k-1}(-) + 1 .
\end{aligned}
$$

When $k \to \infty$ our steady state covariance $P_k(-) = P_\infty(-)$ which we could then solve.

Part (e): If every other measurement is missing then we replace Equations 90 and 91 with

$$
\begin{aligned}
\hat{x}_{2k}(+) &= \hat{x}_{2k}(-) \\
P_{2k}(+) &= P_{2k}(-) ,
\end{aligned}
$$

so that the *total* discrete filter becomes

$$\begin{aligned}
\hat{x}_k(-) &= \frac{1}{2}\hat{x}_{k-1}(+) \\
P_k(-) &= \frac{1}{4}P_{k-1}(+) + 1 \\
\overline{K}_k &= \frac{P_k(-)}{P_k(-)+1} \\
\hat{x}_k(+) &= \hat{x}_k(-) + \overline{K}_k(z_k - \hat{x}_k(-)) \\
P_k(+) &= (1-\overline{K}_k)P_k(-) \\
\hat{x}_{k+1}(-) &= \frac{1}{2}\hat{x}_k(+) \\
P_{k+1}(-) &= \frac{1}{4}P_k(+) + 1 \\
\hat{x}_{k+1}(+) &= \hat{x}_{k+1}(-) \\
P_{k+1}(+) &= P_{k+1}(-)\,.
\end{aligned}$$

Problem 4.3 (filtering a continuous problem using discrete measurements)

I was not sure how to do this problem. Please email me if you have suggestions.

Problem 4.4 (filtering a continuous problem using integrated measurements)

I was not sure how to do this problem. Please email me if you have suggestions.

Problem 4.5 (deriving that $E\langle w_k z_i^T\rangle = 0$)

Consider the expression $E\langle w_k z_i^T\rangle$. By using $z_i = H_i x_i + v_i$ we can write this expression as

$$\begin{aligned}
E\langle w_k z_i^T\rangle &= E\langle w_k (H_i x_i + v_i)^T\rangle \\
&= E\langle w_k x_i^T\rangle H_i^T + E\langle w_k v_i^T\rangle \\
&= E\langle w_k x_i^T\rangle H_i^T\,,
\end{aligned}$$

since w_k and v_k are uncorrelated. Using the discrete dynamic equation $x_i = \Phi_{i-1}x_{i-1} + w_{i-1}$ we can write the above as

$$\begin{aligned}
E\langle w_k z_i^T\rangle &= E\langle w_k (\Phi_{i-1}x_{i-1} + w_{i-1})^T\rangle H_i^T \\
&= E\langle w_k x_{i-1}^T\rangle \Phi_{i-1}^T H_i^T + E\langle w_k w_{i-1}^T\rangle H_i^T \\
&= E\langle w_k x_{i-1}^T\rangle \Phi_{i-1}^T H_i^T\,,
\end{aligned}$$

since $E\langle w_k w_{i-1}^T\rangle = 0$ when $i \leq k$ as w_k is uncorrelated white noise. Continuing to use dynamic equations to replace x_l with an expression in terms of x_{l-1} we eventually get

$$E\langle w_k z_i^T\rangle = E\langle w_k x_0^T\rangle \Phi_0^T \Phi_1^T \cdots \Phi_{i-2}^T \Phi_{i-1}^T H_i^T\,.$$

If we assume x_0 is either fixed (deterministic), independent of w_k, or uncorrelated with w_k this last expectation is zero proving the desired conjecture.

Problem 4.6 (a simpler mathematical model for Example 4.4)

In Exercise 4.4 the system state $\mathbf{x}$, was defined with two additional variables U_k^1 and U_k^2 which are the maneuvering-correlated noise for the range rate $\dot{r}$ and the bearing rate $\dot{\theta}$ respectively. Both are assumed to be given as an AR(1) model with an AR(1) coefficients ρ and r such that

$$\begin{aligned} U_k^1 &= \rho U_{k-1}^1 + w_{k-1}^1 \\ U_k^2 &= r U_{k-1}^2 + w_{k-1}^2 \,, \end{aligned}$$

where w_{k-1}^1 and w_{k-1}^2 are white noise innovations. Because the noise in this formulation is autocorrelated better system modeling results if these two terms are explicitly included in the definition of the state $\mathbf{x}$. In Example 4.4 they are the third and sixth unknowns. If however we take a simpler model where the noise applied to the range rate $\dot{r}$ and the bearing rate $\dot{\theta}$ is in fact *not* colored then we don't need to include these two terms as unknowns in the state and the reduced state becomes simply

$$\mathbf{x}^T = \begin{bmatrix} r & \dot{r} & \theta & \dot{\theta} \end{bmatrix} .$$

The dynamics in this state-space given by

$$x_k = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} x_{k-1} + \begin{bmatrix} 0 \\ w_{k-1}^1 \\ 0 \\ w_{k-1}^2 \end{bmatrix} ,$$

with a discrete observation equation of

$$z_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} x_k + \begin{bmatrix} v_k^1 \\ v_k^2 \end{bmatrix} .$$

To use the same values of P_0, Q, R, σ_r^2, σ_θ^2, σ_1^2, and σ_2^2 as in Example 4.4 with our new state definition we would have

$$P_0 = \begin{bmatrix} \sigma_r^2 & \frac{\sigma_r^2}{T} & 0 & 0 \\ \frac{\sigma_r^2}{T} & \frac{2\sigma_r^2}{T^2} + \sigma_1^2 & 0 & 0 \\ 0 & 0 & \sigma_\theta^2 & \frac{\sigma_\theta^2}{T} \\ 0 & 0 & \frac{\sigma_\theta^2}{T} & \frac{2\sigma_\theta^2}{T^2} + \sigma_2^2 \end{bmatrix} , \quad Q = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_2^2 \end{bmatrix} , \quad R = \begin{bmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_\theta^2 \end{bmatrix} ,$$

with $T = 5, 10, 15$ and parameters given by

$$\begin{aligned} \sigma_r^2 &= (1000\,\mathrm{m})^2 \\ \sigma_1^2 &= (100/3)^2 \\ \sigma_\theta^2 &= (0.017\,\mathrm{rad})^2 \\ \sigma_2^2 &= 1.3\,10^{-8} . \end{aligned}$$

The remaining part of this problem would be to generate plots of $P_k(-)$, $P_k(+)$, and $\overline{K}_k$ for $k = 1, 2, \cdots$, which we can do this since the values of these expressions don't depend on the received measurements but only on the dynamic and measurement model.

Problem 4.8 (Calculating $P_k(-)$ and $P_k(+)$)

Given the system and measurement equations presented the discrete Kalman equations in this case would have $\Phi_k = 1$, $H_k = 1$, $Q_k = 30$, and $R_k = 20$. Now we can simplify our work by just performing iterations on the equations for just the covariance measurement and propagation updates. To do this we recognize that we are given $P_0(+) = P_0 = 150$ and the iterations for $k = 1, 2, 3, 4$ would be done with

$$\begin{aligned} P_k(-) &= \Phi_{k-1} P_{k-1}(+) \Phi_{k-1}^T + Q_{k-1} = P_{k-1}(+) + 30 \\ \overline{K}_k &= P_k(-) H_k^T (H_k P_k(-) H_k^T + R_k)^{-1} = \frac{P_k(-)}{P_k(-) + 20} \\ P_k(+) &= (I - \overline{K}_k H_k) P_k(-) = (1 - \overline{K}_k) P_k(-) \,. \end{aligned}$$

To compute the required values of $P_k(+)$, $P_k(-)$, and $\overline{K}_k$ for $k = 1, 2, 3, 4$ we iterate these equations. See the MATLAB script `prob_4_8.m` where this is done.

To compute $P_\infty(+)$ we put the equation for $P_k(-)$ into the equation for $P_k(+)$ to derive a recursive expression for $P_k(+)$. We find

$$\begin{aligned} P_k(+) &= (1 - \overline{K}_k) P_k(-) \\ &= \left(1 - \frac{P_k(-)}{P_k(-) + 20}\right) P_k(-) = \left(\frac{20}{P_k(-) + 20}\right) P_k(-) \\ &= \frac{20(P_{k-1}(+) + 30)}{P_{k-1}(+) + 50} \,. \end{aligned}$$

Taking the limit where $k \to \infty$ and assuming steady state conditions where $P_k(+) = P_{k-1}(+) \equiv P$ we can solve

$$P = \frac{20(P + 30)}{P + 50} \,,$$

for a positive P to determine the value of $P_\infty(+)$.

Problem 4.9 (a parameter estimation problem)

Part (a): We can solve this problem as if there is no dynamic component to the model i.e. assuming a continuous system model of $\frac{dx}{dt} = 0$ which in discrete form is given by $x_k = x_{k-1}$. To have x_k truly stationary we have no error in the dynamics i.e. the covariance matrix Q_k in the dynamic equation is taken to be zero. Thus the state and error covariance extrapolation equations are given by

$$\begin{aligned} \hat{x}_k(-) &= \hat{x}_{k-1}(+) \\ P_k(-) &= P_{k-1}(+) \,. \end{aligned}$$

Since the system and measurement equations presented in this problem have $\Phi_k = 1$, $H_k = 1$, $Q_k = 0$, and $R_k = R$, given $\hat{x}_0(+)$ and $P_0(+)$ for $k = 1, 2, \cdots$ the discrete Kalman filter would iterate

$$\begin{aligned}
\hat{x}_k(-) &= \hat{x}_{k-1}(+) \\
P_k(-) &= P_{k-1}(+) \\
\overline{K}_k &= P_k(-)H_k^T(H_k P_k(-)H_k^T + R_k)^{-1} = P_k(-)[P_k(-) + R]^{-1} \\
\hat{x}_k(+) &= \hat{x}_k(-) + \overline{K}_k(z_k - \hat{x}_k(-)) \\
P_k(+) &= (I - \overline{K}_k H_k)P_k(-) = (1 - \overline{K}_k)P_k(-) .
\end{aligned}$$

Combining these we get the following iterative equations for $\overline{K}_k$, $\hat{x}_k(+)$, and $P_k(+)$

$$\begin{aligned}
\overline{K}_k &= P_{k-1}(+)[P_{k-1}(+) + R]^{-1} \\
\hat{x}_k(+) &= \hat{x}_{k-1}(+) + \overline{K}_k(z_k - \hat{x}_{k-1}(+)) \\
P_k(+) &= (1 - \overline{K}_k)P_{k-1}(+) .
\end{aligned}$$

Part (b): If $R = 0$ we have no measurement noise and the given measurement should give all needed information about the state. The Kalman update above would predict

$$\overline{K}_1 = P_0(P_0^{-1}) = I ,$$

so that

$$\hat{x}_1(+) = x_0 + I(z_1 - x_0) = z_1 ,$$

thus the first measurement gives the entire estimate of the state and would be exact (since there is no measurement noise).

Part (c): If $R = \infty$ we have infinite measurement noise and the measurement of z_1 should give almost no information on the state x_1. When $R = \infty$ we find the Kalman gain given by $\overline{K}_1 = 0$ so that

$$\hat{x}_1(+) = x_0 ,$$

i.e. the measurement does not change our initial estimate of what x is.

Problem 4.10 (calculating $\overline{K}(t)$)

Part (a): The mean squared estimation error, $P(t)$, satisfies Equation 121 which for this system since $F(t) = -1$, $H(t) = 1$, the measurement noise covariance $R(t) = 20$ and the dynamic noise covariance matrix $Q(t) = 30$ becomes (with $G(t) = 1$)

$$\frac{dP(t)}{dt} = -P(t) - P(t) - \frac{P(t)^2}{20} + 30 = -2P(t) - \frac{P(t)^2}{20} + 30 ,$$

which we can solve. For this problem since it is a scalar-time invariance problem the solution to this differential equation can be obtained as in the book by performing a fractional decomposition. Once we have the solution for $P(t)$ we can calculate $\overline{K}(t)$ from

$$\overline{K}(t) = P(t)H^t R^{-1} = \frac{1}{20}P(t) .$$

Problem 4.11 (the Riccati equation implies symmetry)

In Equation 71 since $P_k(-)$ and R_k are both symmetric covariance matrices, the matrix $P_k(+)$ will be also. In Equation 121, since $P(t_0)$ is symmetric since it represents the initial state covariance matrix, the right hand side of this expression is symmetric. Thus $\dot{P}(t)^T = \dot{P}(t)$ and the continuous matrix $P(t)$ must therefor be symmetric for all times.

Problem 4.12 (observability of a time-invariant system)

The discrete observability matrix M for time-invariant systems is given by

$$M = \left[\begin{array}{ccccc} H^T & \Phi^T H^T & (\Phi^T)^2 H^T & \cdots & (\Phi^T)^{n-1} H^T \end{array} \right] , \tag{92}$$

and must have rank n for the given system to be observable. Note that this matrix can sometimes be more easily constructed (i.e. in Mathematica) by first constructing the transpose of M. We have

$$M^T = \left[\begin{array}{c} H \\ H\Phi \\ H\Phi^2 \\ \vdots \\ H\Phi^{n-1} \end{array} \right] .$$

Now for Example 4.4 we have the dimension on the state space $n = 6$, with Φ and H given by

$$\Phi = \left[\begin{array}{cccccc} 1 & T & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & \rho & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & T & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & r \end{array} \right] \quad \text{and} \quad H = \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array} \right] .$$

From these, the observability matrix M is given by

$$M = \left[\begin{array}{cccccccccccc} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & T & 0 & 2T & 0 & 3T & 0 & 4T & 0 & 5T & 0 \\ 0 & 0 & 0 & 0 & T & 0 & (2+\rho)T & 0 & M_{3,9} & 0 & M_{3,11} & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & T & 0 & 2T & 0 & 3T & 0 & 4T & 0 & 5T \\ 0 & 0 & 0 & 0 & 0 & T & 0 & (2+r)T & 0 & M_{6,10} & 0 & M_{6,12} \end{array} \right] ,$$

with components

$$\begin{aligned} M_{39} &= (3 + 2\rho + \rho^2)T \\ M_{3,11} &= (4 + 3\rho + 2\rho^2 + \rho^3)T \\ M_{6,10} &= (3 + 2r + r^2)T \\ M_{6,12} &= (4 + 3r + 2r^2 + r^3)T , \end{aligned}$$

which can be shown to have rank of six showing that this system is observable. For Problem 4.6 we have the dimension on the state space $n = 4$, with Φ and H given by

$$\Phi = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

From these components, the observability matrix M is given by

$$M = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & T & 0 & 2T & 0 & 3T & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & T & 0 & 2T & 0 & 3T \end{bmatrix},$$

which can be shown to have rank of six showing that this system is observable. The algebra for these examples was done in the Mathematica file `prob_4_12.nb`.

Problem 4.13 (a time varying measurement noise variance R_k)

For the given system we have $\Phi_{k-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $Q_k = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $H_k = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $R_k = 2+(-1)^k$. Then with $P_0 = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$ to evaluate $P_k(+)$, $P_k(-)$, and $\overline{K}_k$ we take $P_0(+) = P_0$ and for $k = 1, 2, \cdots$ iterate the following equations

$$\begin{aligned}
P_k(-) &= \Phi_k P_{k-1}(+)\Phi_k^T + Q_k \\
&= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} P_{k-1}(+) \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\
\overline{K}_k &= P_k(-)H_k^T(H_k P_k(-)H_k^T + R_k)^{-1} \\
&= P_k(-) \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left[\begin{bmatrix} 1 & 0 \end{bmatrix} P_k(-) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (2 + (-1)^k) \right]^{-1} \\
P_k(+) &= (I - \overline{K}_k H_k)P_k(-) \\
&= \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \overline{K}_k \begin{bmatrix} 1 & 0 \end{bmatrix} \right) P_k(-) .
\end{aligned}$$

Chapter 5: Nonlinear Applications

Notes On The Text

Notes on Table 5.3: The Discrete *Linearized* Filter Equations

Since I didn't see this equation derived in the book, in this section of these notes we derive the "predicted perturbation from the measurement" equation which is given in Table 5.3 in the book. The normal discrete Kalman state estimate observational update when we Taylor expand about x_k^{nom} can be written as

$$\begin{aligned}\hat{x}_k(+) &= \hat{x}_k(-) + \overline{K}_k(z_k - h(\hat{x}_k(-))) \\ &= \hat{x}_k(-) + \overline{K}_k(z_k - h(x_k^{\text{nom}} + \widehat{\delta x}_k(-))) \\ &\approx \hat{x}_k(-) + \overline{K}_k(z_k - h(x_k^{\text{nom}}) - H_k^{[1]}\widehat{\delta x}_k(-))\,.\end{aligned}$$

But the perturbation definition $\widehat{\delta x}_k(+) = \hat{x}_k(+) - x_k^{\text{nom}}$, means that $\hat{x}_k(+) = x_k^{\text{nom}} + \widehat{\delta x}_k(+)$ and we have

$$x_k^{\text{nom}} + \widehat{\delta x}_k(+) = x_k^{\text{nom}} + \widehat{\delta x}_k(-) + \overline{K}_k(z_k - h(x_k^{\text{nom}}) - H_k^{[1]}\widehat{\delta x}_k(-))\,,$$

or canceling the value of x_k^{nom} from both sides we have

$$\widehat{\delta x}_k(+) = \widehat{\delta x}_k(-) + \overline{K}_k(z_k - h(x_k^{\text{nom}}) - H_k^{[1]}\widehat{\delta x}_k(-))\,, \tag{93}$$

which is the predicted perturbation update equation presented in the book.

Notes on Example 5.1: Linearized Kalman and Extended Kalman Filter Equations

In this section of these notes we provide more explanation and derivations on Example 5.1 from the book which computes the linearized and the extended Kalman filtering equations for a simple discrete scalar non-linear problem. We first derive the linearized Kalman filter equations and then the extended Kalman filtering equations.

For $x_k^{\text{nom}} = 2$ the linearized Kalman filtering have their state $\hat{x}_k(+)$ determined from the perturbation $\widehat{\delta x}_k(+)$ by

$$\hat{x}_k(+) = \hat{x}_k^{\text{nom}} + \widehat{\delta x}_k(+) = 2 + \widehat{\delta x}_k(+)\,.$$

Linear prediction of the state perturbation becomes

$$\widehat{\delta x}_k(-) = \Phi_{k-1}^{[1]}\widehat{\delta x}_{k-1}(+) = 4\widehat{\delta x}_{k-1}(+)\,,$$

since

$$\Phi_{k-1}^{[1]} = \left.\frac{d{x_{k-1}}^2}{dx_{k-1}}\right|_{x_{k-1}=x_{k-1}^{\text{nom}}} = 2x_{k-1}|_{x_{k-1}=2} = 4\,.$$

The a priori covariance equation is given by

$$\begin{aligned} P_k(-) &= \Phi_k^{[1]} P_{k-1}(+) \Phi_k^{[1]T} + Q_k \\ &= 16 P_{k-1}(+) + 1 \,. \end{aligned}$$

Since $\overline{K}_k$ in the linearized Kalman filter is given by $P_k(-)H_k^{[1]T}[H_k^{[1]}P_k(-)H_k^{[1]T} + R_k]^{-1}$, we need to evaluate $H_k^{[1]}$. For this system we find

$$H_k^{[1]} = \left. \frac{d{x_k}^3}{dx_k} \right|_{x_k = x_k^{\text{nom}}} = \left. 3{x_k}^2 \right|_{x_k=2} = 12 \,.$$

With this then

$$\overline{K}_k = \frac{12 P_k(-)}{144 P_k(-) + 2} \,,$$

and we can compute the predicted perturbation conditional on the measurement

$$\widehat{\delta x}_k(+) = \widehat{\delta x}_k(-) + \overline{K}_k (z_k - h_k(x_k^{\text{nom}}) - H_k^{[1]} \widehat{\delta x}_k(-)) \,.$$

Note that $h_k(x_k^{\text{nom}}) = 2^3 = 8$ and we have

$$\widehat{\delta x}_k(+) = \widehat{\delta x}_k(-) + \overline{K}_k (z_k - 8 - 12 \widehat{\delta x}_k(-)) \,.$$

Finally, the a posteriori covariance matrix is given by

$$\begin{aligned} P_k(+) &= (1 - \overline{K}_k H_k^{[1]}) P_k(-) \\ &= (1 - 12 \overline{K}_k) P_k(-) \,. \end{aligned}$$

The extended Kalman filter equations can be derived from the steps presented in Table 5.4 in the book. For the system given here we first evaluate the needed linear approximations of $f_{k-1}(\cdot)$ and $h_k(\cdot)$

$$\begin{aligned} \Phi_{k-1}^{[1]} &= \left. \frac{\partial f_{k-1}}{\partial x} \right|_{x = \hat{x}_{k-1}(-)} = 2 \hat{x}_{k-1}(-) \\ H_k^{[1]} &= \left. \frac{\partial h_k}{\partial x} \right|_{x = \hat{x}_k(-)} = 3 \hat{x}_k(-)^2 \,. \end{aligned}$$

Using these approximations, given values for $\hat{x}_0(+)$ and $P_0(+)$ for $k = 1, 2, \cdots$ the discrete extended Kalman filter equations become

$$\begin{aligned} \hat{x}_k(-) &= f_{k-1}(\hat{x}_{k-1}(+)) = \hat{x}_{k-1}(+)^2 \\ P_k(-) &= \Phi_{k-1}^{[1]} P_{k-1}(+) \Phi_{k-1}^T + Q_{k-1} = 4 \hat{x}_{k-1}(-)^2 P_{k-1}(+) + 1 \\ \hat{z}_k &= h_k(\hat{x}_k(-)) = \hat{x}_k(-)^3 \\ \overline{K}_k &= P_k(-)(3\hat{x}_k(-)^2)(9\hat{x}_k(-)^4 P_k(-) + 2)^{-1} = \frac{3\hat{x}_k(-)^2 P_k(-)}{9\hat{x}_k(-)^4 P_k(-) + 2} \\ \hat{x}_k(+) &= \hat{x}_k(-) + \overline{K}_k (z_k - \hat{z}_k) \\ P_k(+) &= (1 - \overline{K}_k (3\hat{x}_k(-)^2)) P_k(-) = (1 - 3\overline{K}_k \hat{x}_k(-)^2) P_k(-) \,. \end{aligned}$$

Since we have an explicit formula for the state propagation dynamics we can simplify the state update equation to get

$$\begin{aligned} \hat{x}_k(+) &= \hat{x}_{k-1}(+)^2 + \overline{K}_k(z_k - \hat{x}_k(-)^3) \\ &= \hat{x}_{k-1}(+)^2 + \overline{K}_k(z_k - \hat{x}_{k-1}(+)^6) \,. \end{aligned}$$

These equations agree with the ones given in the book.

Notes on Quadratic Modeling Error

For these notes we assume that $h(\cdot)$ in our measurement equation $z = h(x)$ has the specific quadratic form given by

$$h(x) = H_1 x + x^T H_2 x + v \,.$$

Then with error $\tilde{x}$ defined as $\tilde{x} \equiv \hat{x} - x$ so that the state x in terms of our estimate $\hat{x}$ is given by $x = \hat{x} - \tilde{x}$ we can compute the expected measurement $\hat{z}$ with the following steps

$$\begin{aligned} \hat{z} &= E\langle h(x) \rangle \\ &= E\langle H_1 x + x^T H_2 x \rangle \\ &= E\langle H_1(\hat{x} - \tilde{x}) + (\hat{x} - \tilde{x})^T H_2 (\hat{x} - \tilde{x}) \rangle \\ &= H_1 \hat{x} - H_1 E\langle \tilde{x} \rangle + E\langle \hat{x}^T H_2 \hat{x} \rangle - E\langle \tilde{x}^T H_2 \hat{x} \rangle - E\langle \hat{x}^T H_2 \tilde{x} \rangle + E\langle \tilde{x}^T H_2 \tilde{x} \rangle \,. \end{aligned}$$

Now if we assume that the error $\tilde{x}$ is zero mean so that $E\langle \tilde{x} \rangle = 0$ and $\hat{x}$ is deterministic the above simplifies to

$$\hat{z} = H_1 \hat{x} + \hat{x}^T H_2 \hat{x} + E\langle \tilde{x}^T H_2 \tilde{x} \rangle \,.$$

Since $\tilde{x}^T H_2 \tilde{x}$ is a scalar it equals its own trace and by the trace product permutation theorem we have

$$\begin{aligned} E\langle \tilde{x}^T H_2 \tilde{x} \rangle &= E\langle \text{trace}[\tilde{x}^T H_2 \tilde{x}] \rangle = E\langle \text{trace}[H_2 \tilde{x} \tilde{x}^T] \rangle \\ &= \text{trace}[H_2 E\langle \tilde{x} \tilde{x}^T \rangle] \,. \end{aligned}$$

To simplify this recognize that $E\langle \tilde{x} \tilde{x}^T \rangle$ is the covariance of the state error and should equal $P(-)$ thus

$$\begin{aligned} \hat{z} &= H_1 \hat{x} + \hat{x}^T H_2 \hat{x} + \text{trace}[H_2 P(-)] \\ &= h(\hat{x}) + \text{trace}[H_2 P(-)] \,, \end{aligned}$$

the expression presented in the book.

Notes on Example 5.2: Using the Quadratic Error Correction

For a measurement equation given by $z = sy + b + v$ for a state consisting of the unknowns s, b, and y we compute the matrix, H_2 in its quadratic form representation as

$$H_2 = \frac{1}{2} \left[\begin{array}{ccc} \frac{\partial^2 z}{\partial s^2} & \frac{\partial^2 z}{\partial s \partial b} & \frac{\partial^2 z}{\partial s \partial y} \\ \frac{\partial^2 z}{\partial s \partial b} & \frac{\partial^2 z}{\partial b^2} & \frac{\partial^2 z}{\partial b \partial y} \\ \frac{\partial^2 z}{\partial s \partial y} & \frac{\partial^2 z}{\partial b \partial y} & \frac{\partial^2 z}{\partial y^2} \end{array} \right] = \frac{1}{2} \left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{array} \right] ,$$

therefore the expected measurement $h(\hat{z})$ can be *corrected* at each Kalman step by adding the term

$$\text{trace}\left\{\begin{bmatrix} 0 & 0 & 1/2 \\ 0 & 0 & 0 \\ 1/2 & 0 & 0 \end{bmatrix} P(-)\right\}.$$

Problem Solutions

Problem 5.1 (deriving the linearized and the extended Kalman estimator)

For this problem our non-linear dynamical equation is given by

$$x_k = -0.1x_{k-1} + \cos(x_{k-1}) + w_{k-1}\,, \tag{94}$$

and our non-linear measurement equation is given by

$$z_k = x_k^2 + v_k\,. \tag{95}$$

We will derive the equation for the linearized perturbed trajectory and the equation for the predicted perturbation given the measurement first and then list the full set of discrete Kalman filter equations that would be iterated in an implementation. If our nominal trajectory $x_k^{\text{nom}} = 1$, then the linearized Kalman estimator equations becomes

$$\begin{aligned}
\widehat{\delta x}_k(-) &\approx \left.\frac{\partial f_{k-1}}{\partial x}\right|_{x=x_{k-1}^{\text{nom}}} \widehat{\delta x}_{k-1}(+) + w_{k-1} \\
&= (-0.1 - \sin(x_{k-1}^{\text{nom}}))\widehat{\delta x}_{k-1}(+) + w_{k-1} \\
&= (-0.1 - \sin(1))\widehat{\delta x}_{k-1}(+) + w_{k-1}\,,
\end{aligned}$$

with an predicted a priori covariance matrix given by

$$\begin{aligned}
P_k(-) &= \Phi_{k-1}^{[1]} P_{k-1}(+) \Phi_{k-1}^{[1]\,T} + Q_{k-1} \\
&= (0.1 + \sin(1))^2 P_{k-1}(+) + 1\,.
\end{aligned}$$

The linear measurement prediction equation becomes

$$\begin{aligned}
\widehat{\delta x}_k(+) &= \widehat{\delta x}_k(-) + \overline{K}_k[z_k - h_k(x_k^{\text{nom}}) - H_k^{[1]}\widehat{\delta x}_k(-)] \\
&= \widehat{\delta x}_k(-) + \overline{K}_k\left[z_k - (1^2) - \left(\left.\frac{\partial h_k}{\partial x}\right|_{x_k^{\text{nom}}}\right)\widehat{\delta x}_k(-)\right] \\
&= \widehat{\delta x}_k(-) + \overline{K}_k[z_k - 1 - 2\widehat{\delta x}_k(-)]\,.
\end{aligned}$$

where the Kalman gain $\overline{K}_k$ is given by

$$\overline{K}_k = P_k(-)(2)\left(4P_k(-) + \frac{1}{2}\right)^{-1}.$$

and a posteriori covariance matrix, $P_k(+)$, given by

$$P_k(+) = (1 - 2\overline{K}_k)P_k(-)\,.$$

With all of these components the iterations needed to perform discrete Kalman filtering algorithm are then given by

- Pick/specify $\hat{x}_0(+)$ and $P_0(+)$ say $\hat{x}_0(+) = 0$ and $P_0(+) = 1$.
- Compute $\widehat{\delta x}_0(+) = \hat{x}_0(+) - x_0^{\text{nom}} = 0 - 1 = -1$.
- Set $k = 1$ and begin iterating
- State/Covariance propagation from step $k-1$ to step k
  - $\widehat{\delta x}_k(-) = (-0.1 - \sin(1))\widehat{\delta x}_{k-1}(+)$
  - $P_k(-) = (0.1 + \sin(1))^2 P_{k-1}(+) + 1$
- The measurement update:

$$\begin{aligned} \overline{K}_k &= 2P_k(-)\left(4P_k(-) + \frac{1}{2}\right)^{-1} \\ \widehat{\delta x}_k(+) &= \widehat{\delta x}_k(-) + \overline{K}_k(z_k - 1 - 2\widehat{\delta x}_k(-)) \\ P_k(+) &= (1 - 2\overline{K}_k)P_k(-) \end{aligned}$$

Now consider the extended Kalman filter (EKF) for this problem. The only thing that changes between this and the linearized formulation above is in the state prediction equation and the innovation update equation. Thus in implementing the extended Kalman filter we have the following algorithm (changes from the previous algorithm are shown in bold)

- Pick/specify $\hat{x}_0(+)$ and $P_0(+)$ say $\hat{x}_0(+) = 0$ and $P_0(+) = 1$.
- Set $k = 1$ and begin iterating
- State/Covariance propagation from step $k-1$ to step k
  - $\mathbf{\hat{x}_k}(-) = \mathbf{-0.1\hat{x}_{k-1}}(+) + \cos(\mathbf{\hat{x}_{k-1}}(+))$
  - $P_k(-) = (0.1 + \sin(1))^2 P_{k-1}(+) + 1$
- The measurement update:
  - $\overline{K}_k = 2P_k(-)\left(4P_k(-) + \frac{1}{2}\right)^{-1}$
  - $\mathbf{\hat{x}_k}(+) = \mathbf{\hat{x}_k}(-) + \mathbf{\overline{K}_k(z_k - \hat{x}_k}(-)^{\mathbf{2}})$
  - $P_k(+) = (1 - 2\overline{K}_k)P_k(-)$

Problem 5.2 (continuous linearized and extended Kalman filters)

To compute the continuous linearized Kalman estimator equations we recall that when the dynamics and measurement equations are given by

$$\begin{aligned}
\dot{x}(t) &= f(x(t),t) + G(t)w(t) \\
z(t) &= h(x(t),t) + v(t)\,,
\end{aligned}$$

that introducing the variables

$$\begin{aligned}
\delta x(t) &= x(t) - x^{\text{nom}}(t) \\
\delta z(t) &= z(t) - h(x^{\text{nom}}(t),t)\,,
\end{aligned}$$

representing perturbations from a nominal trajectory the linearized differential equations for δx and δz are given by

$$\begin{aligned}
\dot{\delta x}(t) &= \left(\left. \frac{\partial f(x(t),t)}{\partial x(t)} \right|_{x(t)=x^{\text{nom}}(t)} \right) \delta x(t) + G(t)w(t) \\
&= F^{[1]}\delta x(t) + G(t)w(t) \qquad (96) \\
\delta z(t) &= \left(\left. \frac{\partial h(x(t),t)}{\partial x(t)} \right|_{x(t)=x^{\text{nom}}(t)} \right) \delta x(t) + v(t) \\
&= H^{[1]}\delta x(t) + v(t)\,. \qquad (97)
\end{aligned}$$

Using these two equations for the system governed by $\delta x(t)$ and $\delta z(t)$ we can compute an estimate for $\delta x(t)$, denoted $\delta\hat{x}(t)$, using the continuous Kalman filter equations from Chapter 4 by solving (these are taken from the summary section from Chapter 4 but specified to the system above)

$$\begin{aligned}
\frac{d}{dt}\delta\hat{x}(t) &= F^{[1]}\delta\hat{x}(t) + \overline{K}(t)[\delta z(t) - H^{[1]}\delta\hat{x}(t)] \\
\overline{K}(t) &= P(t)H^{[1]T}(t)R^{-1}(t) \\
\frac{d}{dt}P(t) &= F^{[1]}P(t) + P(t)F^{[1]T} - \overline{K}(t)R(t)\overline{K}^T(t) + G(t)Q(t)G(t)^T\,.
\end{aligned}$$

For this specific problem formulation we have the linearized matrices $F^{[1]}$ and $H^{[1]}$ given by

$$\begin{aligned}
F^{[1]} &= \left. \frac{\partial}{\partial x(t)}(-0.5x^2(t)) \right|_{x(t)=x^{\text{nom}}(t)} = -x^{\text{nom}}(t) = -1 \\
H^{[1]} &= \left. \frac{\partial}{\partial x(t)}(x^3(t)) \right|_{x(t)=x^{\text{nom}}(t)} = 3\,x(t)^2\big|_{x(t)=x^{\text{nom}}} = 3\,.
\end{aligned}$$

Using $R(t) = 1/2$, $Q(t) = 1$, $G(t) = 1$ we thus obtain the Kalman-Bucy equations of

$$\begin{aligned}
\frac{d}{dt}\delta\hat{x}(t) &= -\delta\hat{x}(t) + \overline{K}(t)[z(t) - h(x^{\text{nom}}(t),t) - 3\delta\hat{x}(t)] \\
&= -\delta\hat{x}(t) + \overline{K}(t)[z(t) - 1 - 3\delta\hat{x}(t)] \\
\overline{K}(t) &= 6P(t) \\
\frac{d}{dt}P(t) &= -P(t) - P(t) - \frac{1}{2}\overline{K}(t)^2 + 1\,,
\end{aligned}$$

which would be solved for $\delta\hat{x}(t)$ and $P(t)$ as measurements $z(t)$ come in.

For the extended Kalman filter we only change the dynamic equation in the above. Thus we are requested to solve the following Kalman-Bucy system (these are taken from Table 5.5 in this chapter)

$$\begin{aligned}
\frac{d}{dt}\hat{x}(t) &= f(\hat{x}(t),t) + \overline{K}(t)[z(t) - h(\hat{x}(t),t)] \\
\overline{K}(t) &= P(t)H^{[1]T}(t)R^{-1}(t) \\
\frac{d}{dt}P(t) &= F^{[1]}P(t) + P(t)F^{[1]T} - \overline{K}(t)R(t)\overline{K}^T(t) + G(t)Q(t)G(t)^T .
\end{aligned}$$

Where now

$$\begin{aligned}
F^{[1]} &= \frac{\partial}{\partial x(t)}(f(x(t),t))\bigg|_{x(t)=\hat{x}(t)} = -\hat{x} \\
H^{[1]} &= \frac{\partial}{\partial x(t)}(h(x(t),t))\bigg|_{x(t)=\hat{x}(t)} = 3\hat{x}^2(t) .
\end{aligned}$$

Again with $R(t) = 1/2$, $Q(t) = 1$, $G(t) = 1$ we obtain the Kalman-Bucy equations of

$$\begin{aligned}
\frac{d}{dt}\hat{x}(t) &= -\frac{1}{2}\hat{x}(t)^2 + \overline{K}(t)[z(t) - \hat{x}(t)^3] \\
\overline{K}(t) &= 6P(t)\hat{x}^2(t) \\
\frac{d}{dt}P(t) &= -\hat{x}(t)P(t) - P(t)\hat{x}(t) - \frac{\overline{K}(t)^2}{2} + 1 .
\end{aligned}$$

Problem 5.4 (deriving the linearized and the extended Kalman estimator)

For this problem we derive the linearized Kalman filter for the state propagation equation

$$x_k = f(x_{k-1}, k-1) + Gw_{k-1} , \tag{98}$$

and measurement equation

$$z_k = h(x_k, k) + v_k . \tag{99}$$

We need the definitions

$$\begin{aligned}
\Phi_{k-1}^{[1]} &\equiv \frac{\partial f_{k-1}}{\partial x}\bigg|_{x=x_{k-1}^{\text{nom}}} = f'(x_{k-1}^{\text{nom}}, k-1) \\
H_k^{[1]} &\equiv \frac{\partial h_k}{\partial x}\bigg|_{x=x_k^{\text{nom}}} = h'(x_k^{\text{nom}}, k) .
\end{aligned}$$

Then the linearized Kalman filter algorithm is given by the following steps:

- Pick/specify $\hat{x}_0(+)$ and $P_0(+)$.

- Compute $\widehat{\delta x}_0(+) = \hat{x}_0(+) - x_0^{\text{nom}}$, using these values.
- Set $k = 1$ and begin iterating:
- State/Covariance propagation from $k-1$ to k

$$
\begin{aligned}
\widehat{\delta x}_k(-) &= f'(x_{k-1}^{\text{nom}}, k-1)\widehat{\delta x}_{k-1}(+) \\
P_k(-) &= f'(x_{k-1}^{\text{nom}}, k-1)^2 P_{k-1}(+) + GQ_{k-1}G^T \,.
\end{aligned}
$$

- The measurement update:

$$
\begin{aligned}
\overline{K}_k &= P_k(-)H_k^{[1]}(H_k^{[1]}P_k(-)H_k^{[1]T} + R_k)^{-1} \\
&= h'(x_k^{\text{nom}}, k)P_k(-)(h'(x_k^{\text{nom}}, k)^2 P_k(-) + R_k)^{-1} \\
\widehat{\delta x}_k(+) &= \widehat{\delta x}_k(-) + \overline{K}_k(z_k - h(x_k^{\text{nom}}, k) - h'(x_k^{\text{nom}}, k)\widehat{\delta x}_k(-)) \\
P_k(+) &= (1 - h'(x_k^{\text{nom}}, k)\overline{K}_k)P_k(-) \,.
\end{aligned}
$$

Next we compute the extended Kalman filter (EKF) for this system

- Pick/specify $\hat{x}_0(+)$ and $P_0(+)$.
- Set $k = 1$ and begin iterating
- State propagation from $k-1$ to k

$$
\begin{aligned}
x_k(-) &= f(\hat{x}_{k-1}(+), k-1) \\
P_k(-) &= f'(\hat{x}_{k-1}(+), k-1)^2 P_{k-1}(+) + GQ_{k-1}G^T \,.
\end{aligned}
$$

- The measurement update:

$$
\begin{aligned}
\overline{K}_k &= h'(\hat{x}_k(-), k)P_k(-)(h'(\hat{x}_k(-), k)^2 P_k(-) + R_k)^{-1} \\
\hat{x}_k(+) &= \hat{x}_k(-) + \overline{K}_k(z_k - h(\hat{x}_k(-), k)) \\
P_k(+) &= (1 - h'(\hat{x}_k(-), k)\overline{K}_k)P_k(-) \,.
\end{aligned}
$$

Problem 5.5 (parameter estimation via a non-linear filtering)

We can use non-linear Kalman filtering to derive an estimate the value for the parameter a in the plant model in the same way the book estimated the driving parameter ζ in example 5.3. To do this we consider introducing an additional state $x_2(t) = a$, which since a is a constant has a very simple dynamic equation $\dot{x}_2(t) = 0$. Then the total linear system when we take $x_1(t) \equiv x(t)$ then becomes

$$
\frac{d}{dt}\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} x_2(t)x_1(t) \\ 0 \end{bmatrix} + \begin{bmatrix} w(t) \\ 0 \end{bmatrix} ,
$$

which is non-linear due to the product $x_1(t)x_2(t)$. The measurement equation is

$$z(t) = x_1(t) + v(t) = \left[\begin{array}{cc} 1 & 0 \end{array} \right] \left[\begin{array}{c} x_1(t) \\ x_2(t) \end{array} \right] + v(t) \,.$$

To derive an estimator for a we will use the extended Kalman filter (EKF) equations to derive estimates of $x_1(t)$ and $x_2(t)$ and then the limiting value of the estimate of $x_2(t)$ will be the value of a we seek. In extended Kalman filtering we need

$$\begin{aligned}
F^{[1]} &= \left. \frac{\partial}{\partial x(t)}(f(x(t),t)) \right|_{x(t)=\hat{x}(t)} = \left[\begin{array}{cc} \hat{x}_2(t) & \hat{x}_1(t) \\ 0 & 0 \end{array} \right] \\
H^{[1]} &= \left. \frac{\partial}{\partial x(t)}(h(x(t),t)) \right|_{x(t)=\hat{x}(t)} = \left[\begin{array}{cc} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} \end{array} \right] = \left[\begin{array}{cc} 1 & 0 \end{array} \right] .
\end{aligned}$$

Then the EKF estimate $\hat{x}(t)$ is obtained by recognizing that for this problem $R(t) = 2$, $G(t) = I$, and $Q = \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right]$ and solving the following coupled dynamical system (see table 5.5 from the book)

$$\begin{aligned}
\frac{d}{dt} \left[\begin{array}{c} \hat{x}_1(t) \\ \hat{x}_2(t) \end{array} \right] &= \left[\begin{array}{c} \hat{x}_1(t)\hat{x}_2(t) \\ 0 \end{array} \right] + \overline{K}(t)(z(t) - \hat{x}_1(t)) \\
\overline{K}(t) &= \frac{1}{2} P(t) \left[\begin{array}{c} 1 \\ 0 \end{array} \right] \\
\frac{d}{dt} P(t) &= \left[\begin{array}{cc} \hat{x}_2(t) & \hat{x}_1(t) \\ 0 & 0 \end{array} \right] P(t) + P(t) \left[\begin{array}{cc} \hat{x}_2(t) & 0 \\ \hat{x}_1(t) & 0 \end{array} \right] + \left[\begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right] - 2\overline{K}(t)\overline{K}(t)^T ,
\end{aligned}$$

here $P(t)$ is a two by two matrix with three unique elements (recall $P_{12}(t) = P_{21}(t)$ since $P(t)$ is a symmetric matrix).

Problem 5.9 (the linearized Kalman filter for a space vehicle)

To apply the Kalman filtering framework we need to first write the second order differential equation as a first order system. If we try the state-space representation given by

$$\mathbf{x}(t) = \left[\begin{array}{c} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{array} \right] = \left[\begin{array}{c} r \\ \dot{r} \\ \theta \\ \dot{\theta} \end{array} \right] ,$$

then our dynamical system would then become

$$\dot{x}(t) = \left[\begin{array}{c} \dot{r} \\ \ddot{r} \\ \dot{\theta} \\ \ddot{\theta} \end{array} \right] = \left[\begin{array}{c} x_2 \\ r\dot{\theta}^2 - \frac{k}{r^2} + w_r(t) \\ x_4 \\ -\frac{2\dot{r}}{r}\dot{\theta} - \frac{w_\theta(t)}{r} \end{array} \right] = \left[\begin{array}{c} x_2 \\ x_1 x_4^2 - \frac{k}{x_1^2} \\ x_4 \\ -\frac{2x_2x_4}{x_1} \end{array} \right] + \left[\begin{array}{c} 0 \\ w_r(t) \\ 0 \\ \frac{w_\theta(t)}{x_1} \end{array} \right] .$$

This system will not work since it has values of the state $\mathbf{x}$, namely x_1 in the noise term. Thus instead consider the state definition given by

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} = \begin{bmatrix} r \\ \dot{r} \\ \theta \\ r\dot{\theta} \end{bmatrix},$$

where only the definition of x_4 has changed from earlier. Then we have a dynamical system for this state given by

$$\begin{aligned} \frac{\mathbf{x}(t)}{dt} &= \begin{bmatrix} \dot{r} \\ \ddot{r} \\ \dot{\theta} \\ \frac{d(r\dot{\theta})}{dt} \end{bmatrix} = \begin{bmatrix} x_2 \\ r\dot{\theta}^2 - \frac{k}{r^2} + w_r \\ \frac{x_4}{r} \\ \dot{r}\dot{\theta} + r\ddot{\theta} \end{bmatrix} = \begin{bmatrix} x_2 \\ \frac{x_4^2}{x_1} - \frac{k}{x_1^2} + w_r \\ \frac{x_4}{x_1} \\ \dot{r}\dot{\theta} - 2\dot{r}\dot{\theta} + w_\theta \end{bmatrix} \\ &= \begin{bmatrix} x_2 \\ \frac{x_4^2}{x_1} - \frac{k}{x_1^2} \\ \frac{x_4}{x_1} \\ -\frac{x_2 x_4}{x_1} \end{bmatrix} + \begin{bmatrix} 0 \\ w_r \\ 0 \\ w_\theta \end{bmatrix}. \end{aligned}$$

We can apply extended Kalman filtering (EKF) to this system. Our observation equation (in terms of the components of the state $\mathbf{x}(t)$ is given by $z(t) = \begin{bmatrix} \sin^{-1}(\frac{R_e}{x_1(t)}) \\ \alpha_0 - x_3(t) \end{bmatrix}$. To linearize this system about $r_{\text{nom}} = R_0$ and $\theta_{\text{nom}} = \omega_0 t$ we have $\dot{r}_{\text{nom}} = 0$ and $\dot{\theta}_{\text{nom}} = \omega_0$ so

$$\mathbf{x}^{\text{nom}}(t) = \begin{bmatrix} R_0 \\ 0 \\ \omega_0 t \\ R_0\omega_0 \end{bmatrix}.$$

Thus to perform extended Kalman filtering we need $F^{[1]}$ given by

$$\begin{aligned} F^{[1]} &= \left.\frac{\partial f(x(t),t)}{\partial x(t)}\right|_{x(t)=x^{\text{nom}}(t)} = \left.\begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{x_4{}^2}{x_1{}^2} + \frac{2k}{x_1{}^3} & 0 & 0 & \frac{2x_4}{x_1} \\ -\frac{x_4}{x_1{}^2} & 0 & 0 & \frac{1}{x_1} \\ \frac{x_2 x_4}{x_1{}^2} & -\frac{x_4}{x_1} & 0 & -\frac{x_2}{x_1} \end{bmatrix}\right|_{x(t)=x^{\text{nom}}(t)} \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\omega_0^2 + + \frac{2k}{R_0^3} & 0 & 0 & 2\omega_0 \\ -\frac{\omega_0}{R_0} & 0 & 0 & \frac{1}{R_0} \\ 0 & -\omega_0 & 0 & 0 \end{bmatrix}, \end{aligned}$$

and $H^{[1]}$ given by

$$\begin{aligned} H^{[1]} &= \left.\frac{\partial h(x(t),t)}{\partial x(t)}\right|_{x(t)=x^{\text{nom}}(t)} \\ &= \left.\begin{bmatrix} \frac{1}{\sqrt{1-(R_e/x_1)^2}}\left(-\frac{R_e}{x_1^2}\right) & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}\right|_{x(t)=x^{\text{nom}}(t)} \\ &= \begin{bmatrix} \left(-\frac{R_e}{R_0{}^2}\right)\frac{1}{\sqrt{1-(R_e/R_0)^2}} & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}. \end{aligned}$$

These two expressions would then be used in Equations 96 and 97.

Chapter 6: Implementation Methods

Notes On The Text

Example 6.2: The Effects of round off

Consider the given measurement sensitivity matrix H and initial covariance matrix P_0 supplied in this example. We have in infinite arithmetic and then truncated by dropping the term δ^2 since $\delta^2 < \varepsilon_{\text{roundoff}}$ the product HP_0H^T given by

$$\begin{aligned} HP_0H^T &= \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1+\delta \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1+\delta \end{bmatrix} \\ &= \begin{bmatrix} 3 & 3+\delta \\ 3+\delta & 2+(1+\delta)^2 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 3+\delta \\ 3+\delta & 2+1+2\delta+\delta^2 \end{bmatrix} \\ &\approx \begin{bmatrix} 3 & 3+\delta \\ 3+\delta & 3+2\delta \end{bmatrix} . \end{aligned}$$

If we assume our measurement covariance R is taken to be $R = \delta^2 I$ then adding R to HP_0H^T (as required in computing the Kalman gain $\overline{K}$) does not change the value of HP_0H^T. The problem is that due to roundoff error $HP_0H^T + R \approx HP_0H^T$, which is numerically singular which can be seen by computing the determinant of the given expression. We find

$$|HP_0H^T| = 9 + 6\delta - 9 - 6\delta - \delta^2 \approx 0 \,,$$

when rounded. Thus the inversion of $HP_0H^T + R$ needed in computing the Kalman gain will fail even though the problem as stated in infinite precision is non-singular.

Efficient computation of the expression $(HPH^T + R)^{-1}H$

To compute the value of the expression $(HPH^T + R)^{-1}H$ as required in the Kalman gain we will consider a "modified" Cholesky decomposition of $HPH^T + R$ where by it is written as the product of three matrices as

$$HPH^T + R = UDU^T \,,$$

then by construction the matrix product UDU^T is the inverse of $(HPH^T + R)^{-1}$ we have

$$UDU^T(HPH^T + R)^{-1}H = H \,.$$

Defining the expression we desire to evaluate as X so that $X \equiv (HPH^T + R)^{-1}H$ then we have $UDU^TX = H$. Now the stepwise procedure used to compute X comes from grouping this matrix product as

$$U(D(U^TX)) = H \,.$$

Now define $X_{[1]}$ as $X_{[1]} \equiv D(U^T X)$, and we begin by solving $UX_{[1]} = H$, for $X_{[1]}$. This is relatively easy to do since U is upper triangular. Next defining $X_{[2]}$ as $X_{[2]} \equiv U^T X$ the equation for $X_{[2]}$ is given by $DX_{[2]} = X_{[1]}$, we we can easily solve for $X_{[2]}$, since D is diagonal. Finally, recalling how $X_{[2]}$ was defined, as $U^T X = X_{[2]}$, since we have just computed $X_{[2]}$ we solve this equation for the desired matrix X.

Householder reflections along a single coordinate axis

In this subsection we duplicate some of the algebraic steps derived in the book that show the process of triangulation using Householder reflections. Here x is a *row* vector and v a column vector given by $v = x^T + \alpha e_k^T$ so that

$$v^T v = |x|^2 + 2\alpha x_k + \alpha^2 ,$$

and the inner product xv is

$$xv = x(x^T + \alpha e_k^T) = |x|^2 + \alpha x_k .$$

so the Householder transformation $T(v)$ is then given by

$$T(v) = I - \frac{2}{v^T v} v v^T = I - \frac{2}{(|x|^2 + 2\alpha x_k + \alpha^2)} v v^T .$$

Using this we can compute the Householder reflection of x or $xT(v)$ as

$$\begin{aligned} xT(v) &= x - \frac{2xv}{(|x|^2 + 2\alpha x_k + \alpha^2)} v^T \\ &= x - \frac{2(|x|^2 + \alpha x_k)}{(|x|^2 + 2\alpha x_k + \alpha^2)} (x + \alpha e_k) \\ &= \left[\frac{\alpha^2 - |x|^2}{|x|^2 + 2\alpha x_k + \alpha^2} \right] x - \left[\frac{2\alpha(|x|^2 + \alpha x_k)}{|x|^2 + 2\alpha x_k + \alpha^2} \right] e_k . \end{aligned}$$

In triangularization, our goal is to map x (under $T(v)$) so that the product $xT(v)$ is a multiple of e_k. Thus if we let $\alpha = \mp|x|$, then we see that the coefficient in front of x above vanishes and $xT(v)$ becomes a multiple of e_k as

$$xT(v) = \pm \frac{2|x|(|x|^2 \mp |x| x_k)}{|x|^2 \mp 2|x| x_k + |x|^2} e_k = \pm |x| e_k .$$

This specific result is used to zero all but one of the elements in a given row of a matrix M. For example, if in block matrix form our matrix M has the form $M = \begin{bmatrix} Z \\ x \end{bmatrix}$, so that x is the bottom row and Z represents the rows above x when we pick $\alpha = -|x|$ and form the vector $v = x^T + \alpha e_k$ (and the corresponding Householder transformation matrix $T(v)$) we find that the product $MT(v)$ is given by

$$MT(v) = \begin{bmatrix} ZT(v) \\ xT(v) \end{bmatrix} = \begin{bmatrix} & & & ZT(v) & & \\ 0 & 0 & 0 & \cdots & 0 & |x| \end{bmatrix} ,$$

showing that the application of $T(v)$ has been able to achieve the first step at upper triangularizing the matrix M.

Notes on Carlson-Schmidt square-root filtering

We begin with the stated matrix identity in that if W is the Cholesky factor of the rank one modification of the identity as

$$WW^T = I - \frac{vv^T}{R + |v|^2}$$

then

$$\sum_{k=m}^{j} W_{ik} W_{mk} = \Delta_{im} - \frac{v_i v_m}{R + \sum_{k=1}^{j} v_k^2}\,, \tag{100}$$

for all $1 \leq i \leq m \leq j \leq n$. Now if we take $m = j$ in this expression we have

$$W_{ij} W_{jj} = \Delta_{ij} - \frac{v_i v_j}{R + \sum_{k=1}^{j} v_k^2}\,.$$

If we first consider the case where $i = j$ we have

$$W_{jj}^2 = 1 - \frac{v_j^2}{R + \sum_{k=1}^{j} v_k^2} = \frac{R + \sum_{k=1}^{j} v_k^2 - v_j^2}{R + \sum_{k=1}^{j} v_k^2}$$

or

$$W_{jj} = \sqrt{\frac{R + \sum_{k=1}^{j-1} v_k^2}{R + \sum_{k=1}^{j} v_k^2}}\,.$$

When $i < j$ then we have

$$W_{ij} W_{jj} = 0 - \frac{v_i v_j}{R + \sum_{k=1}^{j} v_k^2}\,,$$

so that with the value of W_{jj} we found above we find

$$W_{ij} = -\left(\frac{v_i v_j}{R + \sum_{k=1}^{j} v_k^2}\right) \frac{\sqrt{R + \sum_{k=1}^{j} v_k^2}}{\sqrt{R + \sum_{k=1}^{j-1} v_k^2}} = -\frac{v_i v_j}{\sqrt{\left(R + \sum_{k=1}^{j} v_k^2\right)\left(R + \sum_{k=1}^{j-1} v_k^2\right)}}\,.$$

when $i < j$. Note that this result is slightly different than what the book has in that the square root is missing the the books result. Since W is upper triangular $W_{ij} = 0$ when $i > j$. Combining these three cases gives the expression found in equation 6.55 in the book.

Some discussion on Bierman's UD observational update

In Bierman's UD observational covariance update algorithm uses the modified Cholesky decomposition of the a-priori and a-posterori covariance matrices $P(-)$ and $P(+)$ defined as

$$P(-) \equiv U(-)D(-)U(-)^T \tag{101}$$

$$P(+) \equiv U(+)D(+)U(+)^T\,, \tag{102}$$

to derive a numerically stable way to compute $P(+)$ based on the factors $U(-)$ and $D(-)$ and the modified Cholesky factorization of an intermediate matrix (defined below). To derive these observational covariance update equations we assume that $l = 1$ i.e. we have only one measurement and recall the scalar measurement observational update equation

$$P(+) = P(-) - P(-)H^T(HP(-)H^T + R)^{-1}HP(-) = P(-) - \frac{P(-)H^THP(-)}{R + HP(-)H^T},$$

since in the scalar measurement case the matrix H is really a row vector and R is a scalar. Now using the definitions in Equations 101 and 102 this becomes

$$U(+)D(+)U(+)^T = U(-)D(-)U(-)^T - \frac{U(-)D(-)U(-)^TH^THU(-)D(-)U(-)^T}{R + HU(-)D(-)U(-)^TH^T}.$$

If we define a *vector* v as $v \equiv U^T(-)H^T$ then the above expression in terms of this vector becomes

$$\begin{aligned} U(+)D(+)U(+)^T &= U(-)D(-)U(-)^T - \frac{U(-)D(-)vv^TD(-)U^T(-)}{R + v^TD(-)v} \\ &= U(-)\left[D(-) - \frac{D(-)vv^TD(-)}{R + v^TD(-)v}\right]U(-)^T. \end{aligned}$$

The expression on the right-hand-side can be made to look exactly like a modified Cholesky factorization if we perform a modified Cholesky factorization on the expression "in the middle" or write it as

$$D(-) - \frac{D(-)vv^TD(-)}{R + v^TD(-)v} = BD(+)B^T. \quad (103)$$

When we do this we see that we have written $P(+) = U(+)D(+)U(+)^T$ as

$$U(+)D(+)U(+)^T = U(-)BD(+)B^TU(-)^T.$$

From which we see that $D(+)$ in the modified Cholesky factorization of $P(+)$ is obtained directly from the diagonal matrix in the modified Cholesky factorization of the left-hand-side of Equation 103 and the matrix $U(+)$ is obtained by computing the product $U(-)B$. These steps give the procedure for implementing the Bierman UD observational update given the a-priori modified Cholesky decomposition $P(-) = U(-)D(-)U(-)^T$, when we have scalar measurements. In steps they are

- compute the vector $v = U^T(-)H^T$.
- compute the matrix
$$D(-) - \frac{D(-)vv^TD(-)}{R + v^TD(-)v}$$
- perform the modified Cholesky factorization on this matrix i.e. Equation 103 the output of which are the matrices $D(+)$ and B.
- compute the non-diagonal factor $U(+)$ in the modified Cholesky factorization of $P(+)$ using the matrix B as $U(+) = U(-)B$.

Operation	Symmetric Implementation Flop Count	Notes
HP	$n^2 l$	$l \times n$ times $n \times n$
$H(HP)^T + R$	$\frac{1}{2}l^2 n + \frac{1}{2}l^2$	adding $l \times l$ matrix R requires $\frac{1}{2}l^2$
$\{H(HP)^T + R\}^{-1}$	$l^3 + \frac{1}{2}l^2 + \frac{1}{2}l$	cost for standard matrix inversion
$K^T = \{H(HP)^T + R\}^{-1}(HP)$	nl^2	$l \times l$ times $l \times n$
$P - (HP)^T K^T$	$\frac{1}{2}n^2 l + \frac{1}{2}n^2$	subtracting $n \times n$ requires $\frac{1}{2}n^2$
Total	$\frac{1}{2}(3l+1)n^2 + \frac{3}{2}nl^2 + l^3$	highest order terms only

Table 1: A flop count of the operations in the traditional Kalman filter implementation. Here P stands for the prior state uncertainty covariance matrix $P(-)$.

Earlier Implementation Methods: The Kalman Formulation

Since this is the most commonly implemented version of the Kalman filter it is instructive to comment some on it in this section. The first comment is that in implementing a Kalman filter using the direct equations one should always focus on the factor $HP(-)$. This factor occurs several times in the resulting equations and computing it first and then reusing this matrix product as a base expression can save computational time. The second observation follows the discussion on Page 49 where with uncorrelated measurements the vector measurement $\mathbf{z}$ is processed a l sequential scalar measurements. Under the standard assumption that H is $l \times n$ and $P(\pm)$ is a $n \times n$ matrix, in Table 1 we present a flop count of the operations requires to compute $P(+)$ given $P(-)$. This implementation uses the common factor $HP(-)$ as much as possible and the flop count takes the symmetry of the various matrices involved into account. This table is very similar to one presented in the book but uses some simplifying notation and corrects several typos.

Some discussion on Potter's square-root filter

Potters Square root filter is similar to the Bierman-Thornton UD filtering method but rather then using the *modified* Cholesky decomposition to represent the covariance matrices it uses the direct Cholesky factorization. Thus we introduce the two factorizations

$$P(-) \equiv C(-)C(-)^T \quad (104)$$

$$P(+) \equiv C(+)C(+)^T, \quad (105)$$

note there is no diagonal terms in these factorizations expressions. Then the Kalman filtering temporal update expression becomes

$$\begin{aligned} P(+) &= P(-) - P(-)H^T(HP(-)H^T + R)^{-1}HP(-) \\ &= C(-)C(-)^T - C(-)C(-)^T H^T (HC(-)C(-)^T H^T + R)^{-1} HC(-)C(-)^T \\ &= C(-)C(-)^T - C(-)V(V^T V + R)^{-1} V^T C(-)^T \\ &= C(-)\left[I - V(V^T V + R)^{-1} V^T\right] C(-)^T . \end{aligned}$$

Where in the above we have introduced the $n \times l$ matrix V as $V \equiv C(-)^T H^T$. We are able to write $P(+)$ in the required factored form expressed in Equation 105 when $l = 1$ (we have one measurement) then H is $1 \times n$ so the matrix $V = C^T(-)H^T$ is actually a $n \times 1$ *vector* say v and the "matrix in the middle" or

$$I_n - V(V^T V + R)^{-1} V^T = I_n - \frac{vv^T}{v^T v + R},$$

is a rank-one update of the $n \times n$ identity matrix I_n. To finish the development of Potters square root filter we have to find the "square root" of this rank one-update. This result is presented in the book section entitled: "symmetric square root of a symmetric elementary matrix", where we found that the square root of the matrix $I - svv^T$ is given by the matrix $I - \sigma vv^T$ with

$$\sigma = \frac{1 + \sqrt{1 - s|v|^2}}{|v|^2}. \tag{106}$$

In the application we want to use this result for we have $s = \frac{1}{v^T v + R}$ so the radicand in the expression for σ is given by

$$1 - s|v|^2 = 1 - \frac{|v|^2}{v^T v + R} = \frac{R}{|v|^2 + R}.$$

and so σ then is

$$\sigma = \frac{1 + \sqrt{R/(R + |v|^2)}}{|v|^2}.$$

Thus we have the factoring

$$I_n - \frac{vv^T}{v^T v + R} = WW^T = (I_n - \sigma vv^T)(I_n - \sigma vv^T)^T, \tag{107}$$

from which we can write the Potter factor of $P(+)$ as $C(+) = C(-)W = C(-)(I_n - \sigma vv^T)$, which is equation 6.122 in the book.

Some discussion on the Morf-Kailath combined observational/temporal update

In the Morf-Kailath combined observational temporal update we desire to take the Cholesky factorization of $P(-)$ at timestep k and produce the Cholesky factorization of $P(-)$ at the next timestep $k+1$. To do this recall that at timestep k we know directly values for G_k, Φ_k, and H_k. In addition, we can Cholesky factor the measurement covariance, R_k, the model noise covariance, and the a-priori state covariance matrix $P_k(-)$ as

$$\begin{aligned} R_k &\equiv C_{R(k)} C_{R(k)}^T \\ Q_k &\equiv C_{Q(k)} C_{Q(k)}^T \\ P_k(-) &\equiv C_{P(k)} C_{P(k)}^T. \end{aligned}$$

From all of this information we compute the block matrix A_k defined as

$$A_k = \begin{bmatrix} G_k C_{Q(k)} & \Phi_k C_{P(k)} & 0 \\ 0 & H_k C_{P(k)} & C_{R(k)} \end{bmatrix}.$$

Then notice that $A_k A_k^T$ is given by

$$
\begin{aligned}
A_k A_k^T &= \begin{bmatrix} G_k C_{Q(k)} & \Phi_k C_{P(k)} & 0 \\ 0 & H_k C_{P(k)} & C_{R(k)} \end{bmatrix} \begin{bmatrix} C_{Q(k)}^T G_k^T & 0 \\ C_{P(k)}^T \Phi_k^T & C_{P(k)}^T H_k^T \\ 0 & C_{R(k)}^T \end{bmatrix} \\
&= \begin{bmatrix} G_k Q_k G_k^T + \Phi_k P_k(-) \Phi_k^T & \Phi_k P_k(-) H_k^T \\ H_k P_k(-) \Phi_k^T & H_k P_k(-) H_k^T + R_k \end{bmatrix} . \qquad (108)
\end{aligned}
$$

Using Householder transformations or Givens rotations we will next triangulate this block matrix A_k triangulate it in the process define the matrices $C_{P(k+1)}$, Ψ_k, and $C_{E(k)}$.

$$
A_k T = C_k = \begin{bmatrix} 0 & C_{P(k+1)} & \Psi_k \\ 0 & 0 & C_{E(k)} \end{bmatrix} .
$$

Here T is the orthogonal matrix that triangulates A_k. At this point the introduced matrices: $C_{P(k+1)}$, Ψ_k, and $C_{E(k)}$ are simply names. To show that they also provide the desired Cholesky factorization of $P_{k+1}(-)$ that we seek consider the product $C_k C_k^T$

$$
\begin{aligned}
C_k C_k^T &= A_k A_k^T \\
&= \begin{bmatrix} 0 & C_{P(k+1)} & \Psi_k \\ 0 & 0 & C_{E(k)} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ C_{P(k+1)}^T & 0 \\ \Psi_k^T & C_{E(k)}^T \end{bmatrix} \\
&= \begin{bmatrix} C_{P(k+1)} C_{P(k+1)}^T + \Psi_k \Psi_k^T & \Psi_k C_{E(k)}^T \\ C_{E(k)} \Psi_k^T & C_{E(k)} C_{E(k)}^T \end{bmatrix} .
\end{aligned}
$$

Equating these matrix elements to the corresponding ones from $A_k A_k^T$ in Equation 108 we have

$$
\begin{aligned}
C_{P(k+1)} C_{P(k+1)}^T + \Psi_k \Psi_k^T &= \Phi_k P_k(-) \Phi_k^T + G_k Q_k G_k^T \qquad (109) \\
\Psi_k C_{E(k)}^T &= \Phi_k P_k(-) H_k^T \qquad (110) \\
C_{E(k)} C_{E(k)}^T &= H_k P_k(-) H_k^T + R_k . \qquad (111)
\end{aligned}
$$

These are the books equations 6.133-6.138. Now Equation 110 is equivalent to

$$
\Psi_k = \Phi_k P_k(-) H_k^T C_{E(k)}^{-T} ,
$$

so that when we use this expression Equation 109 becomes

$$
C_{P(k+1)} C_{P(k+1)}^T + \Phi_k P_k(-) H_k^T C_{E(k)}^{-T} C_{E(k)}^{-1} H_k P_k(-) \Phi_k^T = \Phi_k P_k(-) \Phi_k^T + G_k Q_k G_k^T ,
$$

or solving for $C_{P(k+1)} C_{P(k+1)}^T$

$$
C_{P(k+1)} C_{P(k+1)}^T = \Phi_k \left[P_k(-) - P_k(-) H_k^T (C_{E(k)} C_{E(k)}^T)^{-1} H_k P_k(-) \right] \Phi_k^T + G_k Q_k G_k^T .
$$

Now using Equation 111 we have that the above can be written as

$$
C_{P(k+1)} C_{P(k+1)}^T = \Phi_k \left[P_k(-) - P_k(-) H_k^T (H_k P_k(-) H_k^T + R_k)^{-1} H_k P_k(-) \right] \Phi_k^T + G_k Q_k G_k^T .
$$

The right-hand-side of this expression is equivalent to the expression $P_{k+1}(-)$ showing that $C_{P(k+1)}$ is indeed the Cholesky factor of $P_{k+1}(-)$ and proving correctness of the Morf-Kailath update procedure.

Problem Solutions

Problem 6.1 (Moler matrices)

The **Moler matrix** M is defined as

$$M_{ij} = \begin{cases} i & i = j \\ \min(i,j) & i \neq j \end{cases},$$

so the three by three Moler matrix is given by

$$M = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Using MATLAB and the `chol` command we find the Cholesky decomposition of M given by

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

or an upper-triangular matrix of all ones. In fact this makes me wonder if a Moler matrix is *defined* as the product CC^T where C is an upper-triangular matrix of all ones (see the next problem).

Problem 6.2 (more Moler matrices)

Note one can use the MATLAB command `gallery('moler',n,1)` to generate this definition of a Moler matrix. In the MATLAB script `prob_6_2.m` we call the gallery command and compute the Cholesky factorization for each resulting matrix. It appears that for the Moler matrices considered here the hypothesis presented in Problem 6.1 that the Cholesky factor of a Moler matrix is an upper triangular matrix of all ones is still supported.

Problem 6.8 (the SVD)

For C to be a Cholesky factor for P requires $P = CC^T$. Computing this product for the given expression for $C = ED^{1/2}E^T$ we find

$$CC^T = ED^{1/2}E^T(ED^{1/2}E^T) = EDE^T = P.$$

For C to be a square root of P means that $P = C^2$. Computing this product for the given expression for C gives

$$ED^{1/2}E^T(ED^{1/2}E^T) = EDE^T = P.$$

Problem 6.11 (an orthogonal transformation of a Cholesky factor)

If C is a Cholesky factor of P then $P = CC^T$. Now consider the matrix $\hat{C} = CT$ with T an orthogonal matrix. We find $\hat{C}\hat{C}^T = CTT^TC^T = CC^T = P$, showing that $\hat{C}$ is also a Cholesky factor of P.

Problem 6.12 (some matrix squares)

We have for the first product

$$\begin{aligned}(I - vv^T)^2 &= I - vv^T - vv^T + vv^T(vv^T) \\ &= I - 2vv^T + v(v^Tv)v^T \\ &= I - 2vv^T + vv^T \quad \text{if} \quad v^Tv = 1 \\ &= I - vv^T .\end{aligned}$$

Now if $|v|^2 = v^Tv = 2$ the third equation above becomes

$$I - 2vv^T + 2vv^T = I .$$

Problem 6.17 (a block orthogonal matrix)

If A is an orthogonal matrix this means that $A^TA = I$ (the same holds true for B). Now consider the product

$$\begin{aligned}\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^T \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} &= \begin{bmatrix} A^T & 0 \\ 0 & B^T \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \\ &= \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} ,\end{aligned}$$

showing that $\begin{bmatrix} A & 0 \\ B & 0 \end{bmatrix}$ is also orthogonal.

Problem 6.18 (the inverse of a Householder reflection)

The inverse of the given Household reflection matrix is the reflection matrix itself. To show this consider the required product

$$\begin{aligned}\left(I - \frac{2vv^T}{v^Tv}\right)\left(I - \frac{2vv^T}{v^Tv}\right) &= I - \frac{2vv^T}{v^Tv} - \frac{2vv^T}{v^Tv} + \frac{4vv^T(vv^T)}{(v^Tv)^2} \\ &= I - \frac{4vv^T}{v^Tv} + \frac{4vv^T}{v^Tv} = I ,\end{aligned}$$

showing that $I - \frac{2vv^T}{v^Tv}$ is its own inverse.

Problem 6.19 (the number of Householder transformations to triangulate)

Assume that $n > q$ the first Householder transformation will zero all elements in A_{ij} for A_{nk} where $1 \le k \le q-1$. The second Householder transformation will zero all elements of $A_{n-1,k}$ for $1 \le k \le q-2$. We can continue this $n-q+1$ times. Thus we require $q-1$ Householder transformations to triangulate a $n \times q$ matrix. This does not change if $n = q$.

Now assume $n < q$. We will require n Householder transformations when $n < q$. If $n = q$ the last Householder transformation is not required. Thus we require $n-1$ in this case.

Problem 6.20 (the nonlinear equation solved by $C(t)$)

Warning: There is a step below this is not correct or at least it doesn't seem to be correct for 2x2 matrices. I was not sure how to fix this. If anyone has any ideas please email me.

Consider the differential equation for the continuous covariance matrix $P(t)$ given by

$$\dot{P}(t) = F(t)P(t) + P(t)F^T(t) + G(t)Q(t)G(t)^T\,, \tag{112}$$

We want to prove that if $C(t)$ is the differentiable Cholesky factor of $P(t)$ i.e. $P(t) = C(t)C(t)^T$ then $C(t)$ are solutions to the following nonlinear equation

$$\dot{C}(t) = F(t)C(t) + \frac{1}{2}[G(t)Q(t)G^T(t) + A(t)]C^{-T}(t)\,,$$

where $A(t)$ is a skew-symmetric matrix. Since $C(t)$ is a differentiable Cholesky factor of $P(t)$ then $P(t) = C(t)C(t)^T$ and the derivative of $P(t)$ by the product rule is given by

$$\dot{P}(t) = \dot{C}(t)C(t)^T + C(t)\dot{C}(t)^T\,.$$

When this expression is put into Equation 112 we have

$$\dot{C}(t)C(t)^T + C(t)\dot{C}(t)^T = F(t)C(t)C(t)^T + C(t)C(t)^T F^T + GQG^T\,.$$

Warning: This next step does not seem to be correct.

If I could show that $\dot{C}(t)C(t)^T + C(t)\dot{C}(t)^T = 2\dot{C}(t)C(t)^T$ then I would have

$$2\dot{C}(t)C(t)^T = F(t)C(t)C(t)^T + C(t)C(t)^T F^T + GQG^T\,,$$

Thus when we solve for $\dot{C}(t)$ we find

$$\begin{aligned}\dot{C}(t) &= \frac{1}{2}F(t)C(t) + \frac{1}{2}C(t)C(t)^T F(t)^T C(t)^{-T} + \frac{1}{2}G(t)Q(t)G(t)^T C(t)^{-T} \\ &= F(t)C(t) + \frac{1}{2}\left[G(t)Q(t)G(t)^T - F(t)C(t)C(t)^T + C(t)C(t)^T F(t)^T\right] C(t)^{-T}\,.\end{aligned}$$

From this expression if we define the matrix $A(t)$ as $A(t) \equiv -F(t)C(t)C(t)^T + C(t)C(t)^T F(t)^T$ we note that

$$A(t)^T = -C(t)C(t)^T F(t)^T + F(t)C(t)C(t)^T = -A(t)\,,$$

so $A(t)$ is skew symmetric and we have the desired nonlinear differential equation for $C(t)$.

Problem 6.21 (the condition number of the information matrix)

The information matrix Y is defined as $Y = P^{-1}$. Since a matrix and its inverse have the same condition number the result follows immediately.

Problem 6.22 (the correctness of the observational triangularization in SRIF)

The observation update in the square root information filter (SRIF) is given by producing an orthogonal matrix T_{obs} that performs triangularization on the following block matrix

$$\left[\begin{array}{cc} C_{Y_k(-)} & H_k^T C_{R_k^{-1}} \\ \hat{s}_k^T(-) & z_k^T C_{R_k^{-1}} \end{array}\right] T_{\text{obs}} = \left[\begin{array}{cc} C_{Y_k(+)} & 0 \\ \hat{s}_k^T(+) & \varepsilon \end{array}\right].$$

Following the hint given for this problem we take the product of this expression and its own transpose. We find

$$\left[\begin{array}{cc} C_{Y_k(+)} & 0 \\ \hat{s}_k^T(+) & \varepsilon \end{array}\right] \left[\begin{array}{cc} C_{Y_k(+)}^T & \hat{s}_k(+) \\ 0 & \varepsilon^T \end{array}\right] = \left[\begin{array}{cc} C_{Y_k(-)} & H_k^T C_{R_k^{-1}} \\ \hat{s}_k^T(-) & z_k^T C_{R_k^{-1}} \end{array}\right] \left[\begin{array}{cc} C_{Y_k(-)}^T & \hat{s}_k(-) \\ C_{R_k^{-1}}^T H_k & C_{R_k^{-1}}^T z_k \end{array}\right], \quad (113)$$

since $T_{\text{obs}} T_{\text{obs}}^T = I$. The right-hand-side of Equation 113 is given by

$$\left[\begin{array}{cc} C_{Y_k(-)} C_{Y_k(-)}^T + H_k^T C_{R_k^{-1}} C_{R_k^{-1}}^T H_k & C_{Y_k(-)} \hat{s}_k(-) + H_k^T C_{R_k^{-1}} C_{R_k^{-1}}^T z_k \\ \hat{s}_k(-)^T C_{Y_k(-)}^T + z_k^T C_{R_k^{-1}} C_{R_k^{-1}}^T H_k & \hat{s}_k(-)^T \hat{s}_k(-) + z_k^T C_{R_k^{-1}} C_{R_k^{-1}}^T z_k \end{array}\right]$$

which becomes

$$\left[\begin{array}{cc} Y_k(-) + H_k^T R_k^{-1} H_k & C_{Y_k(-)} \hat{s}_k(-) + H_k^T R_k^{-1} z_k \\ \hat{s}_k^T(-) C_{Y_k(-)}^T + z_k^T R_k^{-1} H_k & \hat{s}_k(-)^T \hat{s}_k(-) + z_k^T R_k^{-1} z_k \end{array}\right]. \quad (114)$$

while the left-hand-side of Equation 113 is given by

$$\left[\begin{array}{cc} Y_k(+) & C_{Y_k(+)} \hat{s}_k(+) \\ \hat{s}_k^T(+) C_{Y_k(+)}^T & \hat{s}_k(+)^T \hat{s}_k(+) + \varepsilon \varepsilon^T \end{array}\right] \quad (115)$$

Equating the $(1,1)$ component in Equations 114 and 115 gives the covariance portion of the observational update

$$Y_k(+) = Y_k(-) + H_k^T R_k^{-1} H_k \,.$$

Equating the $(1,2)$ component in Equations 114 and 115 gives

$$C_{Y_k(+)} \hat{s}_k(+) = C_{Y_k(-)} \hat{s}_k(-) + H_k^T R_k^{-1} z_k \,,$$

or when we recall the definition of the square-root information state $\hat{s}_k(\pm)$ given by

$$\hat{s}_k(\pm) = C_{Y_k(\pm)}^T \hat{x}_k(\pm) \,, \quad (116)$$

we have

$$C_{Y_k(+)} C_{Y_k(+)}^T \hat{x}_k(+) = C_{Y_k(-)} C_{Y_k(-)}^T \hat{x}_k(-) + H_k^T R_k^{-1} z_k \,,$$

or

$$Y_k(+) \hat{x}_k(+) = Y_k(-) \hat{x}_k(-) + H_k^T R_k^{-1} z_k \,,$$

the measurement update equation showing the desired equivalence.

Problem 6.24 (Swerling's informational form)

Consider the suggested product we find

$$
\begin{aligned}
P(+)P(+)^{-1} &= \left(P(-) - P(-)H^T[HP(-)H^T + R]^{-1}HP(-)\right)\left(P(-)^{-1} + H^T R^{-1} H\right) \\
&= I + P(-)H^T R^{-1} H - P(-)H^T[HP(-)H^T + R]^{-1}H \\
&- P(-)H^T[HP(-)H^T + R]^{-1}HP(-)H^T R^{-1} H \\
&= I \\
&+ P(-)H^T\left(R^{-1}H - [HP(-)H^T + R]^{-1}H - [HP(-)H^T + R]^{-1}HP(-)H^T R^{-1} H\right) \\
&= I + P(-)H^T[HP(-)H^T + R]^{-1}\left[[HP(-)H^T + R]R^{-1}H - H - HP(-)H^T R^{-1} H\right] \\
&= I \,,
\end{aligned}
$$

as we were to show.

Problem 6.25 (Cholesky factors of $Y = P^{-1}$)

If $P = CC^T$ then defining Y^{-1} as $Y^{-1} = P = CC^T$ we have that

$$
Y = (CC^T)^{-1} = C^{-T}C^{-1} = (C^{-T})(C^{-T})^T \,,
$$

showing that the Cholesky factor of $Y = P^{-1}$ is given by C^{-T}.

Chapter 7: Practical Considerations

Notes On The Text

Example 7.10-11: Adding Process Noise to the Model

Consider the true *real* world model

$$\begin{aligned} \dot{x}_1(t) &= 0 \\ \dot{x}_2(t) &= x_1 \\ z(t) &= x_2(t) + v(t) \,, \end{aligned} \tag{117}$$

In this model x_1 is a constant say $x_1(0)$ and then the second equation is $\dot{x}_2 = x_1(0)$ so $x_2(t)$ is given by

$$x_2(t) = x_2(0) + x_1(0)t \,, \tag{118}$$

a linear "ramp". Assume next that we have modeled this system *incorrectly*. We first consider processing the measurements $z(t)$ with the incorrect model

$$\begin{aligned} \dot{x}_2(t) &= 0 \\ z(t) &= x_2(t) + v(t) \,. \end{aligned} \tag{119}$$

Using this model the estimated state $\hat{x}_2(t)$ will converge to a constant, say $\hat{x}_2(0)$, and thus the filter error in state $\tilde{x}_2(t) = \hat{x}_2(t) - x_2(t)$ will be given by

$$\tilde{x}_2(t) = \hat{x}_2(0) - x_2(0) + x_1(0)t \,,$$

which will grow without bounds as $t \to +\infty$. This set of manipulations can be summarized by stating that: **with the incorrect world model the state estimate can diverge**.

Note that there is no process noise in this system formulation. One "ad hoc" fix one could try would be to *add* some process noise so that we consider the alternative model

$$\begin{aligned} \dot{x}_2(t) &= w(t) \\ z(t) &= x_2(t) + v(t) \,. \end{aligned} \tag{120}$$

Note that in this model the equation for x_2 is in the same form as Equation 119 but with the addition of a $w(t)$ or a process noise term. This is a scalar system which we can solve explicitly. The time dependent covariance matrix $P(t)$ for this problem can be obtained by solving Equation 121 or

$$\dot{P}(t) = P(t)F(t)^T + F(t)P(t) - P(t)H(t)^T R^{-1}(t)H(t)P(t) + G(t)Q(t)G(t)^T \tag{121}$$

with $F = 0$, $H = 1$, $G = 1$, and $R(t)$ and $Q(t)$ constants to get

$$\dot{P}(t) = -\frac{P(t)^2}{R} + Q \,.$$

If we look for the steady-state solution we have $P(\infty) = \sqrt{RQ}$. The steady-state Kalman gain in this case is given by

$$\overline{K}(\infty) = P(\infty) H^T R^{-1} = \frac{\sqrt{RQ}}{R} = \sqrt{\frac{Q}{R}}.$$

which is a constant and never decays to zero. This is a good property in that it means that the filter will never become so over confident that it will not update its belief with new measurements. For the modified state equations (where we have added process noise) we can explicitly compute the error between our state estimate $\hat{x}_2(t)$ and the "truth" $x_2(t)$. To do this recall that we will be filtering and computing $\hat{x}_2(t)$ using

$$\dot{\hat{x}}_2(t) = F\hat{x}_2 + \overline{K}(t)(z(t) - H\hat{x}_2(t)).$$

When we consider the long time limit we can take $\overline{K}(t) \to \overline{K}(\infty)$ and with $F = 0$, $H = 1$ we find our estimate of the state is the solution to

$$\dot{\hat{x}}_2 + \overline{K}(\infty)\hat{x}_2 = \overline{K}(\infty) z(t).$$

We can solve this equation using Laplace transforms where we get (since $\mathcal{L}(\dot{\hat{x}}_2) = s\hat{x}_2$)

$$[s + \overline{K}(\infty)]\hat{x}_2(s) = \overline{K}(\infty) z(s),$$

so that our steady-state filtered solution $\hat{x}_2(s)$ looks like

$$\hat{x}_2(s) = \frac{\overline{K}(\infty)}{s + \overline{K}(\infty)} z(s).$$

We are now in a position to see how well our estimate of the state $\hat{x}_2$ compares with the *actual* true value given by Equation 118. We will do this by considering the error in the state i.e. $\tilde{x}(t) = \hat{x}_2(t) - x_2(t)$, specifically the Laplace transform of this error or $\tilde{x}(s) = \hat{x}_2(s) - x_2(s)$. Now under the *best* case possible, where there is no measurement noise $v = 0$, our measurement $z(t)$ in these models (Equations 117, 119, and 120) is exactly $x_2(t)$ which we wish to estimate. In this case since we know the functional form of the true solution $x_2(t)$ via. Equation 118, we know then the Laplace transform of $z(t)$

$$z(s) = x_2(s) = \mathcal{L}\{x_2(0) + x_1(0)t\} = \frac{x_2(0)}{s} + \frac{x_1(0)}{s^2}. \tag{122}$$

With this we get

$$\begin{aligned}\tilde{x}_2(s) = \hat{x}_2(s) - x_2(s) &= \left[\frac{\overline{K}(s)}{s + \overline{K}(s)} - 1\right] x_2(s) \\ &= -\frac{s}{s + \overline{K}(\infty)} x_2(s).\end{aligned}$$

Using the final value theorem we have that

$$\begin{aligned}\tilde{x}_2(\infty) = \hat{x}_2(\infty) - x_2(\infty) &= \lim_{s\to 0} s\left[\hat{x}_2(s) - x_2(s)\right] \\ &= \lim_{s\to 0}\left[s\left(-\frac{s}{s + \overline{K}(s)}\right) x_2(s)\right].\end{aligned}$$

But as we argued before $x_2(s) = \frac{x_2(0)}{s} + \frac{x_1(0)}{s^2}$, thus we get

$$\tilde{x}_2(\infty) = \lim_{s\to\infty} \left[s \left(-\frac{s}{s+\overline{K}(\infty)} \right) \left(\frac{x_2(0)}{s} + \frac{x_1(0)}{s^2} \right) \right] = -\frac{x_1(0)}{\overline{K}(\infty)} .$$

Note that this is a constant and does not decay with time and there is an inherent bias in the Kalman solution. This set of manipulations can be summarized by stating that: **with the incorrect world model adding process noise can prevent the state from diverging**.

We now consider the case where we get the number and state equations correct but we add some additional process noise to the constant state x_1. That is in this case we still assume that the real world model is given by Equations 117 but that our Kalman model is given by

$$\begin{aligned} \dot{x}_1(t) &= w(t) \\ \dot{x}_2(t) &= x_1(t) \\ z(t) &= x_2(t) + v(t) , \end{aligned} \tag{123}$$

Then for this model we have

$$F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} , \quad G = \begin{bmatrix} 1 \\ 0 \end{bmatrix} , \quad H = \begin{bmatrix} 0 & 1 \end{bmatrix} , \quad Q = \mathrm{cov}(w) \quad R = \mathrm{cov}(v) .$$

To determine the steady-state performance of this model we need to solve for the steady state value $P(\infty)$ in

$$\dot{P}(t) = FP + PF^T + GQG^T - PH^TR^{-1}HP \quad \text{and} \quad \overline{K} = PH^TR^{-1} .$$

with F, G, Q, and H given by the above, we see that

$$\begin{aligned} FP &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ p_{11} & p_{12} \end{bmatrix} \\ PF^T &= \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & p_{11} \\ 0 & p_{12} \end{bmatrix} \\ GQG^T &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} Q \begin{bmatrix} 1 & 0 \end{bmatrix} = Q \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ PH^TR^{-1}HP &= \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \frac{1}{R} \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \\ &= \frac{1}{R} \begin{bmatrix} p_{12} \\ p_{22} \end{bmatrix} \begin{bmatrix} p_{12} & p_{22} \end{bmatrix} = \frac{1}{R} \begin{bmatrix} p_{12}^2 & p_{12}p_{22} \\ p_{22}p_{12} & p_{22}^2 \end{bmatrix} . \end{aligned}$$

Thus the Ricatti equation becomes

$$\dot{P} = \begin{bmatrix} 0 & 0 \\ p_{11} & p_{12} \end{bmatrix} + \begin{bmatrix} 0 & p_{11} \\ 0 & p_{12} \end{bmatrix} + \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} - \frac{1}{R} \begin{bmatrix} p_{12}^2 & p_{12}p_{22} \\ p_{22}p_{12} & p_{22}^2 \end{bmatrix} = \begin{bmatrix} Q - \frac{p_{12}^2}{R} & p_{11} - \frac{p_{12}p_{22}}{R} \\ p_{11} - \frac{p_{12}p_{22}}{R} & 2p_{12} - \frac{p_{22}^2}{R} \end{bmatrix} .$$

To find the steady-state we take $\frac{dP}{dt} = 0$ we get by using the $(1,1)$ component equation that p_{12} is given by $p_{12} = \pm\sqrt{QR}$. When we put this in the $(2,2)$ component equation we have

$$0 = \pm 2\sqrt{QR} - \frac{p_{22}^2}{R} .$$

Which means that $p_{22}^2 = \pm 2R\sqrt{QR}$. We must take the positive sign as p_{22} must be a positive real number. To take the positive number we have $p_{12} = \sqrt{QR}$. Thus $p_{22}^2 = 2Q^{1/2}R^{3/2}$ or

$$p_{22} = \sqrt{2}(R^3Q)^{1/4}\,.$$

When we put this value into the $(1,2)$ component equation we get

$$p_{11} = \frac{p_{12}p_{22}}{R} = \sqrt{2}\frac{(QR)^{1/2}}{R}(R^3Q)^{1/4} = \sqrt{2}(Q^3R)^{1/4}\,.$$

Thus the steady-state Kalman gain $\overline{K}(\infty)$ then becomes

$$\begin{aligned}\overline{K}(\infty) = P(\infty)H^TR^{-1} &= \frac{1}{R}\begin{bmatrix} p_{11}(\infty) & p_{12}(\infty) \\ p_{12}(\infty) & p_{22}(\infty)\end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \frac{1}{R}\begin{bmatrix} p_{12}(\infty) \\ p_{22}(\infty)\end{bmatrix} = \begin{bmatrix} \left(\frac{Q}{R}\right)^{1/2} \\ \sqrt{2}\left(\frac{Q}{R}\right)^{1/4}\end{bmatrix}\,. \end{aligned} \tag{124}$$

To determine how the steady-state Kalman estimate $\hat{x}(t)$ will compare to the truth $\mathbf{x}$ given via $x_1(t) = x_1(0)$ and Equation 118 for $x_2(t)$. We start with the dynamical system we solve to get the estimate $\hat{x}$ given by

$$\dot{\hat{x}} = F\hat{x} + \overline{K}(z - H\hat{x})\,.$$

Taking the long time limit where $t \to \infty$ of this we have

$$\dot{\hat{x}}(t) = F\hat{x}(t) + \overline{K}(\infty)z(t) - \overline{K}(\infty)H\hat{x}(t) = (F - \overline{K}(\infty)H)\hat{x}(t) + \overline{K}(\infty)z(t)\,.$$

Taking the Laplace transform of the above we get

$$s\hat{x}(s) - \hat{x}(0) = (F - K(\infty)H)\hat{x}(s) + \overline{K}(\infty)z(s)\,,$$

or

$$[sI - F - \overline{K}(\infty)H]\hat{x}(s) = \hat{x}(0) + \overline{K}(\infty)z(s)\,.$$

Dropping the term $\hat{x}(0)$ as $t \to \infty$ and it influence will be negligible we get

$$\hat{x}(s) = [sI - F - \overline{K}(\infty)]^{-1}\overline{K}(\infty)z(s)\,. \tag{125}$$

From the definitions of the matrices above we have that

$$sI - F + \overline{K}(\infty)H = \begin{bmatrix} s & \overline{K}_1(\infty) \\ -1 & s + \overline{K}_2(\infty)\end{bmatrix}\,,$$

and the inverse is given by

$$[sI - F + \overline{K}(\infty)H]^{-1} = \frac{1}{s(s + \overline{K}_2(\infty)) + \overline{K}_1(\infty)}\begin{bmatrix} s + \overline{K}_2(\infty) & -\overline{K}_1(\infty) \\ 1 & s \end{bmatrix}\,.$$

Since we know that $z(s)$ is given by Equation 122 we can use this expression to evaluate the vector $\hat{x}(s)$ via Equation 125. We could compute both $\hat{x}_1(s)$ and $\hat{x}_2(s)$ but since we only want to compare performance of $\hat{x}_2(s)$ we only calculate that component. We find

$$\hat{x}_2(s) = \left(\frac{\overline{K}_1(\infty) + s\overline{K}_2(\infty)}{s(s + \overline{K}_2(\infty)) + \overline{K}_1(\infty)}\right)z(s)\,. \tag{126}$$

Then since $z(t) = x_2(t)$ we have

$$\begin{aligned}\tilde{x}_2(s) = \hat{x}_2(s) - x_2(s) &= \left(\frac{\overline{K}_1(\infty) + s\overline{K}_2(\infty)}{s(s + \overline{K}_2(\infty)) + \overline{K}_1(\infty)} \right) z(s) - z(s) \\ &= - \left(\frac{s^2}{s^2 + \overline{K}_2(\infty)s + \overline{K}_1(\infty)} \right) z(s) \\ &= - \left(\frac{s^2}{s^2 + \overline{K}_2(\infty)s + \overline{K}_1(\infty)} \right) \left[\frac{x_2(0)}{s} + \frac{x_1(0)}{s^2} \right] ,\end{aligned}$$

when we use Equation 122. Then using the final-value theorem we have the limiting value of $\tilde{x}_2(\infty)$ given by

$$\tilde{x}_2(\infty) = \lim_{s \to 0} s\tilde{x}_2(s) = \lim_{s \to 0} \left(\frac{-s^3}{s^2 + \overline{K}_2(\infty)s + \overline{K}_1(\infty)} \right) \left[\frac{x_2(0)}{s} + \frac{x_1(0)}{s^2} \right] = 0 ,$$

showing that this addition of process noise results in a convergent estimate. This set of manipulations can be summarized by stating that: **with the incorrect world model adding process noise can result in good state estimates**.

As the final example presented in the book we consider the case where the real world model *has* process noise in the dynamics of x_1 but the model use to perform filtering does not. That is, in this case we assume that that the real world model is given

$$\begin{aligned}\dot{x}_1(t) &= w(t) \\ \dot{x}_2(t) &= x_1(t) \\ z(t) &= x_2(t) + v(t) ,\end{aligned}$$

and that our Kalman model is given by

$$\begin{aligned}\dot{x}_1(t) &= 0 \\ \dot{x}_2(t) &= x_1(t) \\ z(t) &= x_2(t) + v(t) ,\end{aligned} \tag{127}$$

Then for this *assumed* model we have

$$F = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} , \quad H = \begin{bmatrix} 0 & 1 \end{bmatrix} , \quad R = \text{cov}(v) , \quad \text{and} \quad G = 0 , \quad \text{or} \quad Q = 0 .$$

To determine the steady-state performance of this model we need to solve for the steady state value $P(\infty)$ in

$$\dot{P}(t) = FP + PF^T - PH^TR^{-1}HP \quad \text{and} \quad \overline{K} = PH^TR^{-1} .$$

with F, G, Q, and H given by the above, we have the same expressions as above but without the GQG^T term. Thus the Ricatti equation becomes

$$\dot{P} = \begin{bmatrix} -\frac{p_{12}^2}{R} & p_{11} - \frac{p_{12}p_{22}}{R} \\ p_{11} - \frac{p_{12}p_{22}}{R} & 2p_{12} - \frac{p_{22}^2}{R} \end{bmatrix} .$$

To find the steady-state we take $\frac{dP}{dt} = 0$ we get by using the $(1,1)$ component equation that p_{12} is given by $p_{12} = 0$. When we put this in the $(2,2)$ component equation we have that $p_{22} = 0$. When we put this value into the $(1,2)$ component equation we get $p_{11} = 0$. Thus the steady-state Kalman gain $\overline{K}(\infty)$ is zero. To determine how the steady-state Kalman estimate $\hat{x}(t)$ will compare to the truth $\mathbf{x}$ given via $x_1(t) = x_1(0)$ and Equation 118 for $x_2(t)$. We start with the dynamical system we solve to get the estimate $\hat{x}$ given by

$$\dot{\hat{x}} = F\hat{x}\,.$$

This has the simple solution given by

$$\begin{aligned} \hat{x}_1(t) &= \hat{x}_1(0) \quad \text{or a constant} \\ \hat{x}_2(t) &= \hat{x}_1(0)t + \hat{x}_2(0) \quad \text{or the "ramp" function}\,. \end{aligned}$$

Since the true solution for $x_1(t)$ is not a constant this approximate solution is poor.

References

[1] W. G. Kelley and A. C. Peterson. *Difference Equations. An Introduction with Applications.* Academic Press, New York, 1991.