Chapter 3

RECURSIVE ESTIMATION AND KALMAN FILTERING

3.1 The Discrete Time Kalman Filter

Consider the following estimation problem. Given the stochastic system

$$x_{k+1} = Ax_k + Gw_k \tag{3.1}$$

$$y_k = Cx_k + Hv_k \tag{3.2}$$

with

$$x(k_0) = x_0$$

find the linear least squares estimate of x_k based on past observations $y_{k_0}, ..., y_{k-1}$. We denote this by either $\hat{E}\{x_k|\mathcal{Y}_{k-1}\}$ where $\mathcal{Y}_{k-1} = \{y_{k_0}, ..., y_{k-1}\}$, or by $\hat{x}_{k|k-1}$, or by $\hat{x}(k|k-1)$. We also use y^k for $\mathcal{Y}_k$.

In general, the problem of estimating x_k based on $\mathcal{Y}_j$ is called the prediction problem, the filtering problem, and the smoothing or interpolation problem, for $j < k$, $j = k$, and $j > k$, respectively. Since there is a one-step delay in the information available for computing $\hat{x}_{k|k-1}$, we often call $\hat{x}_{k|k-1}$ the one-step ahead predictor.

We make the following assumptions concerning the system (3.1) and (3.2), which will be in force throughout the rest of this chapter.

(i) w_k is an uncorrelated sequence of zero mean random vectors, with $Ew_kw_k^T = Q$.

(ii) v_k is an uncorrelated sequence of zero mean random vectors, with $Ev_kv_k^T = R$.

(iii) The initial random vector x_0 has mean m_0 and covariance P_0.

(iv) w_k, v_j and x_0 are mutually uncorrelated for all k and j, except $Ew_kv_k^T = T$

(v) The matrix HRH^T is assumed to be nonsingular, hence positive definite.

Although we have assumed constant matrices for simplicity, all derivations and results, with the exception of those on asymptotic behaviour, remain unchanged for systems with time-varying matrices as well.

Let us analyze the estimation problem. Obviously, $\hat{x}_{k_0|k_0-1} = m_0$ because $\mathcal{Y}_{k_0-1}$ has no observations and hence no new information. If we define the error covariance $P_{k|k-1} = E\{(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T\}$, we also have $P_{k_0|k_0-1} = P_0$. Now assume that $y_{k_0}, ..., y_{k-1}$ have been observed, giving rise to the estimate

$\hat{x}_{k|k-1}$ and the error covariance $P_{k|k-1}$. The new measurement y_k improves our estimate and according to the results of Section 1.7,

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + P_{k|k-1}C^T(CP_{k|k-1}C^T + HRH^T)^{-1}(y_k - C_k\hat{x}_{k|k-1}) \tag{3.3}$$

Similarly,

$$P_{k|k} = P_{k|k-1} - P_{k|k-1}C^T(CP_{k|k-1}C^T + R)^{-1}CP_{k|k-1} \tag{3.4}$$

These results follow because by our assumptions, v_k is orthogonal to x_k and to $\mathcal{Y}_{k-1}$.

To continue the process by induction, we need to find $\hat{x}_{k+1|k}$ in terms of $\hat{x}_{k|k}$. But $\hat{x}_{k+1|k}$ is given by

$$\begin{aligned} \hat{E}\{x_{k+1}/\mathcal{Y}_k\} &= \hat{E}\{Ax_k + Gw_k/\mathcal{Y}_k\} \\ &= A_k\hat{x}_{k|k} + G\hat{w}_{k|k} \end{aligned}$$

Since w_k is orthogonal to $\mathcal{Y}_{k-1}$,

$$\hat{w}_{k|k} = E(w_k\tilde{y}_k^T)E(\tilde{y}_k\tilde{y}_k^T)^{-1}\tilde{y}_k$$

where $\tilde{y}_k = y_k - C\hat{x}_{k|k-1}$. Now

$$E(w_k\tilde{y}_k^T) = E(w_k v_k^T H^T)$$

so that

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k} + GTH^T(CP_{k|k-1}C^T + HRH^T)^{-1}\tilde{y}_k \tag{3.5}$$

Substituting (3.3) into (3.5), we obtain

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + (AP_{k|k-1}C^T + GTH^T)(CP_{k|k-1}C^T + HRH^T)^{-1}(y_k - C\hat{x}_{k|k-1})$$

To obtain a difference equation for the error covariance $P_{k|k-1}$, let K_k denote the Kalman gain

$$K_k = (AP_{k|k-1}C^T + GTH^T)(CP_{k|k-1}C^T + HRH^T)^{-1}$$

and $\tilde{x}_{k|k-1}$ denote the estimation error

$$\tilde{x}_{k|k-1} = x_k - \hat{x}_{k|k-1}$$

The estimation error satisfies the equation

$$\tilde{x}_{k+1|k} = (A - K_kC)\tilde{x}_{k|k-1} + Gw_k - K_kHv_k \tag{3.6}$$

Since $P_{k|k-1} = E(\tilde{x}_{k|k-1}\tilde{x}_{k|k-1}^T)$, we obtain the following difference equation for $P_{k|k-1}$:

$$\begin{aligned} P_{k+1|k} &= (A - K_kC)P_{k|k-1}(A - K_kC)^T + GQG^T \\ &\quad - GTH^TK_k^T - K_kHT^TG^T + K_kHRH^TK_k^T \\ &= AP_{k|k-1}A^T + GQG^T \\ &\quad - (AP_{k|k-1}C^T + GTH^T)(CP_{k|k-1}C^T + HRH^T)^{-1}(CP_{k|k-1}A^T + HT^TG^T) \end{aligned}$$

Combining the above results, we obtain the discrete time Kalman filter in the one-step ahead prediction form.

Theorem 3.1: The linear least squares estimator of x_k given $\mathcal{Y}_{k-1}$ is generated by the following recursive relations:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + (AP_{k|k-1}C^T + GTH^T)(CP_{k|k-1}C^T + HRH^T)^{-1}(y_k - C\hat{x}_{k|k-1}) \tag{3.7}$$

$$\begin{aligned} P_{k+1|k} &= AP_{k|k-1}A^T + GQG^T \\ &\quad - (AP_{k|k-1}C^T + GTH^T)(CP_{k|k-1}C^T + HRH^T)^{-1}(CP_{k|k-1}A^T + HT^TG^T) \end{aligned} \tag{3.8}$$

with

$$\begin{aligned} \hat{x}_{k_0|k_0-1} &= m_0 \\ P_{k_0|k_0-1} &= P_0 \end{aligned}$$

The above theorem was first given by Kalman in his famous paper, "A new approach to linear filtering and prediction problem", J. Basic Engineering, ASME, 82 (March 1960), 34-45.

Equation (3.8) is especially important in control and estimation theory and is referred to as the *discrete time Riccati difference equation.* The "filtered" estimate and error covariance can be obtained using (3.3) and (3.4). For simplicity, we shall refer to all these equations as Kalman filter equations.

If we make the stronger assumptions that w_k, v_j and x_0 are Gaussian random vectors, then $\hat{E}(x_k/\mathcal{Y}_{k-1})$ is in fact the conditional expectation $E(x_k/\mathcal{Y}_{k-1})$. Thus in this case, the process $\hat{x}_{k|k-1}$ is the minimum mean square error estimator for x_k. In addition, since x_k and y_k are jointly Gaussian, the conditional error covariance

$$E\{[x_k - \hat{x}_{k|k-1}][x_k - \hat{x}_{k|k-1}]/\mathcal{Y}_{k-1}\}$$

does not depend on the observations $\mathcal{Y}_{k-1}$ and hence is equal to $P_{k|k-1}$. So the Kalman filter completely characterizes the conditional probability distribution in this case.

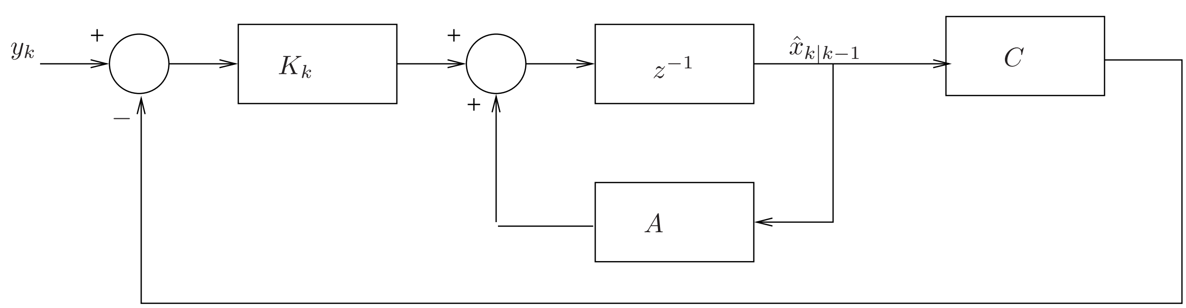
Equation (3.7) may be written as

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k-1} + K_k(y_k - C\hat{x}_{k|k-1})$$

where

$$K_k = (AP_{k|k-1}C^T + GTH^T)[CP_{k|k-1}C^T + HRH^T]^{-1}$$

is the Kalman gain. This shows that the Kalman filter uses the same dynamics as the system state equation, with the new information contained in $\tilde{y}_k$ fed back into the system through the Kalman gain. The block diagram description of the Kalman filter is given in the following figure.



A very important feature of the Kalman filter is that the error covariance does not depend on the observations. Hence $P_{k|k-1}$ can be *pre-computed* and the accuracy of the filter assessed before the observations are made. In particular, we may investigate the asymptotic behaviour of the filter by analyzing the discrete time Riccati equation. This we shall do in a later section.

Example 3.1.1:

Consider the following scalar value process

$$\begin{aligned} x_{k+1} &= x_k \\ y_k &= x_k + v_k \end{aligned}$$

This corresponds to a constant random variable observed in noise. The Riccati difference equation is given by (we denote $p_{k|k-1}$ by p_k for simplicity)

$$p_{k+1} = p_k - \frac{p_k^2}{p_k + r} = \frac{p_k r}{p_k + r} \tag{3.9}$$

To solve this nonlinear difference equation, observe that if we write the equation for $\frac{1}{p_k}$, we get

$$\frac{1}{p_{k+1}} = \frac{1}{p_k} + \frac{1}{r}$$

This is now a linear equation, which can be solved immediately to give

$$\frac{1}{p_k} = \frac{1}{p_0} + \frac{k}{r}$$

Inverting, we obtain the solution to (3.9)

$$p_k = \frac{p_0 r}{r + p_0\, k} \tag{3.10}$$

The Kalman filter equation for this system can be readily written down

$$\hat{x}_{k+1|k} = \hat{x}_{k|k-1} + \frac{p_k}{p_k + r}(y_k - \hat{x}_{k|k-1}) \tag{3.11}$$

Substituting (3.10) into (3.11), we obtain

$$\begin{aligned}
\hat{x}_{k+1|k} &= \hat{x}_{k|k-1} + \frac{\frac{p_0 r}{r+p_0\, k}}{\frac{p_0 r}{r+p_0\, k} + r}(y_k - \hat{x}_{k|k-1}) \\
&= \hat{x}_{k|k-1} + \frac{p_0}{r + p_0(k+1)}(y_k - \hat{x}_{k|k-1}) \\
&= \frac{r + p_0\, k}{r + p_0(k+1)}\hat{x}_{k|k-1} + \frac{p_0}{r + p_0(k+1)} y_k
\end{aligned} \tag{3.12}$$

Equation (3.12) is a linear time-varying difference equation which can be readily solved. First observe that the transition function associated with $\frac{r+p_0\, k}{r+p_0(k+1)}$ is given by

$$\Phi(k;j) = \frac{r + p_0(k-1)}{r + p_0\, k}\frac{r + p_0(k-2)}{r + p_0(k-1)} \cdots \frac{r + p_0 j}{r + p_0(j+1)} = \frac{r + p_0 j}{r + p_0\, k} \tag{3.13}$$

Using the results of Chapter 2, we can immediately write down

$$\begin{aligned}
\hat{x}_{k|k-1} &= \frac{r}{r + p_0\, k} m_0 + \sum_{j=0}^{k-1} \frac{r + p_0(j+1)}{r + p_0\, k}\, \frac{p_0}{r + p_0(j+1)} y_j \\
&= \frac{r}{r + p_0\, k} m_0 + \frac{1}{k}\sum_{j=0}^{k-1} \frac{1}{1 + \frac{r}{p_0\, k}} y_j
\end{aligned} \tag{3.14}$$

Example 3.1.2:

Consider the scalar process

$$x_{k+1} = a x_k + w_k \tag{3.15}$$

$$y_k = x_k \tag{3.16}$$

It is immediately seen that the solution of the Riccati difference equation is given by

$$p_{k|k-1} = q \quad \text{for} \ \ k \geq 1$$

The Kalman one-step ahead predictor equation is given by

$$\begin{aligned} \hat{x}_{k+1|k} &= a\hat{x}_{k|k-1} + \frac{aq}{q}(y_k - \hat{x}) \\ &= ay_k \end{aligned} \tag{3.17}$$

To get the filtered estimate, we have

$$\begin{aligned} \hat{x}_{k|k} &= \hat{x}_{k|k-1} + (y_k - \hat{x}_{k|k-1}) \\ &= y_k = x_k \end{aligned}$$

Finally, the error covariance for the filtered estimate is given by

$$P_{k|k} = q - \frac{q^2}{q} = 0$$

These results make good physical sense. For, since x_k is perfectly observed at time k, we expect that the filtered estimate will be a perfect estimate of x_k. The filtered error covariance should therefore be 0. On the other hand, the one-step ahead predictor should give

$$\hat{x}_{k+1|k} = a\hat{x}_{k|k} = ay_k$$

which is exactly what the one-step predictor equation provides.

3.2 The Innovations Process

The process $\nu_k = y_k - C_k\hat{x}_{k|k-1}$ is called the innovations process. It represents the new information contained in y_k for the estimation of x_k. The following is an important property of the innovations.

Theorem 3.2: ν_k and ν_j are orthogonal for $k \neq j$ and

$$\text{cov}(\nu_k) = CP_{k|k-1}C^T + HRH^T$$

Proof: Suppose $j < k$. We may write the innovations process in 2 ways:

$$\begin{aligned} \nu_k &= C\tilde{x}_{k|k-1} + Hv_k \\ \nu_j &= y_j - C\hat{x}_{j|j-1} \end{aligned}$$

where

$$\tilde{x}_{j|j-1} = x_j - \hat{x}_{j|j-1}$$

is the estimation error.

Since $\hat{x}_{j|j-1}$ is a linear function of the observations $y(s)$, $k_0 \leq s \leq j-1$, ν_j is a linear function of $y(s)$, $k_0 \leq s \leq j$. By the Projection Theorem, $\tilde{x}_{k|k-1}$ is orthogonal to $y(\sigma)$, $k_0 \leq \sigma \leq k-1$. Since v_k is also orthogonal to $\mathcal{Y}_{k-1}$, we find that ν_k is orthogonal to ν_j

$$\begin{aligned} \text{cov}(\nu_k) &= E\{(C\tilde{x}_{k|k-1} + Hv_k)(C\tilde{x}_{k|k-1} + Hv_k)^T\} \\ &= CP_{k|k-1}C^T + HRH^T \end{aligned}$$

by the orthogonality between $\tilde{x}_{k|k-1}$ and v_k.

Recall that an uncorrelated sequence is called white noise. The above result shows that the innovations process is a white noise with covariance which depends on the estimation error covariance.

3.3 The Discrete Time Riccati Equation

Further analysis of the Kalman filter hinges on the analysis of the discrete time Riccati difference equation

$$\begin{aligned} P_{k+1|k} &= AP_{k|k-1}A^T + GQG^T \\ &\quad - (AP_{k|k-1}C^T + GTH^T)(CP_{k|k-1}C^T + HRH^T)^{-1}(CP_{k|k-1}A^T + HT^TG^T) \end{aligned} \tag{3.18}$$

and, in the time-invariant case, the algebraic Riccati equation

$$P = APA^T - (APC^T + GTH^T)(CPC^T + HRH^T)^{-1}(HT^TG^T + CPA^T) + GQG^T \tag{3.19}$$

The full details are rather complex and we shall not go into them. We shall only state the most useful results.

Definition: A pair of matrices (A, B) with A, $n \times n$ and B $n \times m$ is called **stabilizable** if there exists a $m \times n$ matrix L such that $A - BL$ is stable, i.e. that $|\lambda(A - BL)| < 1$.

Definition: A pair of matrices (A, C) with A $n \times n$ and C $p \times n$ is called **detectable** if there exists a $n \times p$ matrix K such that $A - KC$ is stable.

Since a matrix M is stable if and only if M^T is stable, the detectability definition is equivalent to requiring the existence of K^T such that $A^T - C^TK^T$ is stable. This is the definition of (A^T, C^T) stabilizable. We conclude that (C, A) is detectable if and only if (A^T, C^T) is stabilizable.

Algebraic tests for stabilizability and detectability are as follows:

(A, B) is stabilizable if and only if $Rank[A - \lambda I \quad B] = n$ for all λ which is an unstable eigenvalue of A. Similarly, (C, A) is detectable if and only if $Rank \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} = n$ for all λ which is an unstable eigenvalue of A. These tests are referred to as the **PBH tests** for stabilizability and detectability.

Let

$$\begin{aligned} \breve{A} &= A - GTH^T(HRH^T)^{-1}C \\ \breve{G} &= G(Q - TH^T(HRH^T)^{-1}HT^T)^{\frac{1}{2}} \\ K &= (APC^T + GTH^T)(CPC^T + HRH^T)^{-1} \end{aligned}$$

We have the following important result.

Theorem 3.3: Suppose $(\breve{A}, \breve{G})$ is stabilizable and (C, A) detectable. Then the algebraic Riccati equation (3.19) has a unique solution P in the class of positive semidefinite matrices. The matrix $A - (APC^T + GTH^T)(CPC^T + HRH^T)^{-1}C$ is stable. Furthermore, for any $P_0 \geq 0$, $P_{k|k-1} \longrightarrow P$ as $k \longrightarrow \infty$.

There are many proofs of this result. See, e.g. M.H.A. Davis and R.B. Vinter, *Stochastic Modelling and Control*.

Remark 1: The matrices $\breve{A}$ and $\breve{G}$ arise in the following manner. When $T \neq 0$, w_k and Hv_k are not orthogonal. However, we can create a process $\tilde{w}_k$ which is orthogonal to Hv_k by setting

$$\tilde{w}_k = w_k - TH^T(HRH^T)^{-1}Hv_k$$

It is easy to verify that $E\tilde{w}_kv_k^TH^T = 0$. Write

$$w_k = \tilde{w}_k + TH^T(HRH^T)^{-1}Hv_k = \tilde{w}_k + TH^T(HRH^T)^{-1}(y_k - Cx_k)$$

Putting this into the system equation (3.1), we obtain

$$x_{k+1} = \check{A}x_k + G\tilde{w}_k + TH^T(HRH^T)^{-1}y_k$$

Observe that

$$E\tilde{w}_k\tilde{w}_k^T = Q - TH^T(HRH^T)^{-1}HT^T$$

so that $EG\tilde{w}_k\tilde{w}_k^T G^T = \check{G}\check{G}^T$. If we let ξ_k be a zero mean white noise process with $cov(\xi_k) = I$, we can write $G\tilde{w}_k = \check{G}\xi_k$ without changing the second order properties of the equations. This yields

$$x_{k+1} = \check{A}x_k + \check{G}\xi_k + TH^T(HRH^T)^{-1}y_k$$

which explains how the condition of stabilizability of $(\check{A}, \check{G})$ arises.

Remark 2: Any $\check{G}$ satisfying $\check{G}\check{G}^T = G(Q - TH^T(HRH^T)^{-1}HT^T)G^T$ can also be used in the stabilizability test. In fact, $(\check{A}, \check{G})$ is stabilizable if and only if $(\check{A}, \check{G}\check{G}^T)$ Note also that if $T = 0$ so that w_k and v_j are uncorrelated for all k, j, then $\check{A} = A$ and $\check{G} = GQ^{\frac{1}{2}}$. Checking the stabilizability condition is much simplified in this case.

Remark 3: The assumptions made in Theorem 3.3 are sufficient conditions. When they are not satisfied, the solution of the algebraic Riccati equation can be quite complicated. We observe that (C, A) detectable is clearly necessary for stability of $A - KC$, hence it is also a necessary condition for the existence of a stabilizing solution, i.e., a solution P whose corresponding K results in $A - KC$ stable. Delving into the detailed structure of all solutions to the algebraic Riccati equation without requiring stabilizability and detectability is beyond the scope of this course.

Let us discuss the meaning of Theorem 3.3. The algebraic Riccati equation (3.19) is a quadratic matrix equation in P. In general, there can be many solutions to such an equation. Theorem 3.3 asserts that if stabilizability and detectability holds, there is a unique positive semidefinite one. If we compare (3.19) with (3.18), we see that (3.19) is the "steady state" version of (3.18). Theorem 3.3 also shows, by the convergence of $P_{k|k-1}$ to P, that P is indeed the steady state error covariance.

If we put in this steady state version in (3.7) in place of $P_{k|k-1}$, we would get

$$\begin{aligned} \hat{x}_s(k+1|k) &= A\hat{x}_s(k|k-1) + (APC^T + GTH^T)(CPC^T + HRH^T)^{-1}(y_k - C\hat{x}_s(k|k-1)) \quad (3.20) \\ &= [A - (APC^T + GTH^T)(CPC^T + HRH^T)^{-1}C]\hat{x}_s(k|k-1) \\ &\quad + (APC^T + GTH^T)(CPC^T + HRH^T)^{-1}y_k \quad (3.21) \end{aligned}$$

Theorem 3.3 says that (3.21) is a stable system driven by the observations y_k. The estimation error process satisfies

$$\begin{aligned} \tilde{x}_s(k+1|k) &= [A - (APC^T + GTH^T)(CPC^T + HRH^T)^{-1}C]\tilde{x}_s(k|k-1) \\ &\quad + Gw_k - (APC^T + GTH^T)(CPC^T + HRH^T)^{-1}Hv_k \quad (3.22) \end{aligned}$$

This is a stable linear system driven by white noise. If we compute the error covariance associated with $\hat{x}_s(k+1|k)$, we get

$$P_s(k+1|k) = (A - KC)P_s(k|k-1)(A - KC)^T + GQG^T - GTH^TK^T - KHT^TG^T + KHRH^TK^T \quad (3.23)$$

Since, by Theorem 3.3, $(A - KC)$ is stable, $P_s(k+1|k)$ converges to a unique steady state solution as $k \longrightarrow \infty$. It is easily verified that P, the unique solution defined by the algebraic Riccati equation, is a steady state solution of (3.23). By uniqueness, we see that $P_s(k+1|k) \longrightarrow P$ as $k \longrightarrow \infty$. This means that the "steady state" filter is stable in the sense that its error covariance is bounded. But since the "steady state" filter is only suboptimal, this implies that the optimal filter defined by (3.7) must be stable also.

The stability property of the Kalman filter is one of the most important theoretical results. Filters which are not stable would have little engineering significance. This shows also the importance of system-theoretic properties like stabilizability and detectability. In this connection, we describe here 2 related important properties.

Definition: A pair of matrices (A, B) with A, $n \times n$ and B $n \times m$ is called **controllable** if Rank $[B \ \ AB \cdots A^{n-1}B] = n$.

The matrix $\mathcal{C}_{AB} = [B \ \ AB \cdots A^{n-1}B]$ is called the controllability matrix.

Definition: A pair of matrices (C, A) is **observable** if

$$\mathcal{N}\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} = \{0\}$$

where $\mathcal{N}(Q)$ denotes the nullspace of the matrix Q. From linear algebra, we know that this is equivalent to $[C^T \ A^TC^T \ (A^T)^2C^T \ \cdots (A^T)^{n-1}C^T]$ has rank n. The matrix

$$\mathcal{O}_{CA} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is called the observability matrix. The above observation says that $\mathcal{N}(\mathcal{O}_{CA}) = \{0\}$ if and only if $\text{Rank}(\mathcal{C}_{A^TC^T}) = n$. From this, we deduce the

Theorem on Duality between Controllability and Observability:
(A, B) is controllable if and only if (B^T, A^T) is observable.

A very important theorem in linear systems theory which relies on the notion of controllability is the Pole Assignment Theorem. To state it, we first define a symmetric set of complex numbers.

Definition:
A set of complex numbers is called symmetric if for any complex number belonging to the set, its complex conjugate also belongs to the set.

Note that eigenvalues of real matrices form a symmetric set of complex numbers. The roots of a polynomial with real coefficients also form a symmetric set.

We can now state

The Pole Assignment Theorem:
There exists a matrix L such that the eigenvalues of the matrix $A - BL$ coincide with any given set of symmetric numbers if and only if (A, B) is controllable.

By the Pole Assignment Theorem, we see that controllability allows the choice of a matrix L such that $A - BL$ is stable, since we can pick the symmetric set of complex numbers to be stable. Hence, it is clear that stabilizability is implied by controllability. Dually, we see that detectability is implied by observability.

Since there are algebraic tests for controllability and observability, these give verifiable sufficient conditions for the stability of the Kalman filter.

Example 3.3.1:
Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0.5 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The pair (A, B) is not controllable but is stabilizable. We can readily check that the only unstable eigenvalue is 1, and that

$$Rank(A - I\ B) = Rank \begin{bmatrix} 0 & 1 & 1 \\ 0 & -0.5 & 0 \end{bmatrix} = 2$$

Example 3.3.2:
Let

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0.5 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

The pair (A, B) is controllable and hence stabilizable.

In practice, the steady state filter (3.22) is often used instead of the optimal filter (3.7). This is because for many systems, $P_{k|k-1}$ converges to P reasonably fast, so that (3.22) is almost optimal. The steady state filter does not require the computation of $P_{k|k-1}$ at every stage, and is therefore a lot simpler to implement.

If we write the solution to (3.8) as $P(k; k_0)$ to indicate explicitly the dependence on the initial time, we see that by the time-invariance of the system, $P(k; k_0)$ actually depends only on $(k - k_0)$. From this we also obtain that $P(k; k_0) \underset{k_0 \to -\infty}{\longrightarrow} P$. We may interpret this as saying that if we had started the filter in the infinitely remote past, then we would have reached the steady state and that the steady state filter (3.21) would be optimal.

Example 3.3.3:
Consider the scalar system

$$x_{k+1} = ax_k + w_k$$
$$y_k = x_k + v_k$$

where w_k and v_j are assumed to be uncorrelated for all k, j. The algebraic Riccati equation is given by

$$p = a^2 p - \frac{a^2 p^2}{p + r} + q = \frac{a^2 r}{p + r} p + q$$

The solution for p is determined by solving the quadratic equation

$$p^2 + (r - a^2 r - q)p - rq = 0$$

which has the following unique positive solution

$$p = \frac{(a^2 - 1)r + q + \sqrt{[(a^2 - 1)r + q]^2 + 4rq}}{2}$$

The steady state Kalman filter has as its system matrix

$$a - \frac{ap}{p + r} = \frac{ar}{p + r} = \frac{2ar}{(a^2 + 1)r + q + \sqrt{[(a^2 - 1)r + q]^2 + 4rq}} < 1$$

Example 3.3.4:

Consider the 2nd order system

$$x_{k+1} = Ax_k + Gw_k$$
$$y_k = Cx_k + v_k$$

where

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad C = \begin{bmatrix} 0.5 & 1 \end{bmatrix}$$

Assume also that w_k and v_j are uncorrelated for all k, j, with $Q = 1$, $R = 1$. It is easily verified that (A, G) is controllable, hence stabilizable, and that (C, A) is observable, hence detectable. There exists therefore a unique positive semidefinite solution to the algebraic Riccati equation. The solution is computed using the Matlab command dlqe to give

$$P = \begin{bmatrix} 1.2557 & 0.4825 \\ 0.4825 & 1.6023 \end{bmatrix}$$

The eigenvalues of $(A - KC)$ are given by $0.5115 \pm 0.1805i$ which are inside the unit circle.

3.4 Exercises

1. In discrete time estimation problems, certain matrix identities are very useful. We examine some of them in this problem.

   (i) Verify the identity $(I_n + AB)^{-1} = I_n - A(I_m + BA)^{-1}B$ where A is $n \times m$, B is $m \times n$ and I_p is the $p \times p$ identity matrix.

   (ii) Using (i), prove that

   (a) $(A+BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$ whenever A^{-1} and C^{-1} exists. This, together with (i), are often referred to as the matrix inversion lemma.

   (b) If $P_2^{-1} = P_1^{-1} + C^TR^{-1}C$, then $P_2 = P_1 - P_1C^T(CP_1C^T + R)^{-1}CP_1$. ($C$ here is in general rectangular so that C^{-1} may not exist.)

   (c) If $P_2 = P_1 - P_1C^T(CP_1C^T + R)^{-1}CP_1$ then $P_2C^TR^{-1} = P_1C^T(CP_1C^T + R)^{-1}$.

2. In this problem, we analyze further the asymptotic behaviour of the scalar discrete Riccati equation

$$\begin{aligned} p(k+1) &= a^2p(k) - \tfrac{a^2p(k)^2}{p(k)+r} + q \\ p(0) &= p_0 \end{aligned} \qquad (ex2.1)$$

   and its relation to the algebraic Riccati equation

$$p = a^2p - \frac{a^2p^2}{p+r} + q \qquad (ex2.2)$$

   We have considered the cases (i) $r > 0$, $q > 0$, (ii) $a = 1$, $r > 0$, $q = 0$, and (iii) $r = 0$ in the notes. In this problem, you are asked to examine the other cases.

   (i) Assume $r > 0$, $q = 0$, $|a| \neq 1$. Solve (ex2.1) explicitly. Show that for all nonzero p_0, $p(t)$ converges for $|a| < 1$ and $|a| > 1$, and determine the limiting values. (Note that for $|a| > 1$, the system is not stabilizable.) Do these limiting values correspond to positive semidefinite solutions of (ex2.2)?

   (ii) The system associated with (ex2.1) and (ex2.2) when $q = 0$ is given by

$$\begin{aligned} x_{k+1} &= ax_k \\ y_k &= x_k + v_k \end{aligned}$$

   The time-varying Kalman filter is given by

$$\hat{x}_{k+1|k} = \left[a - \frac{ap(k)}{p(k)+r}\right]\hat{x}_{k|k-1} + \frac{ap(k)}{p(k)+r}y_k$$

   In each of the cases examined in (i) above, what value does $a - \frac{ap(k)}{p(k)+r}$ converge to? Is the resulting time-invariant filter stable? What conclusions can you draw from the case $|a| > 1$?

3. We have derived the covariance propagation equation for $P_{k|k-1} = cov(\tilde{x}_{k|k-1})$. One can also obtain an equation for the propagation of $P_{k|k} = cov(\tilde{x}_{k|k})$. It is most common to do this in two steps:

   (i) Express $P_{k|k}$ in terms of $P_{k|k-1}$. This has already been done in class.

   (ii) Then express $P_{k+1|k}$ in terms of $P_{k|k}$.

Carry out the derivation for (ii) for the case $Ew_k v_k^T = T$ using the following steps.

(a) Show that

$$(HRH^T)^{-1}(y_k - C\hat{x}_{k|k}) = (CP_{k|k-1}C^T + HRH^T)^{-1}\tilde{y}_{k|k-1}$$

(b) Show that

$$\tilde{x}_{k+1|k} = \check{A}\tilde{x}_{k|k} + Gw_k - GTH^T(HRH^T)^{-1}Hv_k$$

where $\check{A} = A - GTH^T(HRH^T)^{-1}C$.

(c) Finally, show that

$$P_{k+1|k} = \check{A}P_{k|k}\check{A}^T + \check{G}\check{G}^T$$

where $\check{G} = G(Q - TH^T(HRH^T)^{-1}HT^T)^{\frac{1}{2}}$.

Note that this means the Riccati difference equation can also be written as

$$P_{k+1|k} = \check{A}P_{k|k-1}\check{A}^T - \check{A}P_{k|k-1}C^T(CP_{k|k-1}C^T + HRH^T)^{-1}CP_{k|k-1}\check{A}^T + \check{G}\check{G}^T \qquad (ex3.1)$$

4. Consider the 2nd order system

$$x_{k+1} = Ax_k + Gw_k$$
$$y_k = Cx_k + w_k$$

where

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad G = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad C = \begin{bmatrix} 0.5 & 1 \end{bmatrix}$$

with $Q = 1$, $R = 1$. Note that this is almost the same system as Example 3.3.4, except that the observation noise is now w_k. Check if the relevant stabilizability and detectability conditions hold. If they do, determine the unique positive semidefinite solution of the algebraic Riccati equation (you may find it helpful to express the algebraic Riccati equation in the form suggested by $(ex3.1)$). Hence write down the steady state Kalman filter equations for the one-step ahead predictor.

5. Consider the system

$$x_{k+1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x_k + w_k$$
$$y_k = [1 \quad 0]x_k + v_k$$

w_k and v_k are independent Gaussian white noise processes with $cov(w_k) = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}$ and $cov(v_k) = 2$. Determine the equations satisfied by the components of the steady state error covariance with q_1 and q_2 as parameters. Solve the equations for the case $q_1 = 0$, $q_2 = 1$ and verify your answer using the routine dlqe on Matlab.

6. In this problem, we illustrate the application of Kalman filtering to problems in which processes may not described in state space form. Let y_k, the observed process, be given by

$$y_k = z_k + v_k$$

with z_k having spectral density $\Phi_z(\omega)$ given by

$$\Phi_z(\omega) = \frac{0.36(2 + 2\cos\omega)}{2.04 + 0.8\cos\omega + 2\cos 2\omega}$$

and v_k is zero mean white noise with variance 1. We would like to obtain the steady-state Kalman filter for the process z_k.

(a) Express $\Phi_z(\omega)$ in the form

$$\Phi_z(\omega) = \lambda \frac{C(e^{i\omega})C(e^{-i\omega})}{A(e^{i\omega})A(e^{-i\omega})}$$

with $\lambda > 0$. Interpret z_k as the solution of the ARMA equation

$$A(q^{-1})z_k = C(q^{-1})w_k$$

with w_k having variance λ. Now obtain a state space description for the process z_k.
Hint: The observable representation does not give the most convenient state space equations for computation. Try writing it in the form

$$x_{k+1} = Ax_k + Gw_k$$

$$z_k = [1\ \ 0]x_k = Cx_k$$

(b) Determine the steady-state Kalman filter for the estimate $\hat{z}_{k|k}$. You may use Matlab to solve the algebraic Riccati equation.

(c) Find the transfer function from y_k to $\hat{z}_{k|k}$. This then is the optimal steady-state filter in the frequency domain.

7. Recall that estimation problems of the form: estimate x_j based on y^k are generally classified into 3 categories: for $j > k$, it is called a prediction problem; for $j = k$, it is called a filtering problem; for $j < k$, it is called a smoothing problem. Smoothing problems are noncausal as far as information flow is concerned, but they arise in situations where estimates do not have to be generated in real-time. This problem shows how Kalman filtering equations may be used to solve smoothing problems.

(a) Consider the linear stochastic system

$$x_{k+1} = A_k x_k + w_k$$

$$y_k = C_k x_k + v_k$$

where w_k and v_k are zero mean independent sequences of random vectors with covariances Q and $R > 0$, respectively. For $j > 0$ fixed, determine $\hat{x}_{j|k}$, the linear least squares estimate of x_j based on observations y_t, $0 \le t \le k$ for all $k > j$, and determine the associated error covariance $P_{j|k} = E\{[x(j) - \hat{x}_{j|k}][x(j) - \hat{x}_{j|k}]^T\}$. This is called the fixed point smoothing problem.
(Hint: Introduce an auxiliary process ξ_k, $k \ge j$ by

$$\xi_{k+1} = \xi_k$$

$$\xi_j = x_j$$

Now examine the augmented process $z_k = \begin{bmatrix} x_k \\ \xi_k \end{bmatrix}$. What happens if you apply Kalman filtering to z_k?)

(b) Determine the improvement provided by smoothing by finding $P_{j/j-1} - P_{j|k}$ and verify that the improvement is nondecreasing with k.