2.160 IDENTIFICATION, ESTIMATION, AND LEARNING

LECTURE NOTES NO. 5

6. Continuous Kalman Filter (The Kalman-Bucy Filter)

6.1 Converting the Discrete-Time Filter to a Continuous-Time Filter

Consider a linear, continuous-time dynamical system given by

$$\dot{x}(t) = F(t)x + G(t)w(t) \tag{49}$$

$$y = H(t)x + v(t) \tag{50}$$

where $F(t)$ and $H(t)$ are, respectively, state transition and observation matrices, which are in general time-varying, and $w(t)$ and $v(t)$ are process and observation noise. As before, without loss of generality the input has been set to zero: $u(t) \equiv 0$.

The figure below compares the continuous-time dynamical system to the discrete-time system given by eqs. (5-4) and (5-5) in the previous section. Note that the process noise $w(t)$ leads to an integrator in the continuous time system, whereas the process noise w_t in the discrete time system is not integrated in the forward path of the block diagram. Therefore, the time integral of the continuous time process noise is analogous to the discrete-time process noise.

$$w_t = \int_{t-\Delta t}^{t} w(\tau)d\tau \tag{51}$$

where Δt is the sampling interval of the discrete time system. The covariance of the process noise can be computed using this formula.

$$\begin{aligned} Q_t &= E[w_t w_t^T] = E\left[\int_{t-\Delta t}^{t} w(\tau)d\tau \int_{t-\Delta t}^{t} w^T(\tau')d\tau'\right] = \int_{t-\Delta t}^{t}\int_{t-\Delta t}^{t} E[w(\tau)w^T(\tau')]d\tau' d\tau \\ &= \int_{t-\Delta t}^{t}[\int_{t-\Delta t}^{t} Q(\tau)\delta(\tau-\tau')d\tau']d\tau = \int_{t-\Delta t}^{t} Q(\tau)d\tau \cong Q(t)\Delta t \end{aligned} \tag{52}$$

where $\delta(\tau-\tau')$ is Dirac's delta function, which is zero unless $\tau = \tau'$. We assume that the process noise is uncorrelated in the continuous time system, too. Therefore, the process noise covariance Q_t is related to that of the continuous time as $Q_t = Q(t)\Delta t$.

The noise v_t in the discrete-time system, on the other hand, can be interpreted as a time average of the noise $v(t)$ in the continuous system.

$$v_t = \frac{1}{\Delta t}\int_{t-\Delta t}^{t} v(\tau)d\tau \tag{53}$$

The measurement covariance is therefore given by

$$
\begin{aligned}
R_t &= E[v_t v_t^T] = E\left[\frac{1}{\Delta t}\int_{t-\Delta t}^{t} v(\tau)d\tau \cdot \frac{1}{\Delta t}\int_{t-\Delta t}^{t} v^T(\tau')d\tau'\right] = \frac{1}{(\Delta t)^2}\int_{t-\Delta t}^{t}\int_{t-\Delta t}^{t} E[v(\tau)v^T(\tau')]d\tau' d\tau \\
&= \frac{1}{(\Delta t)^2}\int_{t-\Delta t}^{t}[\int_{t-\Delta t}^{t} R(\tau)\delta(\tau-\tau')d\tau']d\tau = \frac{1}{(\Delta t)^2}\int_{t-\Delta t}^{t} R(\tau)d\tau \cong \frac{1}{\Delta t}R(t)
\end{aligned} \tag{54}
$$

Again the measurement noise is assumed to be uncorrelated, and the process noise and measurement noise, too, are uncorrelated to each other; $E\left[w(t)w^T(s)\right] = Q(t)\delta(t-s)$, $E\left[v(t)v^T(s)\right] = R(t)\delta(t-s)$, and $E\left[v(t)w^T(s)\right] = 0$, for all t and s.

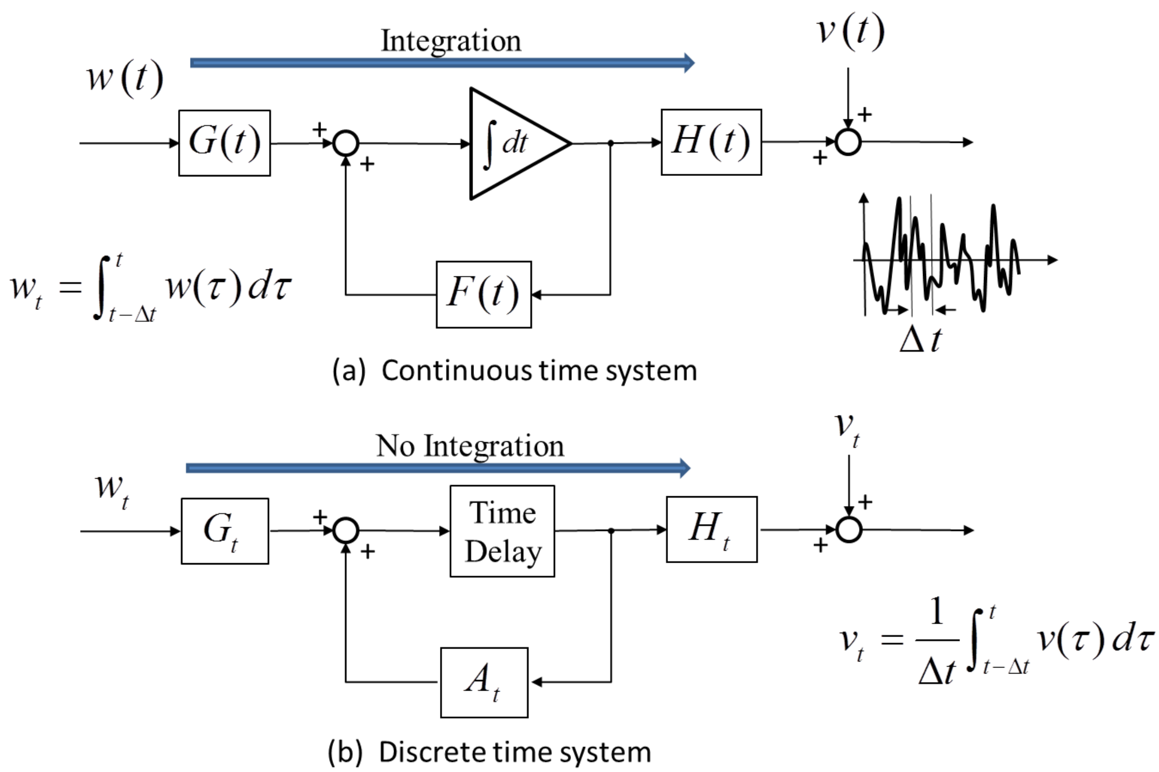


Figure 6-1 Comparison between continuous time and discrete time dynamical system with process and measurement noise.

From (38)

$$
\begin{aligned}
K_t &= P_{t|t-1}H_t^T\left(H_t P_{t|t-1}H_t^T + R_t\right)^{-1} \qquad R_t \to \frac{R}{\Delta t} \\
&= \Delta t P_{t|t-1}H_t^T\left(\Delta t \cdot H_t P_{t|t-1}H_t^T + R\right)^{-1} \\
&\cong \Delta t P_{t|t-1}H_t^T R^{-1} \qquad \text{for } |\Delta t| << 1
\end{aligned} \tag{55}
$$

Define $K \overset{\Delta}{=} P_{t|t-1}H_t^T R^{-1}$ (56)

From (45)

$$
\begin{aligned}
P_{t+1|t} &= A_t P_t A_t^T + G_t Q_t G_t^T \\
&= A_t \left(I - K_t H_t \right) P_{t|t-1} A_t^T + G_t Q_t G_t^T
\end{aligned}
\tag{57}
$$

$\Delta t K$

$A_t = I + \Delta t F$

Note that $\dfrac{x_{t+1} - x_t}{\Delta t} = F \cdot x_t + \cdots \Rightarrow x_{t+1} = (I + F\Delta t) \cdot x_t + \cdots$ is used in the above expression.

Ignoring higher-order small quantities; $O\left(\Delta t^2\right) \cong 0$ (58)

$$
P_{t+1|t} = P_{t|t-1} + \Delta t F P_{t|t-1} + \Delta t P_{t|t-1} F^T - \Delta t K H_t P_t + G_t \Delta t Q G_t^T \tag{59}
$$

$$
\frac{P_{t+1|t} - P_{t|t-1}}{\Delta t} = F P_{t|t-1} + P_{t|t-1} F^T - P_{t|t-1} H_t^T R^{-1} H_t P_t + G_t Q G_t^T \tag{60}
$$

$$
\Delta t \to 0 \qquad \lim_{\Delta t \to 0} P_{t|t-1} = P_{t-1} \tag{61}
$$

$$
\dot{P} = FP + PF^T - PH^T R^{-1} HP + GQG^T \tag{62}
$$

This is called the ***Matrix Riccati Equation***.

Similarly, we can reduce the discrete time form of state estimation correction to the one of continuous time. Combining (23) and (21) and replacing variables yield

$$
\dot{\hat{x}} = F\hat{x} + K\left(y - H\hat{x}\right) \tag{63}
$$

where the Kalman gain is given by

$$
K = PH^T R^{-1} \tag{64}
$$

This estimator given by (62), (63), and (64) is called the Kalman-Bucy Filter (1961).

The physical interpretation of the Matrix Riccati Equation

$$
\dot{P} = \underbrace{FP + PF^T} - PH^T R^{-1} HP + GQG^T \tag{62}
$$

Unforced State Transition: The effect of the unforced system dynamics upon the covariance propagation

The expected decrease of uncertainty as a result of measurement R

The expected increase of uncertainty due to the process disturbance Q

6.2 The Algebraic Riccati Equation

Assume that the Riccati differential equation has an asymptotically stable solution for $P(t)$:

$$\lim_{t\to\infty} P(t) = P_\infty \tag{65}$$

Then the time derivative vanishes

$$\lim_{t\to\infty} \frac{dP(t)}{dt} = 0 \tag{66}$$

Substituting this into the Riccati equation yields

$$0 = FP_\infty + P_\infty F^T - P_\infty H^T R^{-1} H P_\infty + GQG^T \tag{67}$$

This is called the ***Algebraic Riccati Equation***. This is a nonlinear matrix equation, and need a numerical solver to obtain a solution for P_∞.

Consider a scalar case; $P_\infty \in R^{1\times1}$, $F, H, Q, R, G \in R^{1\times1}$. The Algebraic Riccati Equation can be solved analytically

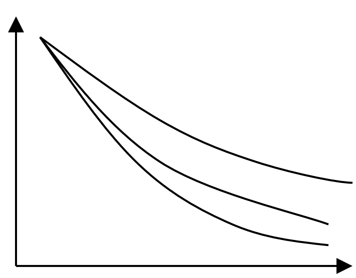
$$\frac{H^2}{R} P_\infty^2 - 2FP_\infty - G^2 Q = 0 \tag{68}$$

$$P_\infty = \frac{R}{H^2}\left(F \pm \sqrt{F^2 + \frac{Q}{R} H^2 G^2} \right) \tag{69}$$

There are two solutions; one positive and the other negative.
Taking the positive solution

$$\lim_{t\to\infty} P(t) = P_\infty = \frac{R}{H^2}\left(F + \sqrt{F^2 + \frac{Q}{R} H^2 G^2} \right) \tag{70}$$

Note that, regardless of the sign of F ($F < 0$ means a stable process dynamics), the above limit P_∞ is positive.



Remarks

1) As the sensor variance R increases, P_∞ increases

2) As the process noise variance Q increases, P_∞ increases

3) When the process noise variance Q is zero, and the process is stable, $F < 0$, P_∞ becomes zero.

6.3 Convergence Analysis

6.3.1 Transient Response of the Covariance Matrix

The Discrete Kalman Filter is hinged on the covariance matrix update law:

$$P_t = (I - K_t H_t) P_{t|t-1} \tag{41}$$

$$P_{t+1|t} = A_t P_t A_t^T + G_t Q_t G_t^T \tag{45}$$

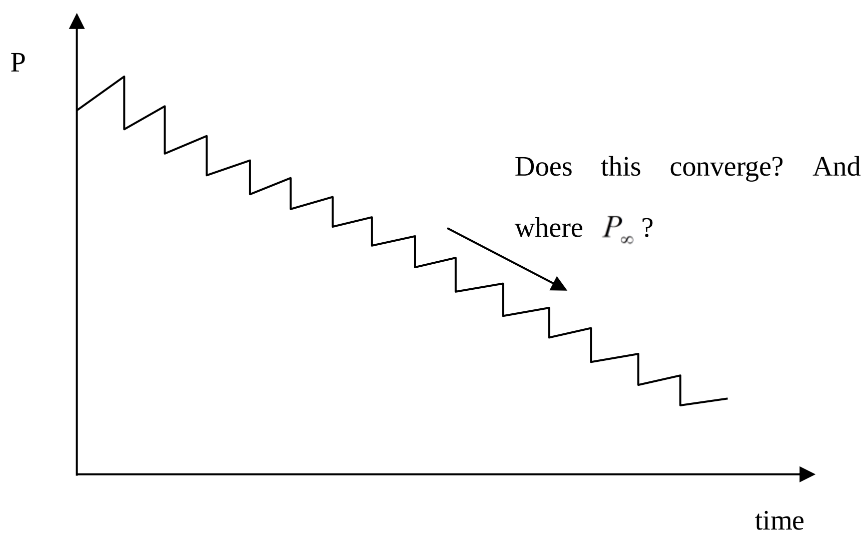


Figure 6-2 Time evolution of estimation error covariance

Continuous Kalman Filter:
The covariance matrix is given by the Riccati Differential equation:

$$\frac{d}{dt} P(t) = FP(t) + P(t)F^T - P(t)H^T R^{-1} HP(t) + GQG^T \tag{62}$$

where F is a state transition matrix:

$$\frac{d}{dt} x(t) = F(t)x(t) + G(t)w(t) \tag{49}$$

Let us examine the properties of the Riccati differential equation in order to gain insights as to whether the covariance of Kalman filter converges or not.

6.3.2 Matrix Fraction Decomposition

The Riccati Differential Equation (62) can be solved by using a technique, called the Matrix Fraction Decomposition

Consider a square matrix $L(t)$ decomposed into two square matrices *A(t)* and *B(t)*,

$$L(t) = A(t)B^{-1}(t) \tag{71}$$

where B is non-singular and both A and B are differentiable with respect to time t. The above expression is called a fraction decomposition of Matrix L.

Differentiating $B(t)B(t)^{-1} = I$ (identify matrix) with respect to time t,

$$\dot{B}B^{-1} + B\dot{B}^{-1} = 0$$

Therefore

$$\frac{d}{dt}B^{-1}(t) = -B^{-1} \cdot \frac{d}{dt}B(t) \cdot B^{-1} \tag{72}$$

Now let us represent the covariance matrix $P(t)$ by

$$P(t) = A(t)B^{-1}(t) \tag{73}$$

and applying eq.(72)

$$\begin{aligned} \frac{dP(t)}{dt} &= \dot{A}B^{-1} + A\dot{B}^{-1} \\ &= \dot{A}B^{-1} - AB^{-1}\dot{B}B^{-1} \end{aligned} \tag{74}$$

From the Riccati equation (62)

$$\frac{dP(t)}{dt} = FAB^{-1} + AB^{-1}F^T - AB^{-1}H^T R^{-1}HAB^{-1} + GQG^T \tag{75}$$

Equating the right hand sides of (74) and (75), and post-multiplying B yield

$$\dot{A} - AB^{-1}\dot{B} = \left(FA + GQG^T B\right) - AB^{-1}\left(H^T R^{-1}HA - F^T B\right) \tag{76}$$

Therefore, if we find A and B that satisfy:

$$\dot{A} = FA + GQG^T B$$

$$\dot{B} = H^T R^{-1} HA - F^T B \tag{77}$$

then $P(t) = A(t)B^{-1}(t)$ satisfies the Riccati differential equation. Note that (77) is linear differential equations with respect to matrices A and B . This can be rearranged as

$$\frac{d}{dt}\begin{pmatrix} A(t) \\ B(t) \end{pmatrix} = \overbrace{\begin{bmatrix} F(t) & G(t)Q(t)G^T(t) \\ H^T(t)R^{-1}(t)H(t) & -F(t)^T \end{bmatrix}}^{\text{A Hamiltonian Matrix, } \boldsymbol{M}}\begin{pmatrix} A(t) \\ B(t) \end{pmatrix} \tag{78}$$

As for the initial conditions, we can set

$$A(0) = P_0 \quad \text{and} \quad B(0) = I\,. \tag{79}$$

6.3.3 Convergence Properties of a Scalar, Time-Invariant Case

Consider a scalar case : $A(t) \to a(t)$, and $B(t) \to b(t)$

and assume that the process and measurement equations are time-invariant

$$\left.\begin{aligned} F(t) &= F \\ G(t) &= G \\ Q(t) &= Q \\ R(t) &= R \\ H(t) &= H \end{aligned}\right\} \text{Scalar}$$

Eq.(78) reduces to

$$\begin{pmatrix} \dot{a} \\ \dot{b} \end{pmatrix} = \begin{bmatrix} F & G^2 Q \\ H^2/R & -F \end{bmatrix}\begin{pmatrix} a \\ b \end{pmatrix} \tag{80}$$

This can be solved with initial condition of $a(0) = P_0$ and $b(0) = 1$.

The eigenvalues of the Hamiltonian Matrix are

$$\lambda_{1,}\lambda_2 = \pm\sqrt{F^2 + \frac{Q}{R}G^2H^2} = \pm\lambda \tag{81}$$

The solution of (80) is given by

$$\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = e^{\mathbf{M}\cdot t}\begin{pmatrix} P_0 \\ 1 \end{pmatrix} \tag{82}$$

Using the eigen vectors $\mathbf{v}_1$, $\mathbf{v}_2$ associated with eigenvalues λ_1, λ_2, respectively, the Hamiltonian matrix can be diagonalized as

$$\mathbf{M} = [\mathbf{v}_1, \mathbf{v}_2]\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}[\mathbf{v}_1, \mathbf{v}_2]^{-1} \tag{83}$$

Using this in the above solution yields

$$\begin{pmatrix} a(t) \\ b(t) \end{pmatrix} = [\mathbf{v}_1, \mathbf{v}_2]\cdot\begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix}[\mathbf{v}_1, \mathbf{v}_2]^{-1}\begin{pmatrix} P_0 \\ 1 \end{pmatrix}$$

This leads to

$$a(t) = \frac{1}{2\lambda}\left\{[P_0(\lambda + F) + q]e^{\lambda t} + [P_0(\lambda - F) - q]e^{-\lambda t}\right\}$$

$$b(t) = \frac{1}{2\lambda q}\left\{(\lambda - F)[P_0(\lambda + F) + q]e^{\lambda t} - (\lambda + F)[P_0(\lambda - F) - q]e^{-\lambda t}\right\} \tag{84}$$

where $q = G^2 Q$. Therefore, the covariance is given by

$$P(t) = \frac{a(t)}{b(t)} = q\frac{[P_0(\lambda + F) + q] + [P_0(\lambda - F) - q]e^{-2\lambda t}}{(\lambda - F)[P_0(\lambda + F) + q] - (\lambda + F)[P_0(\lambda - F) - q]e^{-2\lambda t}} \tag{85}$$

The steady-state solution is given by

$$P_\infty = \lim_{t\to\infty} P(t) = \frac{q}{\lambda - F} = \frac{R}{H^2}\left(F + \sqrt{F^2 + \frac{Q}{R}H^2G^2}\right)$$

This agrees with the previous result, eq.(70).

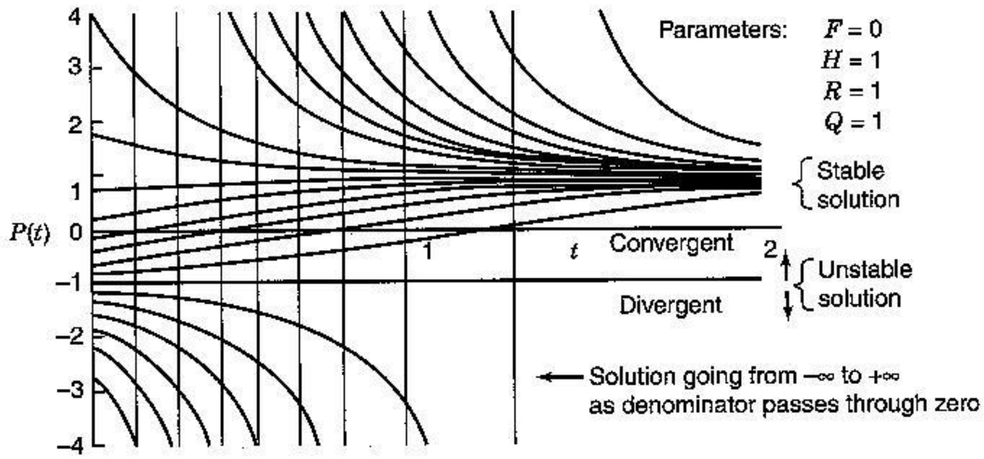


Figure 6-3 Solutions of the scalar, time-invariant Riccati equation (from Grewal and Andrews, "Kalman Filtering: Theory and Practice", Wiley 2001, Section 4.8,)

The figure above shows the solutions for parameter values of F=0, H=R=Q=1. Note that the denominator of (85) may become zero at time:

$$t_{diverge} = \frac{1}{2\lambda} \ln \frac{(\lambda + F)[P_0(\lambda - F) - q]}{(\lambda - F)[P_0(\lambda + F) + q]} \tag{86}$$

This implies that the solution becomes discontinuous, going from $-\infty$ to ∞ when passing the zero point. This undesirable discontinuity does not happen if $P_0 > \frac{R}{H^2}\left(F - \sqrt{F^2 + \frac{H^2}{R} G^2 Q} \right) \triangleq P_-$, the negative solution of the Algebraic Riccati Equation, as illustrated in the figure.

An important property of the Riccati Differential Equation (RDE):

If the system is observable, i.e. (F, H), Observable Pair, then the RDE has a positive-definite, symmetric solution for an arbitrary positive-definite initial value of matrix P_o>0;

$$\exists P(t) \ for \ \forall P_0 > 0 \ p.d., \quad \text{such that } P(t) > 0 \, p.d., P(t) = P^T(t) \in R^{n \times n}, \quad \forall t > 0, \tag{87}$$