COMPUTATIONAL ASPECTS OF ESTIMATION

Implementation of Linear Estimation

The Information filter implementation of Kalman filter

10/4/2022

Implementation of Linear Estimation

The two properties of a covariance matrix

- symmetry and
- positive definiteness

can be lost due to round-off errors in the course of calculating its propagation equations — the covariance prediction and the covariance update

Joseph form covariance update

The covariance propagation equations, propensity for causing loss of symmetry and/or loss of positive definiteness

The covariance prediction equation

$$P(k+1|k) = F(k)P(k|k)F(k)' + \Gamma(k)Q(k)\Gamma(k)'$$

can affect only the symmetry of the resulting matrix. A suitable implementation of the products of three matrices will avoid this problem.

More significant numerical problems arise in the covariance update equation

$$P(k+1|k+1) = [I - W(k+1)H(k+1)]P(k+1|k)$$

This is **very sensitive to round-off errors and is bound to lead to loss of symmetry as well as positive definiteness**

$$P(k+1|k+1) = P(k+1|k) - W(k+1)S(k+1)W(k+1)'$$

This will avoid loss of symmetry with a suitable implementation of the last term, which is a product of three matrices (, but it can still lead to loss of positive definiteness due to numerical errors in the subtraction.

$$P(k+1|k+1) = [I - W(k+1)H(k+1)]P(k+1|k)$$

$$\cdot[I - W(k+1)H(k+1)]' + W(k+1)R(k+1)W(k+1)'$$

Joseph form covariance update, while computationally more expensive, is less sensitive to round-off errors.

With the proper implementation of the products of three matrices, it will **preserve symmetry**. Furthermore, since the only place it has a subtraction is in the term I − WH, which appears “squared,” this form of the covariance update has the property of **preserving the positive definiteness** of the resulting updated covariance.

Information Filter

- carries out the recursive computation of the inverse of the covariance matrix.
- an alternative to the “standard“ Kalman filter formulation and is less demanding computationally for systems with dimension of the measurement vector larger than that of the state.
- has the advantage that it allows the start-up of the estimation without an initial estimate, i.e., with a noniformative prior
- Suitable for distributed estimation for example in swarm robotics

Recursions for the Information Matrices

The **standard version of the Kalman filter** calculates the gain in conjunction with a recursive computation of the **state covariance**

$$P(k|k-1) = F(k-1)P(k-1|k-1)F(k-1)' + \Gamma(k-1)Q(k-1)\Gamma(k-1)'$$

$$W(k) = P(k|k-1)H(k)'[H(k)P(k|k-1)H(k)' + R(k)]^{-1}$$

$$\begin{aligned} P(k|k) &= P(k|k-1) - P(k|k-1)H(k)'[H(k)P(k|k-1)H(k)' \\ &\quad + R(k)]^{-1}H(k)P(k|k-1) \end{aligned}$$

The **information filter** calculates recursively the **inverses of the covariance** matrices, **both for the prediction and the update**.

The term “information” is used in the sense of the Cramer-Rao lower bound, where the (Fisher) information matrix is the inverse of the covariance matrix

The update

$$P(k|k)^{-1} = P(k|k-1)^{-1} + H(k)'R(k)^{-1}H(k)$$

The prediction

$$A(k-1)^{-1} \triangleq F(k-1)P(k-1|k-1)F(k-1)'$$

$$A(k-1) = [F(k-1)^{-1}]'P(k-1|k-1)^{-1}F(k-1)^{-1}$$

$$\begin{aligned} P(k|k-1)^{-1} &= [F(k-1)P(k-1|k-1)F(k-1)' + \Gamma(k-1) \\ &\quad \cdot Q(k-1)\Gamma(k-1)']^{-1} \\ &= [A(k-1)^{-1} + \Gamma(k-1)Q(k-1)\Gamma(k-1)']^{-1} \end{aligned}$$

with the matrix inversion lemma 2

$$(P^{-1} + H'R^{-1}H)^{-1} = P - PH'(HPH' + R)^{-1}HP$$

$$P(k|k-1)^{-1} = A(k-1) - A(k-1)\Gamma(k-1)$$

$$\cdot[\Gamma(k-1)'A(k-1)\Gamma(k-1) + Q(k-1)^{-1}]^{-1}\Gamma(k-1)'A(k-1)$$

The expression of the gain can be written in form

$$W(k) = [P(k|k-1)^{-1} + H(k)'R(k)^{-1}H(k)]^{-1}H(k)'R(k)^{-1}$$

that is equivalent to the alternate form of the gain

$$W(k) = P(k|k)H(k)'R(k)^{-1}$$

Duality Between the Covariance and Information Equations

Covariance prediction $\longleftrightarrow$ Information update

Covariance update $\longleftrightarrow$ Information prediction

the following replacements show the **duality**
In Information Filter:

Prediction

$$P(k+1|k) \rightarrow P(k+1|k+1)^{-1}$$

$$A(k) \rightarrow P(k+1|k)$$

$$\Gamma(k) \rightarrow H(k)'$$

$$Q(k) \rightarrow R(k)^{-1}$$

Update

$$P(k+1|k)^{-1} \leftarrow P(k+1|k+1)$$

$$A(k) \leftarrow P(k+1|k)$$

$$\Gamma(k) \leftarrow H(k)'$$

$$Q(k)^{-1} \leftarrow R(k)$$

Overview of the Information Filter Algorithm

The state estimate can be computed in the same manner as in the conventional Kalman filter.

There is an alternative, however, that is discussed in the next slides

Updated State Information Matrix at $k-1$

$$P(k-1|k-1)^{-1}$$

Noiseless State Prediction Information Matrix

$$A(k-1) = \left[F(k-1)^{-1}\right]' P(k-1|k-1)^{-1} F(k-1)^{-1}$$

State Prediction Information Matrix

$$P(k|k-1)^{-1} = A(k-1) - A(k-1)\Gamma(k-1) \cdot \left[\Gamma(k-1)' A(k-1) \Gamma(k-1) + Q(k-1)^{-1}\right]^{-1} \Gamma(k-1)' A(k-1)$$

Gain

$$W(k) = \left[P(k|k-1)^{-1} + H(k)' R(k)^{-1} H(k)\right]^{-1} H(k)' R(k)^{-1}$$

Updated State Information Matrix at k

$$P(k|k)^{-1} = P(k|k-1)^{-1} + H(k)' R(k)^{-1} H(k)$$

Benefits of Information filter

- This implementation of the Kalman filter is advantageous when the **dimension n_z of the measurement** vector is **larger than the dimension n_x of the state**.
- Note that the inversions in the above sequence of calculations are for $n_x \times n_x$ matrices (and Q, which is $n_v \times n_v$, with $n_v \leq n_x$), while the conventional Kalman requires the inversion of the innovation covariance, which is $n_z \times n_z$. The
- inverse of the measurement noise covariance $R(k)$ is usually simple to obtain because it is, in many cases, diagonal.

Recursion for the Information Filter State

For the situations where there is no initial estimate, one can start the estimation with the initial information matrix as zero $P(0|0)^{-1} = 0$

This amounts to a ***noninformative (diffuse) prior*** because of the infinite uncertainty (infinite variance) associated with it. No initial estimate of the system's state $\hat{x}(0|0)$ is needed, see below

Alternative formulation for the state

Information filter state y is defined**,** one can obtain recursions from

$$\hat{y}(k-1|k-1) \triangleq P(k-1|k-1)^{-1}\hat{x}(k-1|k-1)$$

to

$$\hat{y}(k|k-1) \triangleq P(k|k-1)^{-1}\hat{x}(k|k-1)$$

and then to

$$\hat{y}(k|k) \triangleq P(k|k)^{-1}\hat{x}(k|k)$$

The prediction

$$
\begin{aligned}
\hat{y}(k|k-1) &= P(k|k-1)^{-1}\hat{x}(k|k-1) \\
&= [F(k-1)P(k-1|k-1)F(k-1)' \\
&\quad +\Gamma(k-1)Q(k-1)\Gamma(k-1)']^{-1}F(k-1)\hat{x}(k-1|k-1)
\end{aligned}
$$

...

$$
\begin{aligned}
&= [F(k-1)']^{-1}\{I - P(k-1|k-1)^{-1}F(k-1)^{-1}\Gamma(k-1) \\
&\quad \cdot[\Gamma(k-1)'(F(k-1)')^{-1}P(k-1|k-1)^{-1}F(k-1)^{-1}\Gamma(k-1) \\
&\quad +Q(k-1)^{-1}]^{-1}\Gamma(k-1)'(F(k-1)')^{-1}\hat{y}(k-1|k-1)
\end{aligned}
$$

$$
\boxed{
\begin{aligned}
\hat{y}(k|k-1) = (F(k-1)')^{-1}\{I - P(k-1|k-1)^{-1}F(k-1)^{-1}\Gamma(k-1) \\
\cdot[\Gamma(k-1)'A(k-1)\Gamma(k-1) + Q(k-1)^{-1}]^{-1} \\
\cdot\Gamma(k-1)'(F(k-1)')^{-1}\hat{y}(k-1|k-1)
\end{aligned}
}
$$

The update

$$\begin{aligned}\hat{y}(k|k) &= P(k|k)^{-1}\hat{x}(k|k) \\ &= [P(k|k-1)^{-1} + H(k)'R(k)^{-1}H(k)]\hat{x}(k|k-1) \\ &\quad +P(k|k)^{-1}P(k|k)H(k)'R(k)^{-1}[z(k) - H(k)\hat{x}(k|k-1)]\end{aligned}$$

$$= P(k|k-1)^{-1}\hat{x}(k|k-1) + H(k)'R(k)^{-1}z(k)$$

$$\boxed{\hat{y}(k|k) = \hat{y}(k|k-1) + H(k)'R(k)^{-1}z(k)}$$

one can recover the state estimate from the information filter state

$$\hat{x}(k|k) = P(k|k)\hat{y}(k|k)$$

THE CONTINUOUS-TIME LINEAR STATE ESTIMATION FILTER

The linear minimum mean square error (LMMSE) filter for this *continuous time* problem, known as the ***Kalman-Bucy filter***

The *duality* of the LMMSE estimation with the linear-quadratic (LQ) control problem is discussed. These two problems have their solutions determined by the same Riccati equation

The **continuous-time state equation** is (*without input, for simplicity*)

$$\dot{x}(t) = A(t)x(t) + D(t)\tilde{v}(t)$$

$$E[\tilde{v}(t)] = 0 \qquad E[\tilde{v}(t)\tilde{v}(\tau)'] = \tilde{Q}(t)\delta(t-\tau)$$

$\tilde{Q}(t)$ is the intensity of the white noise. This becomes the power spectral density if it is time-invariant

The **continuous measurement equation** is

$$z(t) = C(t)x(t) + \tilde{w}(t)$$

$$E[\tilde{w}(t)] = 0 \qquad E[\tilde{w}(t)\tilde{w}(\tau)'] = \tilde{R}(t)\delta(t-\tau)$$

The white process noise, the white measurement noise and the initial state are mutually uncorrelate

The continuous time filter will be derived by taking the limit of the discrete time filter as $\Delta \triangleq (t_{k+1} - t_k) \to 0$

The state equations are sometimes written in mathematically correct ***Ito differential equation*** form e.g.

$$dx(t) = A(t)x(t)\,dt + D(t)\,d\mathbf{w}_v(t)$$

where $d\mathbf{w}_v(t)$ is the infinitesimal increment of the Wiener process $\mathbf{w}_v(t)$.

Discrete-time state equation

$$x(t_{k+1}) = F(t_{k+1}, t_k)x(t_k) + v(t_k)$$

taking a small time increment $\Delta = t_{k+1} - t_k \quad \rightarrow 0$

$$\begin{aligned}\dot{x}(t_k) &= \frac{1}{\Delta}[x(t_k+\Delta) - x(t_k)] = \frac{1}{\Delta}[F(t_k+\Delta, t_k)x(t_k) + v(t_k) - x(t_k)] \\ &= \frac{1}{\Delta}[F(t_k+\Delta, t_k) - I]x(t_k) + \frac{1}{\Delta}v(t_k)\end{aligned}$$

$$F(t,t) = I$$

$$\frac{\partial}{\partial \tau}F(\tau, t) = A(\tau)F(\tau, t)$$

as $\Delta \rightarrow 0$ $\quad \frac{1}{\Delta}[F(t+\Delta, t) - I] = A(t)$

and the discrete-time process noise

$$v(t_k) = \int_{t_k}^{t_{k+1}} F(t_{k+1}, \tau) D(\tau) \tilde{v}(\tau) \, d\tau$$

the transition matrix over a very short interval tends to the identity matrix

$$D(t_k)\tilde{v}(t_k) = \frac{1}{\Delta} v(t_k)$$

The discrete state equation converges to continuous form.

The **corresponding discrete and continuous covariances**

$$Q(t_k) = E[v(t_k)v(t_k)'] = E\left[\left[\int_{t_k}^{t_k+\Delta} D(\tau_1)\tilde{v}(\tau_1)\, d\tau_1\right]\left[\int_{t_k}^{t_k+\Delta} D(\tau_2)\tilde{v}(\tau_2)\, d\tau_2\right]'\right]$$

…

$$D(t_k)\tilde{Q}(t_k)D(t_k)' = \frac{1}{\Delta} Q(t_k)$$

The ***discrete-time measurement*** can be viewed as a ***short-term average*** of the continuous-time measurement taking a small time increment

$$z(t_k) = \frac{1}{\Delta}\int_{t_k}^{t_k+\Delta} z(\tau)\,d\tau = C(t_k)x(t_k) + \frac{1}{\Delta}\int_{t_k}^{t_k+\Delta} \tilde{w}(\tau)\,d\tau$$

$$z(t_k) = H(t_k)x(t_k) + w(t_k) \qquad H(t_k) = C(t_k)$$

$$w(t_k) = \frac{1}{\Delta}\int_{t_k}^{t_k+\Delta} \tilde{w}(\tau)\,d\tau \qquad \Delta = t_{k+1} - t_k \;\longrightarrow 0$$

the covariance of the discrete time measurement noise and the intensity of the continuous time measurement noise

$$R(t_k) \;=\; E[w(t_k)w(t_k)'] = E\left[\frac{1}{\Delta^2}\int\int \tilde{w}(\tau_1)\tilde{w}(\tau_2)'\,d\tau_1\,d\tau_2\right]$$

$$\ldots \;=\; \frac{1}{\Delta}\tilde{R}(t_k) \qquad \tilde{R}(t_k) = R(t_k)\Delta$$

The Continuous-Time Filter, *Kalman-Bucy filter*

$$\hat{x}(k+1|k+1) - \hat{x}(k|k) = F(k)\hat{x}(k|k) - \hat{x}(k|k) + W(k+1)[z(k+1) - H(k+1)\hat{x}(k+1|k)]$$

$$\frac{1}{\Delta}[\hat{x}(t_{k+1}|t_{k+1}) - \hat{x}(t_k|t_k)] = \frac{1}{\Delta}[F(t_{k+1}, t_k) - I]\hat{x}(t_k|t_k) + P(t_{k+1}|t_{k+1})C(t_{k+1})'[R(t_{k+1})\Delta]^{-1} \cdot [z(t_{k+1}) - C(t_k)\hat{x}(t_{k+1}|t_k)]$$

$$\Delta = t_{k+1} - t_k \quad \rightarrow 0$$

$$\hat{x}(t_{k+1}|t_k) - \hat{x}(t_k|t_k) = O(\Delta) \qquad \hat{x}(t_{k+1}|t_k) \xrightarrow{\Delta \rightarrow 0} \hat{x}(t_k|t_k) \triangleq \hat{x}(t_k)$$

$$\boxed{\dot{\hat{x}}(t) = A(t)\hat{x}(t) + L(t)[z(t) - C(t)\hat{x}(t)]}$$

continuous-time filter gain $\quad L(t) \triangleq P(t)C(t)'\tilde{R}(t)^{-1}$

The Continuous-Time Filter, covariances

$$
\begin{aligned}
P(t_{k+1}|t_{k+1}) &= P(t_{k+1}|t_k) - P(t_{k+1}|t_{k+1})H(t_{k+1})'R(t_{k+1})^{-1}H(t_{k+1})P(t_{k+1}|t_k) \\
&= F(t_{k+1},t_k)P(t_k|t_k)F(t_{k+1},t_k)' + Q(t_k) \\
&\quad - P(t_{k+1}|t_{k+1})H(t_{k+1})'R(t_{k+1})^{-1}H(t_{k+1})P(t_{k+1}|t_k) \\
&= [I + A(t_k)\Delta]P(t_k|t_k)[I + A(t_k)\Delta]' + Q(t_k) \\
&\quad - P(t_{k+1}|t_{k+1})H(t_{k+1})'R(t_{k+1})^{-1}H(t_{k+1})P(t_{k+1}|t_k) \\
&= P(t_k|t_k) + A(t_k)P(t_k|t_k)\Delta + P(t_k|t_k)A(t_k)'\Delta + Q(t_k) \\
&\quad - P(t_{k+1}|t_{k+1})H(t_{k+1})'R(t_{k+1})^{-1}H(t_{k+1})P(t_{k+1}|t_k)
\end{aligned}
$$

$$
\begin{aligned}
\frac{1}{\Delta}[P(t_{k+1}|t_{k+1}) - P(t_k|t_k)] &= A(t_k)P(t_k|t_k) + P(t_k|t_k)A(t_k)' + \frac{1}{\Delta}Q(t_k) \\
&\quad - P(t_{k+1}|t_{k+1})H(t_{k+1})'[R(t_{k+1})\Delta]^{-1} \\
&\quad \cdot H(t_{k+1})P(t_{k+1}|t_k)
\end{aligned}
$$

the matrix **Riccati differential equation** for the state estimate Covariance

$$P(t_{k+1}|t_k) \xrightarrow{\Delta \to 0} P(t_{k+1}|t_{k+1}) \triangleq P(t_{k+1})$$

$$\boxed{\dot{P}(t) = A(t)P(t) + P(t)A(t)' + D(t)\tilde{Q}(t)D(t)' - P(t)C(t)'\tilde{R}(t)^{-1}C(t)P(t)}$$

The structure of the **Kalman-Bucy filter** is the same as that of the **Luenberger observer**

The ***continuous-time innovation***

$$\nu(t) \triangleq z(t) - C(t)\hat{x}(t) = C(t)x(t) + \tilde{w}(t) - C(t)\hat{x}(t) = C(t)\tilde{x}(t) + \tilde{w}(t)$$

Is zero mean and white

The steady-state covariance is the solution of the ***algebraic matrix Riccati equation***

$$AP + PA' + D\tilde{Q}D' - PC'\tilde{R}^{-1}CP = 0$$

which has a *unique positive definite solution*. To this corresponds a ***steady-state filter***

Kalman-Bucy Filter