Random Linear Systems

ECE 618
Lecture 9

Consider $\underline{x}_{k+1} = \underline{\Phi}_k \underline{x}_k + \underline{\Gamma}_k \underline{w}_k$

where $\underline{w}_k$ is white noise (in the sence of Gelb – i.e., Gaussian)

$E\{\underline{w}_k\} = \underline{0}$

$E\{\underline{w}_k \underline{w}_\ell^T\} = \underline{Q}_k \delta(k-\ell)$

$E\{\underline{x}_0\} = \underline{0}$

$\underline{w}_k$ and $\underline{x}_0$ Gaussian and independent r.v. $\forall\, k$.

$\underline{w}_k, \underline{w}_\ell$ indep. $\forall\, k \neq \ell$.

Then

$$E\{\underline{x}_{k+1}\} = E\{\underline{\Phi}_k \underline{x}_k + \underline{\Gamma}_k \underline{w}_k\}$$

$$= \underline{\Phi}_k E\{\underline{x}_k\} + \underline{\Gamma}_k E\{\underline{w}_k\} = \underline{0} \quad \text{(by induction)}$$

If we have an estimate $\hat{\underline{x}}_k$ of $\underline{x}_k$, define error in estimate as

$$\underline{e}_k = \hat{\underline{x}}_k - \underline{x}_k$$

Assume for the moment that $E\{e_k\} = 0$ (estimate is unbiased).

Then $\mathrm{cov}\{\underline{e}_k\} = \mathrm{cor}\{\underline{e}_k\} = E\{\underline{e}_k \underline{e}_k^T\}$ is a measure of uncertainty about the true value of $\underline{x}_k$.

Suppose $\underline{x}_k$ is known and an <u>unbiased</u> $\left(E\{\underline{e}_{k+1}\} = \underline{0}\right)$ estimate of $\underline{x}_{k+1}$, $\hat{\underline{x}}_{k+1}$, is desired

This estimate is given by

$$\hat{\underline{x}}_{k+1} = \underline{\Phi}_k \underline{x}_k .$$

<u>Proof</u>: $\underline{e}_{k+1} = \hat{\underline{x}}_{k+1} - \underline{x}_{k+1} = \underline{\Phi}_k \underline{x}_k - \underline{\Phi}_k \underline{x}_k - \underline{\Gamma}_k \underline{w}_k = -\underline{\Gamma}_k \underline{w}_k$

and $E\{\underline{e}_{k+1}\} = E\{-\underline{\Gamma}_k \underline{w}_k\} = -\underline{\Gamma}_k E\{\underline{w}_k\} = \underline{0}$

If a <u>known</u> input is given at time k, then

$$\hat{\underline{x}}_{k+1} = \underline{\Phi}_k \underline{x}_k + \underline{B}_k \underline{u}_k, \text{ where } \underline{x}_{k+1} = \underline{\Phi}_k \underline{x}_k + \underline{\Gamma}_k \underline{w}_k + \underline{B}_k \underline{u}_k$$

This is the <u>one-step predictor</u> of the state.

Suppose $\underline{P}_k = E\{\underline{e}_k \underline{e}_k^T\}$ is known (covariance of error in estimate at time k), along with $\hat{\underline{x}}_k$.

Then $\underline{P}_{k+1} = E\{\underline{e}_{k+1} \underline{e}_{k+1}^T\}$ and

$$\underline{e}_{k+1} \underline{e}_{k+1}^T = \left(\underline{\Phi}_k \hat{\underline{x}}_k - \underline{\Phi}_k \underline{x}_k - \underline{\Gamma}_k \underline{w}_k\right)\left(\underline{\Phi}_k \hat{\underline{x}}_k - \underline{\Phi}_k \underline{x}_k - \underline{\Gamma}_k \underline{w}_k\right)^T$$

$$= \left(\underline{\Phi}_k \underline{e}_k - \underline{\Gamma}_k \underline{w}_k\right)\left(\underline{\Phi}_k \underline{e}_k - \underline{\Gamma}_k \underline{w}_k\right)^T$$

$$= \underline{\Phi}_k \underline{e}_k \underline{e}_k^T \underline{\Phi}_k^T - \underline{\Gamma}_k \underline{w}_k \underline{e}_k^T \underline{\Phi}_k^T - \underline{\Phi}_k \underline{e}_k \underline{w}_k^T \underline{\Gamma}_k^T + \underline{\Gamma}_k \underline{w}_k \underline{w}_k^T \underline{\Gamma}_k^T$$

Since w_k is white and assumed uncorrelated with $\underline{e}_k$,

$$\underline{P}_{k+1} = E\{\underline{e}_{k+1}\underline{e}_{k+1}^T\} = \underline{\Phi}_k E\{\underline{e}_k\underline{e}_k^T\}\underline{\Phi}_k^T + \underline{\Gamma}_k E\{\underline{w}_k\underline{w}_k^T\}\underline{\Gamma}_k^T$$

(the other 2 terms are zero.)

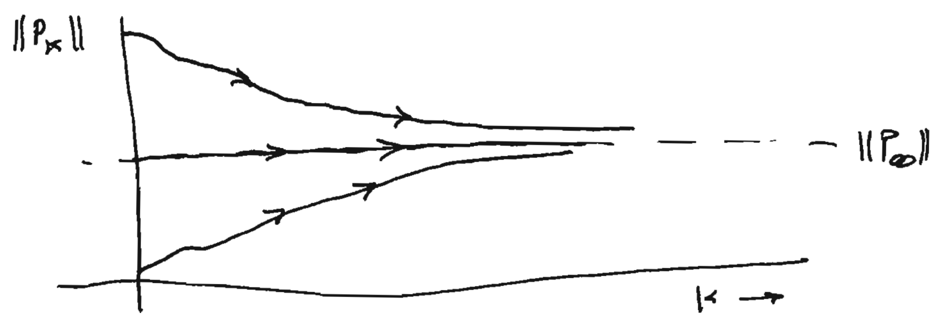
or,

$$\boxed{\underline{P}_{k+1} = \underline{\Phi}_k\underline{P}_k\underline{\Phi}_k^T + \underline{\Gamma}_k\underline{Q}_k\underline{\Gamma}_k^T}$$

Discrete-Time Linear Variance Equation

The DTLVE relates error in the estimate at time k+1 to error in the estimate at time k and the effect of additive white process noise.

We will learn that, if $\{\underline{\Phi}_k\}_{k=0}^{\infty}$ is stable, $\|P_k\|$ is bounded (and converges for constant $\underline{\Phi}_k$ and Q_k).



In continuous time,

$$\dot{\underline{x}}_t = \underline{F}_t \underline{x}_t + \underline{G}_t \underline{w}_t$$

$$\underline{\Gamma}_k \underline{Q}_k \underline{\Gamma}_k^T = \int_{t_k}^{t_{k+1}} \underline{\Phi}(t_{k+1}, \tau) \underline{G}(\tau) \underline{Q}(\tau) \underline{G}^T(\tau) \underline{\Phi}^T(t_{k+1}, \tau) \, d\tau$$

If $\Delta t = t_{k+1} - t_k \to 0$, then

$$\underline{\Gamma}_k \underline{Q}_k \underline{\Gamma}_k^T \to \underline{G}\underline{Q}\underline{G}^T \Delta t$$

$\underline{\Phi}(t)$ satisfies $\dot{\underline{\Phi}}(t) = \underline{F}_t \underline{\Phi}(t)$ (actually, $\underline{\Phi}(t, \tau)$)

or $$\frac{\underline{\Phi}(t+\Delta t) - \underline{\Phi}(t)}{\Delta t} \Rightarrow F(t)\underline{\Phi}(t) \quad \text{as } \Delta t \to 0$$

or $$\underline{\Phi}_k \longrightarrow \underline{I} + \underline{F}_t \Delta t$$

Therefore,

$$\begin{aligned}
\underline{P}_{k+1} &= (\underline{I} + \underline{F}\Delta t)\underline{P}_k(\underline{I} + \underline{F}\Delta t)^T + \underline{G}\underline{Q}\underline{G}^T \Delta t \\
&= \underline{P}_k + (\underline{F}\underline{P}_k + \underline{P}_k\underline{F}^T + \underline{G}\underline{Q}\underline{G}^T)\Delta t + \underline{F}\underline{P}_k\underline{F}^T(\Delta t)^2
\end{aligned}$$

or $$\frac{\underline{P}_{k+1} - \underline{P}_k}{\Delta t} = \underline{F}\underline{P}_k + \underline{P}_k\underline{F}^T + \underline{G}\underline{Q}\underline{G}^T + \underline{F}\underline{P}_k\underline{F}^T \Delta t$$

as $\Delta t \to 0$:

$$\boxed{\dot{\underline{P}}_t = \underline{F}_t \underline{P}_t + \underline{P}_t \underline{F}_t^T + \underline{G}_t \underline{Q}_t \underline{G}_t^T}$$

Continuous-Time Linear Variance Equation

So, what is (state) estimation — at least in the linear case?

(1) We know that if no measurements are available the unbiased estimator is:

$$\underline{\hat{x}}_{k+1} = \underline{\Phi}_k \underline{\hat{x}}_k$$

(+ deterministic inputs, if present)

or

$$\underline{\dot{x}}_t = \underline{F}_t \underline{x}_t$$

This is the (optimal) one-step predictor.

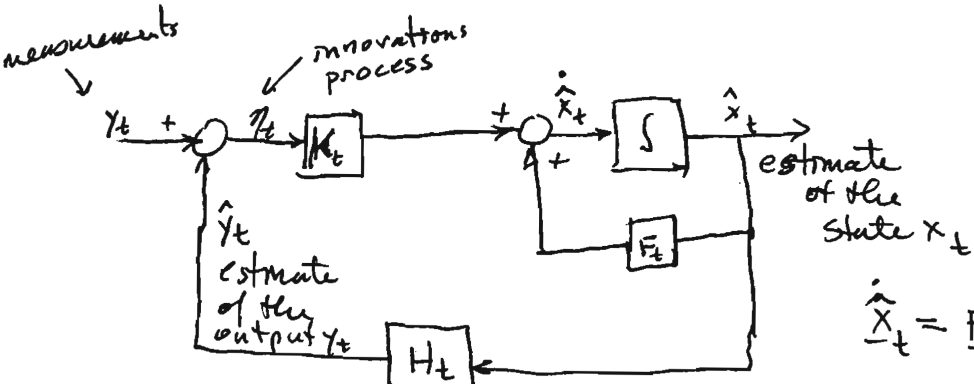
(2) The estimation (prediction) error propagates (again, without measurements) according to:

$$\underline{P}_{k+1} = \underline{\Phi}_k \underline{P}_k \underline{\Phi}_k^T + \underline{\Gamma}_k \underline{Q}_k \underline{\Gamma}_k^T$$

or

$$\underline{\dot{P}}_t = \underline{F}_t \underline{P}_t + \underline{P}_t \underline{F}_t^T + \underline{G}_t \underline{Q}_t \underline{G}_t^T$$

(3) Processes are linear, and noises are Gaussian zero-mean, so it seems reasonable to expect the effect of measurements on estimates to be combined linearly with prediction:



$$\underline{\dot{x}}_t = \underline{F}_t \underline{x}_t + \underline{G}_t \underline{w}_t$$

$$\underline{y}_t = \underline{H}_t \underline{x}_t + \underline{v}_t$$

$$\underline{\dot{\hat{x}}}_t = \underline{F}_t \underline{\hat{x}}_t + \underline{K}_t (\underline{y}_t - \underline{H}_t \underline{\hat{x}}_t)$$

In discrete time,

$$\hat{x}_{k+1} = \underline{\Phi}_k \hat{\underline{x}}_k + \underline{K}_k \left[\cancel{\underline{\Phi}_k \hat{\underline{x}}_k} \; y_k - \underline{H}_k \hat{\underline{x}}_k \right]$$

where

$$\underline{x}_{k+1} = \underline{\Phi}_k \underline{x}_k + \underline{\Gamma}_k \underline{w}_k$$

$$y_k = \underline{H}_k \underline{x}_k + \underline{v}_k \qquad E[\underline{v}_k] = 0, \; E[\underline{v}_k \underline{v}_\ell^T] = \underline{R}_k \, \delta(k-\ell)$$

Let's look ahead at what happens to the error

$$\underline{e}_k = \hat{\underline{x}}_k - \underline{x}_k$$

$$\begin{aligned}
\underline{e}_{k+1} &= \hat{\underline{x}}_{k+1} - \underline{x}_{k+1} \\
&= \underline{\Phi}_k \hat{\underline{x}}_k + \underline{K}_k \left[y_k - \underline{H}_k \hat{\underline{x}}_k \right] - \underline{\Phi}_k \underline{x}_k - \underline{\Gamma}_k \underline{w}_k \\
&= \left[\underline{\Phi}_k - \underline{K}_k \underline{H}_k \right] \hat{\underline{x}}_k - \left[\underline{\Phi}_k - \underline{K}_k \underline{H}_k \right] \underline{x}_k - \underline{\Gamma}_k \underline{w}_k + \underline{K}_k \underline{v}_k
\end{aligned}$$

$$(y_k = \underline{H}_k \underline{x}_k + \underline{v}_k)$$

$$\boxed{\underline{e}_{k+1} = \left[\underline{\Phi}_k - \underline{K}_k \underline{H}_k \right] \underline{e}_k - \underline{\Gamma}_k \underline{w}_k + \underline{K}_k \underline{v}_k}$$

<u>Observation 1</u>: $\underline{K}_k$ must be chosen such that $\{ \underline{\Phi}_k - \underline{K}_k \underline{H}_k \}$ is asymptotically stable in order for $\{ \underline{e}_k \}$ to decay.

Suppose $\underline{P}_k = E\{\underline{e}_k \underline{e}_k^T\}$ is known, as before.

$$E\left[\underline{e}_{k+1}\underline{e}_{k+1}^T\right] = E\left\{\left[[\underline{\Phi}_k - \underline{K}_k\underline{H}_k]\underline{e}_k - \underline{\Gamma}_k\underline{w}_k + \underline{K}_k\underline{v}_k\right]\left[-\text{same}-\right]^T\right\}$$

$$= [\underline{\Phi}_k - \underline{K}_k\underline{H}_k]E[\underline{e}_k\underline{e}_k^T][\underline{\Phi}_k - \underline{K}_k\underline{H}_k]^T$$

$$+ \underline{\Gamma}_k E[\underline{w}_k\underline{w}_k^T]\underline{\Gamma}_k^T + \underline{K}_k E[\underline{v}_k\underline{v}_k^T]\underline{K}_k^T$$

(Assuming $\underline{e}_k$, $\underline{w}_k$, & $\underline{v}_k$ are independent & zero-mean)

or,

$$\boxed{\underline{P}_{k+1} = [\underline{\Phi}_k - \underline{K}_k\underline{H}_k]\underline{P}_k[\underline{\Phi}_k - \underline{K}_k\underline{H}_k]^T + \underline{\Gamma}_k\underline{Q}_k\underline{\Gamma}_k^T + \underline{K}_k\underline{R}_k\underline{K}_k^T}$$

<u>Observation 2</u>: Three things influence covariance at $k+1$:

(i) Process noise covariance.

(ii) Measurement noise covariance.

(iii) Noise covariance in previous estimate.

<u>Observation 3</u>: This equation is quadratic in $\underline{K}_k$, and quadratic equations — at least scalar ones, — either have a minimum or a maximum.

(In the matrix case, there are lots of solutions $\underline{K}_k$ which yield minima, maxima, or inflection points, but it will turn out we are interested in only one.)

Therefore,

Observation 4: It should be possible to select a $\underline{K}_k$ (which we already know must stabilize $\Phi_k - K_k H_k$) to minimize the covariance of the error in the estimate at time $k+1$.

In other words, there is an optimal linear combination of the prediction and the innovations process that minimizes estimate error.

This is the Kalman Filter (in discrete time).