Estimation: Recap

Given the discrete time system

$$x_{k+1} = A_k x_k + B_k u_k + G_k\, \eta_k;$$

$$y_k = H_k x_k + \omega_k;$$

With assumptions

- A_k, B_k, H_k are constant matrices
- x_k is the state at time t_k; u_k is the control at time t_k
- the initial state x_0 is Gaussian distributed with mean $\bar{x}_0$ and covariance P_0

$$p(x_0) = \frac{1}{\sqrt{(2\pi)^n |P_0|}} e^{-\frac{1}{2}(x_0 - \bar{x}_0)^T P_0^{-1} (x_0 - \bar{x}_0)} \qquad \text{where } |P_0| = \det(P_0)$$

- Both η_k, ω_k are zero mean, Gaussian, and "white" random processes
  - $\eta_k \sim N(0, Q_k)$; $\omega_k \sim N(0, R_k)$
  - $E\left[\eta_k\, \eta_l^T\right] = Q_k\, \delta_{kl}$; $E\left[\omega_k\, \omega_l^T\right] = R_k\, \delta_{kl}$ (uncorrelated across time)
  - η_k, ω_k, and x_0 are independent (which implies uncorrelated)

Find state estimate which minimizes a loss function:

Jointly Gaussian Random Variables

If CRVs $x_1, x_2, \cdots, x_n$ are *jointly distributed*, then ∃ $p(x_1, \cdots, x_n)$ such that:

$$\mathrm{P}(x_{1l} \leq x_1 \leq \ x_{1u};\ \cdots; x_{nl} \leq x_n \ \leq \ x_{nu}) \equiv \int_{x_{1l}}^{x_{1u}} \cdots \int_{x_{nl}}^{x_{nu}} p(x_1, \dots, x_n) dx_1 \cdots dx_n$$

Defn: A collection of CRVs $x_1, x_2, \cdots, x_n$ are "jointly Gaussian" if

$$\sum_{i=1}^{n} a_i x_i$$

Is a Gaussian CRV for any real $\{a_i\}\ i = 1, \dots, n$

Joint Gaussian Variables can be represented by a joint pdf: $\vec{x} = N(\mu, K)$

– $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix};\qquad \vec{\mu} = E[\vec{x}];\qquad K = E[(\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})^T]$

Affine transformation of jointly Gaussian CRVs is jointly Gaussian CRV:

– Let $\vec{x} \sim N(\mu, K)$. Then $\vec{A}\{x\} + \vec{b} \sim N(A\vec{\mu} + \vec{b}, AKA^T)$

Jointly Gaussian Random Variables

Partial Proof: (see Anderson & Moore Appendix A for full details)

- **Mean:** $E[A\vec{x} + b] = E[A\vec{x}] + E[\vec{b}] = A\,E[\vec{x}] + b = A\vec{\mu} + b;$

- **Variance:**

$$E\big[\,[(A\vec{x} + b) - (A\vec{\mu} + b)][(A\vec{x} + b) - (A\vec{\mu} + b)]\big] =$$
$$E[\,[A(\vec{x} - \vec{\mu})][A(\vec{x} - \vec{\mu})]^T\,] = A\,E[(\vec{x} - \vec{\mu})(\vec{x} - \vec{\mu})^T]A^T = A\,K\,A^T$$

Fact: Jointly Gaussian RVs are independent iff they are uncorrelated

Conditional Density of Jointly Gaussian RVs

Assume that x (state) and y (measurements) are jointly Gaussian RVs Let $z = [\vec{x} \;\; \vec{y}]^T$

- *Mean:* $\bar{z} = E\begin{bmatrix}\vec{x}\\ \vec{y}\end{bmatrix} = E\begin{bmatrix}\bar{x}\\ \bar{y}\end{bmatrix}$ Variance: $Q_{zz} = \begin{bmatrix}Q_{xx} & Q_{xy}\\ Q_{yx} & Q_{yy}\end{bmatrix}$, $Q_{xy} = Q_{yx}^T$

Goal: find $p(\vec{x}|\vec{y})$, which will define the probability of the state given measurements (with n=dim(x) and m=dim(y)):

$$p(x|y) = \frac{p(x,y)}{p(y)} = \frac{[(2\pi)^{n+m}|Q_{zz}|]^{-1/2}}{\left[(2\pi)^{m}\left|Q_{yy}\right|\right]^{-1/2}} \frac{e^{-\frac{1}{2}(z-\bar{z})^T Q_{zz}^{-1}(z-\bar{z})}}{e^{-\frac{1}{2}(y-\bar{y})^T Q_{yy}^{-1}(y-\bar{y})}}$$

How do we simplify this?

- Find expression for $|Q_{zz}|$
- Simplify exponents

Conditional Density (continued)

$$p(\vec{x}|\vec{y}) = \frac{p(\vec{x},y)}{p(y)} = \frac{[(2\pi)^{n+m}\boxed{|Q_{zz}|}]^{-1/2}}{\left[(2\pi)^m |Q_{yy}|\right]^{-1/2}} \frac{e^{-\frac{1}{2}(z-\bar{z})^T Q_{zz}^{-1}(z-\bar{z})}}{e^{-\frac{1}{2}(y-\bar{y})^T Q_{yy}^{-1}(y-\bar{y})}}$$

Determinant:

If $T = \begin{bmatrix} I & -Q_{xy}Q_{yy}^{-1} \\ 0 & I \end{bmatrix}$, then $\quad TQ_{zz}T^T = \begin{bmatrix} (Q_{xx} - Q_{xy}Q_{yy}^{-1}Q_{xy}^T) & 0 \\ 0 & Q_{yy} \end{bmatrix} \equiv \Sigma$

Then: $\det(TQ_{zz}T^T) = \det(T)\det(Q_{zz})\det(T) = \det(Q_{zz})$

$$= \det\left(Q_{yy}\right)\det\left(Q_{xx} - Q_{xy}Q_{yy}^{-1}Q_{xy}^T\right) = \boxed{|Q_{yy}|\,|Q_{x|y}|}$$

Exponent:

$$(\vec{z} - \bar{z})^T Q_{zz}(\vec{z} - \bar{z}) = \begin{bmatrix} \vec{x} - \bar{x} \\ \vec{y} - \bar{y} \end{bmatrix}^T (T^{-1}\Sigma T^{-T})^{-1} \begin{bmatrix} \vec{x} - \bar{x} \\ \vec{y} - \bar{y} \end{bmatrix}$$

$$= \left(\vec{x} - \bar{\mu}_{x|y}\right)^{\mathrm{T}} \mathrm{Q}_{x|y}^{-1}\left(\vec{x} - \bar{\mu}_{x|y}\right) + (\vec{y} - \bar{y})^{\mathrm{T}} \mathrm{Q}_{yy}^{-1}(\vec{y} - \bar{y})$$

Where: $\quad \bar{\mu}_{x|y} = \bar{x} + Q_{xy}Q_{yy}^{-1}(\vec{y} - \bar{y})$

$\quad Q_{x|y} = Q_{xx} - Q_{xy}Q_{yy}^{-1}Q_{xy}^T$

Conditional Density (continued)

Substitute the newly derived expressions into p(x|y) to yield:

$$p(\vec{x}|\vec{y}) = \frac{\left[(2\pi)^{n+m}\left|\cancel{Q_{yy}}\right|\left|Q_{x|y}\right|\right]^{-1/2}}{\left[(2\pi)^{m}\left|\cancel{Q_{yy}}\right|\right]^{-1/2}} \frac{e^{-\frac{1}{2}\left[\left(\vec{x}-\bar{\mu}_{x|y}\right)^{\mathrm{T}} \mathrm{Q}_{x|y}^{-1}\left(\vec{x}-\bar{\mu}_{x|y}\right)+\cancel{(\vec{y}-\bar{y})^{\mathrm{T}} \mathrm{Q}_{yy}^{-1}(\vec{y}-\bar{y})^{T}}\right]}}{\cancel{e^{-\frac{1}{2}(y-\bar{y})^{T} Q_{yy}^{-1}(y-\bar{y})}}}$$

$$= \frac{1}{\left[(2\pi)^{n}\left|Q_{x|y}\right|\right]^{1/2}} e^{-\frac{1}{2}\left(\vec{x}-\bar{\mu}_{x|y}\right)^{T} Q_{x|y}^{-1}\left(\vec{x}-\bar{\mu}_{x|y}\right)}$$

Result: The Conditional Density *is a Gaussian* pdf

Next: Develop expressions for each of the key terms

Linear Discrete Time Systems

Consider the system:

Add constant term later

$$\vec{x}_{k+1} = A_k\vec{x}_k + \cancel{B_k\vec{u}_k} + G_k\,\vec{\eta}_k \;=\; A_k\vec{x}_k + G_k\,\vec{\eta}_k;$$

Where η_k is white, Gaussian, and zero mean:

$$\vec{x}_{k+2} = A_{k+1}\vec{x}_{k+1} + G_{k+1}\;\vec{\eta}_{k+1} \;=\; A_{k+1}A_k\vec{x}_k + A_{k+1}G_k\,\vec{\eta}_k + G_{k+1}\,\vec{\eta}_{k+1};$$

$$\vdots$$

$$\vec{x}_n = \Phi_{n,0}x_0 + \sum_{l=0}^{n-1}\Phi_{n,l+1}\,G_l\,\vec{\eta}_l \qquad (*)$$

Where $\Phi_{k,l} = A_{k-1}A_{k-2}\cdots A_l$ and $\Phi_{k,k} = I$

From (*), $\vec{x}_n$ is a linear combination of $\vec{x}_0$(Gaussian RV) and linear transform of process noise samples (drawn from Gaussian). Since they are assumed *independent,* they are “jointly Gaussian”

Linear Discrete Time Systems

Proposition: (see Appendix A of Anderson & Moore)

- Linear transformation and additions of Gaussian RVs are Gaussian RVs
- $\therefore \vec{x}_n$ is a Gaussian RV

Measurement Equation

- $\vec{y}_k = H_k \vec{x}_k + \vec{\omega}_k$ is the sum of two independent Gaussians, and therefore jointly Gaussian
- $\therefore \vec{y}_n$ is a Gaussian RV

Means of these Gaussians:

- $\bar{x}_n = E[\vec{x}_n] = E[\Phi_{n,0}\vec{x}_0 + \sum_{l=0}^{n-1} \Phi_{n,l+1}\, G_l\, \vec{\eta}_l]$
  $= \Phi_{n,0}E[\vec{x}_0] + \sum_{l=0}^{n-1} \Phi_{n,l+1}\, G_l\, E[\vec{\eta}_l] = \Phi_{n,0}\bar{x}_0$
    - *Note Recursion*: $\bar{x}_{n+1} = A_n \bar{x}_n$
- $E[\vec{y}_k] = E[H_k \vec{x}_k + \vec{\omega}_k] = H_k E[\vec{x}_k] + E[\vec{\omega}_k] = H_k \bar{x}_k$

Linear Discrete Time Systems

Covariance of the Gaussian state distribution:

– $P_{k,l} = E[(\vec{x}_k - \bar{x}_k)(\vec{x}_l - \bar{x}_l)^T]$

$$= E\{[\Phi_{k,0}(\vec{x}_0 - \bar{x}_0) + \sum_{m=0}^{k-1} \Phi_{k,m+1}\, G_m\, \vec{\eta}_m][\Phi_{l,0}(\vec{x}_0 - \bar{x}_0) + \sum_{n=0}^{l-1} \Phi_{l,n+1}\, G_n\, \vec{\eta}_n]^T\}$$

– But $(\vec{x}_0 - \bar{x}_0)$, and $\vec{\eta}_0, \vec{\eta}_1, \ldots, \vec{\eta}_{k-1}$ are *independent,* while $\vec{\eta}_0, \vec{\eta}_1, \ldots, \vec{\eta}_{k-1}$ are uncorrelated. The expectation of many product terms will be zero.

– E.g. $E\left[\vec{x}_0 \vec{\eta}_k^T\right] = E[\vec{x}_0]E\left[\vec{\eta}_k^T\right] = E[\vec{x}_k] \cdot 0 = 0$ (since η_k is zero mean)

– E.g. $E\left[\vec{\eta}_j \eta_k^T\right] = 0$ for $j \neq k$

– $P_{k,l} = \Phi_{k,0} E[(\vec{x}_0 - \bar{x}_0)(\vec{x}_0 - \bar{x}_0)^T]\Phi_{l,0}^T + \sum_{m=0}^{l-1} \Phi_{k,m+1} G_m Q_m G_m^T \Phi_{l,m+1}^T$

– We will be particularly interested in $P_{k,k}$, which can be found recursively!

- $P_{k+1,k+1} = \underbrace{A_k P_{k,k} A_k^T}_{\substack{\text{Dynamic} \\ \text{covariance} \\ \text{propagation}}} + \underbrace{G_k Q_k G_k^T}_{\substack{\text{Effect of} \\ \text{Process} \\ \text{Noise}}}$ (state covariance propagation)

Linear Discrete Time Systems

Covariance of the Gaussian measurement distribution:

- $\vec{y}_k = H_k \vec{x}_k + \vec{\omega}_k$

- $cov(y_k, y_l) = E[(\vec{y}_k - \bar{y}_k)(\vec{y}_l - \bar{y}_l)^T]$
  $= E\{[H_k(\vec{x}_k - \bar{x}_k) + \vec{\omega}_k][H_k(\vec{x}_k - \bar{x}_k) + \vec{\omega}_k]^T\}$

- But ω_k is independent of $(\vec{x}_k - \bar{x}_k)$ since $\vec{x}_{\boldsymbol{k}}, \bar{x}_k$ are functions of $\vec{x}_0$ and $\eta_0, \eta_1, \dots, \eta_{k-1}$, which are independent of ω_k

- $\therefore cov(y_k, y_l) = H_k \Phi_{k,0} P_{0,0} \Phi_{l,0}^T H_l^T + R_k \delta_{k,l}$

- Recursion: $cov(y_k, y_k) = \underbrace{H_k P_{k,k} \mathrm{H}_k^T}_{\text{Effect of state uncertainty on measurement uncertainty}} + \underbrace{R_k}_{\text{Effect of measurement noise}}$