Recap

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 $\{a_i\}$ $i = 1, \ldots, n.$ Or, if the joint pdf of two CRVs x, y can be expressed as

$$p(x, y) = \frac{1}{2\pi\sigma_x\sigma_y[1-\rho^2]^{1/2}} e^{-\frac{1}{2(1-\rho^2)}\left[\frac{(x-\bar{x})^2}{\sigma_x^2} - \frac{2\rho(x-\bar{x})(y-\bar{y})}{\sigma_x\sigma_y} + \frac{(y-\bar{y})^2}{\sigma_y^2}\right]}$$

Where $\rho = \frac{E[(x-\bar{x})(y-\bar{y})]}{\sigma_x\sigma_y}$ = "correlation coefficient"

- X and Y are *uncorrelated* if $\rho = 0$, or equivalently, $E[XY] = E[X]E[Y]$ since $E[(x - \bar{x})(y - \bar{y})] = \mathrm{E}[\mathrm{XY}] - \mathrm{E}[\mathrm{X}]\mathrm{E}[\mathrm{Y}]$

Reality: vectors of Gaussian RVs are jointly Gaussian.

Proposition: (Appendix A of Anderson & Moore) Linear transformation and additions of Gaussian RVs are Gaussian RVs

Recap

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

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

Since x_0 is Gaussian distributed, and η_k is zero mean white Gaussian,

- x_k is a jointly Gaussian RV for each k

Since x_k is Gaussian distributed, and ω_k is zero mean white Gaussian,

- y_k is a jointly Gaussian RV for each k

Conditional pdf of joint Gaussian variables is Gaussian:

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

Where

$$\bar{\mu}_{x|y} = \bar{x} + Q_{xy} Q_{yy}^{-1} (\vec{y} - \bar{y}) \qquad Q_{x|y} = Q_{xx} - Q_{xy} Q_{yy}^{-1} Q_{xy}^T$$

I.e, what do measurements $\vec{y}$ tell us about state $\vec{x}$?

Linear Discrete Time Systems

Means of state and measurement Gaussians:

- $\bar{x}_n = \quad \Phi_{n,0}\bar{x}_0 \,; \qquad \Rightarrow \qquad \bar{x}_{n+1} = A_n$
- $\bar{y}_k = E[H_k\vec{x}_k + \vec{\omega}_k] = H_k\bar{x}_k$

Covariance of state and measurement Gaussians:

- $P_{k+1,k+1} = A_k P_{k,k} A_k^T + G_k Q_k G_k^T;$
- $cov(y_k, y_k) = H_k P_{k,k} \mathrm{H}_k^T + R_k$

Next: Select a criteria for estimator design

- ***Minimum Covariance***

Minimum Variance Design

Estimator is a random function

- Takes measurements $y_1, y_2, \cdots, y_n$ as input, and produces a random variable, $\hat{X}_n$, with $\hat{x}_n$ as a specific estimate.
- The variance associated with the estimator is the “uncertainty” in the estimate. Minimum variance design is the “least uncertain”

Minimum Variance Design (Kalman 1960)

- Choose the state estimate, $\hat{x}_k$, according to

$$\min_{\hat{x}_k} E[(X_k - \hat{x}_k)^T (X_k - \hat{x}_k)] \qquad (*)$$

- Note, the cost function is a *scalar*, as opposed to nxn covariance

$$E[(X_k - \hat{x}_k)(X_k - \hat{x}_k)^T] \qquad (**)$$

- ***Lemma:*** *Let $\vec{x}$ be a vector CRV, and let $||\vec{x}|| = \sqrt{E[\vec{x}^T x]}$. Then*
  - $||A\vec{x}||^2 = E[\vec{x}^T A^T A\vec{x}] = trace\{E[A^T A\vec{x}\vec{x}^T]\}$
  - If A=I, $||\vec{x}||^2 = E[\vec{x}^T\vec{x}] = trace[E[\vec{x}\vec{x}^T]]$
- So, minimizing(*) minimizes trace of (**)

Minimum Variance Estimator

Theorem 3.1 (Anderson & Moore, p. 26)

- Let X, Y be two joint distributed (not necessarily Gaussian) vector Rvs. Let Y be the “measurement,” which takes value y.
- The minimum variance estimate is given by the *conditional mean* of X given Y.

$$\hat{x} = E[X|Y = y] = \int_{-\infty}^{\infty} x\, p(x|y) dx$$

- Proof: (Brute force—see Anderson & Moore p. 27)

Consequence: With jointly Gaussian state (x) and measurements (y),

- *Mean:* $E\begin{bmatrix}\vec{x}\\ \vec{y}\end{bmatrix} = E\begin{bmatrix}\bar{x}\\ \bar{y}\end{bmatrix}$ Variance: $\begin{bmatrix}\Sigma_{xx} & \Sigma_{xy}\\ \Sigma_{yx} & \Sigma_{yy}\end{bmatrix}$, $\Sigma_{xy} = \Sigma_{yx}^T$
- The minimum variance estimate of $\vec{x}$ given $\vec{y}$

$$\hat{x} = \bar{x} + \Sigma_{xy}\Sigma_{yy}^{-1}(\vec{y} - \bar{y}) \qquad \Sigma_{x|y} = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{xy}^T$$