Advanced Control Engineering

- Full State Feedback for MIMO
- Stabilizability
- State Estimation & Output Feedback
- Observability & Duality
- Exogenous Inputs, Integral Control
- Optimal Control (LQR/LQG)
- Robustness & Sensitivity
- Kalman Filtering
- Introduction to Discrete Time

Background - I

Remember that the objective of the LQR problem was to find a constant state-feedback control gain matrix $\mathbf{G}$ that minimized the quadratic cost function:

$$V(\mathbf{G}) = \int_0^\infty \mathbf{x}^T(\tau)\mathbf{Q}\mathbf{x}(\tau) + \mathbf{u}^T(\tau)\mathbf{R}\mathbf{u}(\tau)d\tau$$

This feedback control was considered "optimal", but still required a state estimator or dynamic observer.

Background - II

So, what is a Kalman Filter?

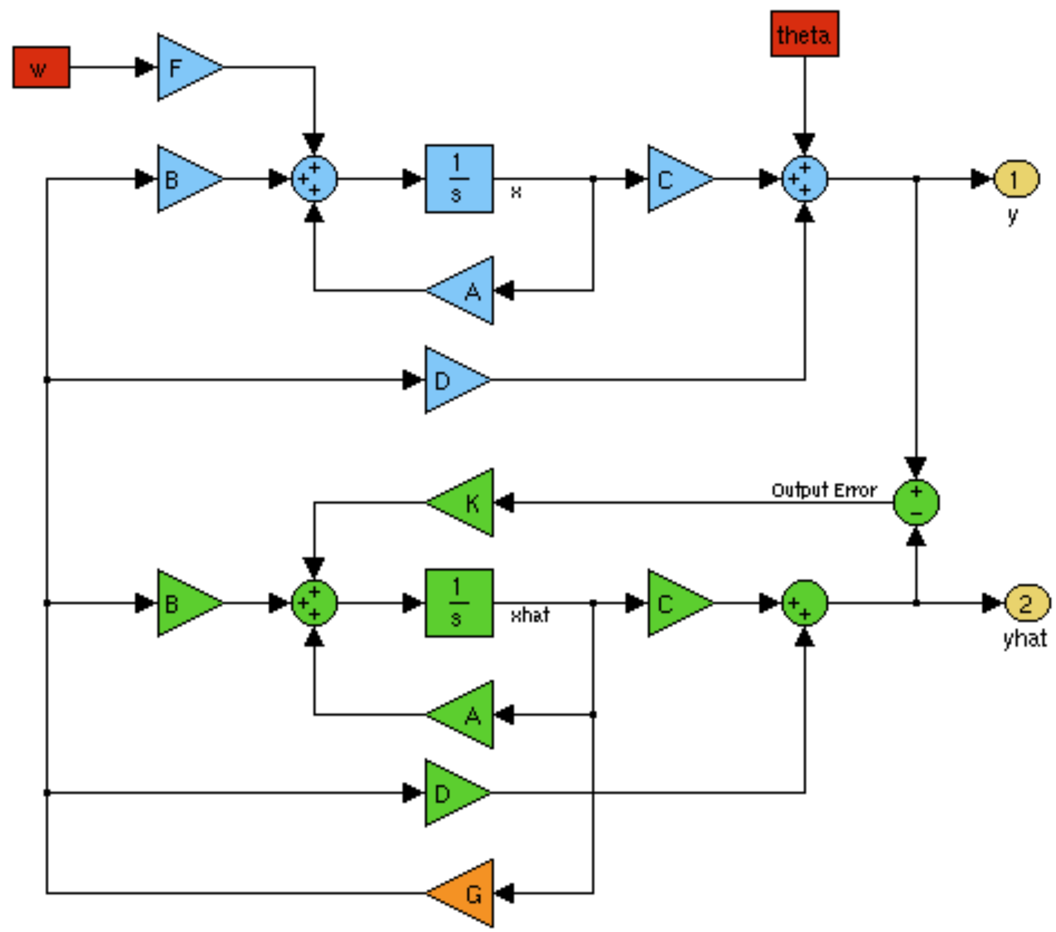
The Kalman Filter is an optimal linear state estimator.

This dynamic observer is the optimal solution to the Linear Quadratic Gaussian (LQG) problem of estimating the instantaneous states of a dynamic system when the process noise and measurement noise are white and Gaussian.

Background - III

The Kalman block diagram is exactly the same as for a Luenberger Observer:

A Luenberger Observer is sometimes considered to be a <u>sub-optimal</u> Kalman Filter



Background - IV

According to Grewal & Andrews, the Kalman filter:

- Was developed around 1959 by the Hungarian-born Rudolph Kalman
- Enabled trajectory estimation and control for the Apollo project (and others)
- Has been an integral part of nearly every onboard aircraft trajectory estimation and control system
- Is one of the greatest discoveries in the history of statistical estimation theory

Background - V

Kalman's idea was met with skepticism by the EE community so he published the original work in the *ASME Journal of Basic Engineering* in 1960.

The original Kalman filter was developed for discrete-time implementation.

In 1961, Kalman and Richard Bucy published the continuous-time version known as the <u>Kalman-Bucy Filter</u>.

Background - VI

Some interesting facts:

- The sub-optimal Luenberger observer came after the optimal Kalman filter
- The Kalman filter is optimal under ANY reasonable performance criterion if the random inputs are white and Gaussian
- A non-linear filter could be better if the random processes are not Gaussian
- The Kalman filter does not require that the random processes have stationary properties

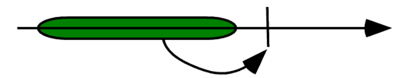
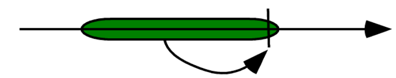
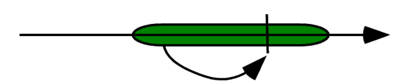
Background - VII

Although the Kalman Filter was developed in a community of control theorists, and is the most widely used result of Modern Control Theory, it is no longer viewed as a control result!

It is a result within the field of Estimation Theory and is more properly classified as a Signal Processing result.

Background - VIII

Within the field of Estimation Theory, there are three general types of estimators for the LQG problem:

- Predictors use observations strictly prior to the time of the estimated state 
- Filters use observations up to and including the time of the estimated state 
- Smoothers use observations beyond the time of the estimated state 

Background - IX

You (should) already have experience with optimal estimation.

In 1795, Gauss introduced the method of “least-squares” for extracting optimal estimates from noisy data.

The same principles are used to develop the optimal Kalman filter, but first we need to become familiar with <u>random variables</u> and <u>stochastic systems</u>.

Stochastics - I

The word *Stochastic* is derived from a Greek word meaning *to aim or guess*.

Knowledge of the state of a random variable at any instant is NOT sufficient to predict the future evolution.

This discussion is restricted to stationary random processes.

Let's start with a scalar continuous random variable $x(t)$.

Stochastics - II

A <u>cumulative distribution function</u> (CDF) completely describes the probability distribution of a real random variable x.

For every real number x_1, the CDF is given by:

$$P(x_1) = \Pr\left[x(t) \leq x_1\right]$$

The CDF is monotone increasing and continuous from the right.

Stochastics - III

A <u>probability density function</u> (PDF) $p(x)$ for a random signal $x(t)$ is related to the CDF by the following identity:

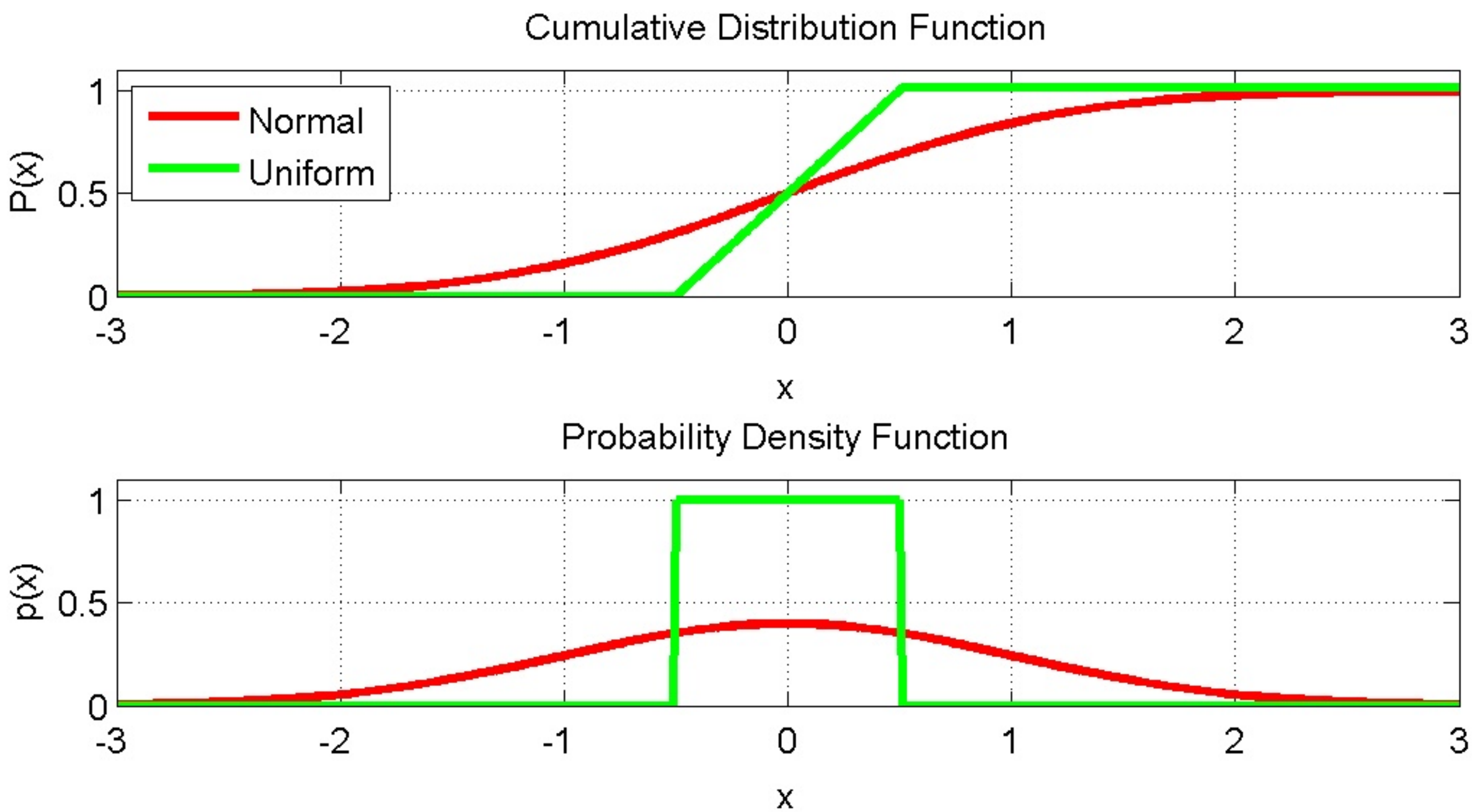
$$p(x_1) = \frac{d}{dx_1} P(x_1)$$

The probability density function is used to describe the probability of finding x with a specific value at a specific time.

$$\Pr\left[x_1 < x(t) < x_1 + \Delta x\right] = \int_{x_1}^{x_1 + \Delta x} p(x_1)\,dx \approx p(x_1)\Delta x$$

Stochastics - IV

Let's look at two common examples:



$$p_{uniform}(x) = \begin{cases} 1 & |x| \leq 0.5 \\ 0 & |x| > 0.5 \end{cases}$$

$$p_{normal}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\bar{x})^2}{2\sigma^2}\right)$$

Stochastics - V

The most common statisical properties are first order:

Mean $\bar{x} = E\langle x \rangle = \int_{-\infty}^{\infty} x p(x) dx$

Mean Square $E\langle x^2 \rangle = \int_{-\infty}^{\infty} x^2 p(x) dx$

Variance $\sigma^2 = E\langle (x - \bar{x})^2 \rangle = \int_{-\infty}^{\infty} (x - \bar{x})^2 p(x) dx$

Where E is the <u>expectation operator</u>, and means “average”.

Stochastics - VI

Second-order statistics are used to evaluate correlation functions.

Auto-correlation:

$$\rho_{xx}(\tau) = E\langle x(t)x(t+\tau)\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 \underbrace{p\big(x_1(t), x_2(t+\tau)\big)}_{\text{Joint Density Function}} dx_1 dx_2$$

Cross-correlation:

$$\rho_{xy}(\tau) = E\langle x(t)y(t+\tau)\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, \underbrace{p\big(x(t), y(t+\tau)\big)}_{\text{Joint Density Function}} dx\, dy$$

Stochastics - VII

Second-order statistics are also used to evaluate covariance functions.

Auto-covariance:

$$C_{xx}(\tau) = E\left\langle \left(x(t) - \bar{x}\right)\left(x(t+\tau) - \bar{x}\right) \right\rangle$$

Covariance:

$$C_{xy}(\tau) = E\left\langle \left(x(t) - \bar{x}\right)\left(y(t+\tau) - \bar{y}\right) \right\rangle$$

Stochastics - VIII

The covariance is related to the cross-correlation because the expectation is a linear operator :

$$C_{xy}(\tau) = E\left\langle \left(x(t) - \bar{x}\right)\left(y(t+\tau) - \bar{y}\right)\right\rangle$$

$$= E\left\langle x(t)y(t+\tau) - \bar{x}y(t+\tau) - \bar{y}x(t) + \bar{x}\bar{y}\right\rangle$$

$$= E\left\langle x(t)y(t+\tau)\right\rangle - E\left\langle \bar{x}y(t+\tau)\right\rangle - E\left\langle \bar{y}x(t)\right\rangle + E\left\langle \bar{x}\bar{y}\right\rangle$$

$$= \rho_{xy}(\tau) - \bar{x}E\left\langle y(t)\right\rangle - \bar{y}E\left\langle x(t)\right\rangle + \bar{x}\bar{y}$$

$$= \rho_{xy}(\tau) - \bar{x}\bar{y} - \bar{y}\bar{x} + \bar{x}\bar{y}$$

$$= \rho_{xy}(\tau) - \bar{x}\bar{y}$$

Stochastics - IX

Two scalar random variables x and y are called <u>uncorrelated</u> if:

$$E\langle xy\rangle = E\langle x\rangle\, E\langle y\rangle \quad \Leftrightarrow \quad \rho_{xy}(\tau) = \bar{x}\;\bar{y}$$

They are called <u>orthogonal</u> if:

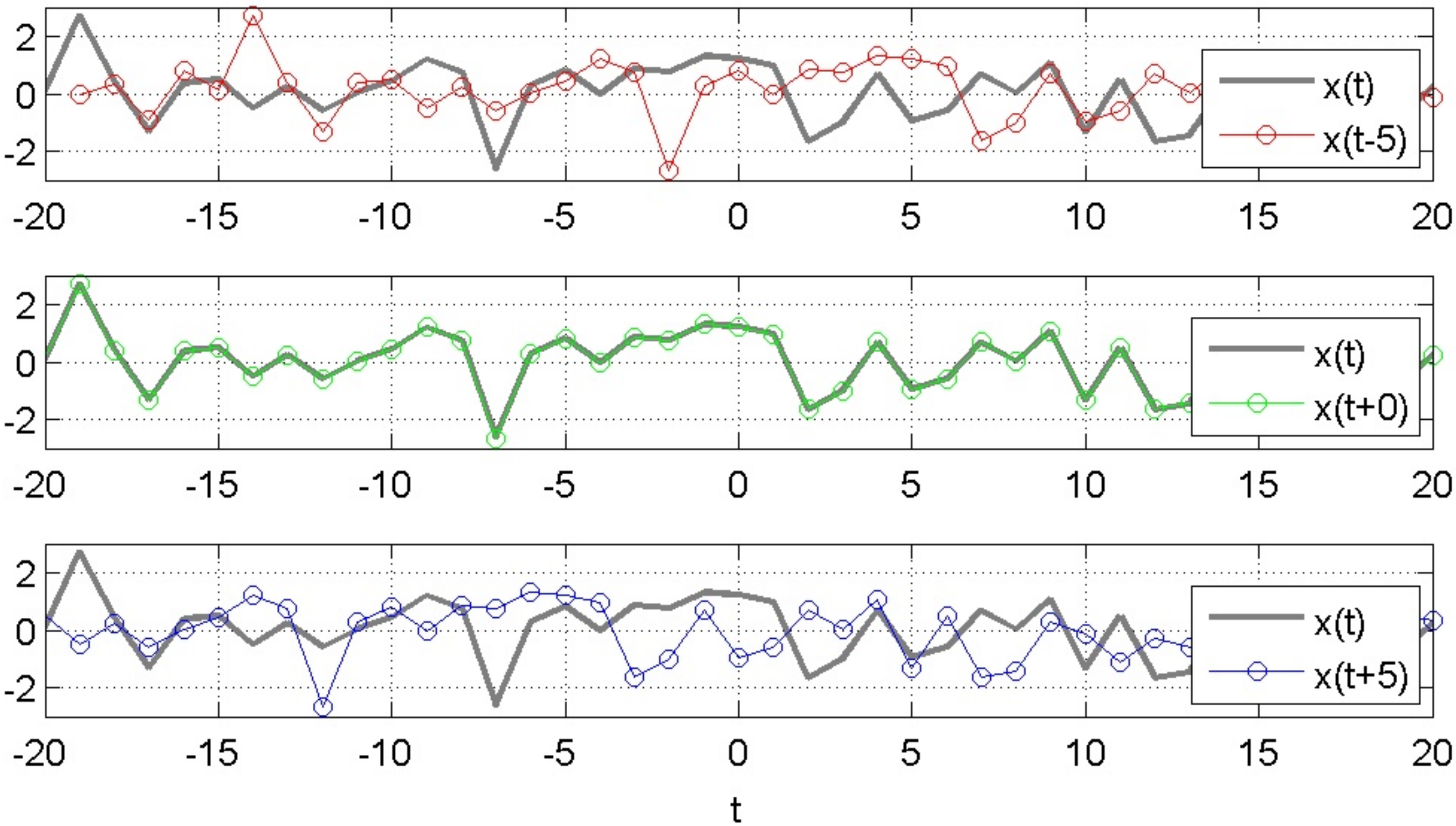
$$E\langle xy\rangle = 0 \quad \Leftrightarrow \quad \rho_{xy}(\tau) = 0$$

and <u>independent</u> if:

$$p(x,y) = p(x)\, p(y)$$

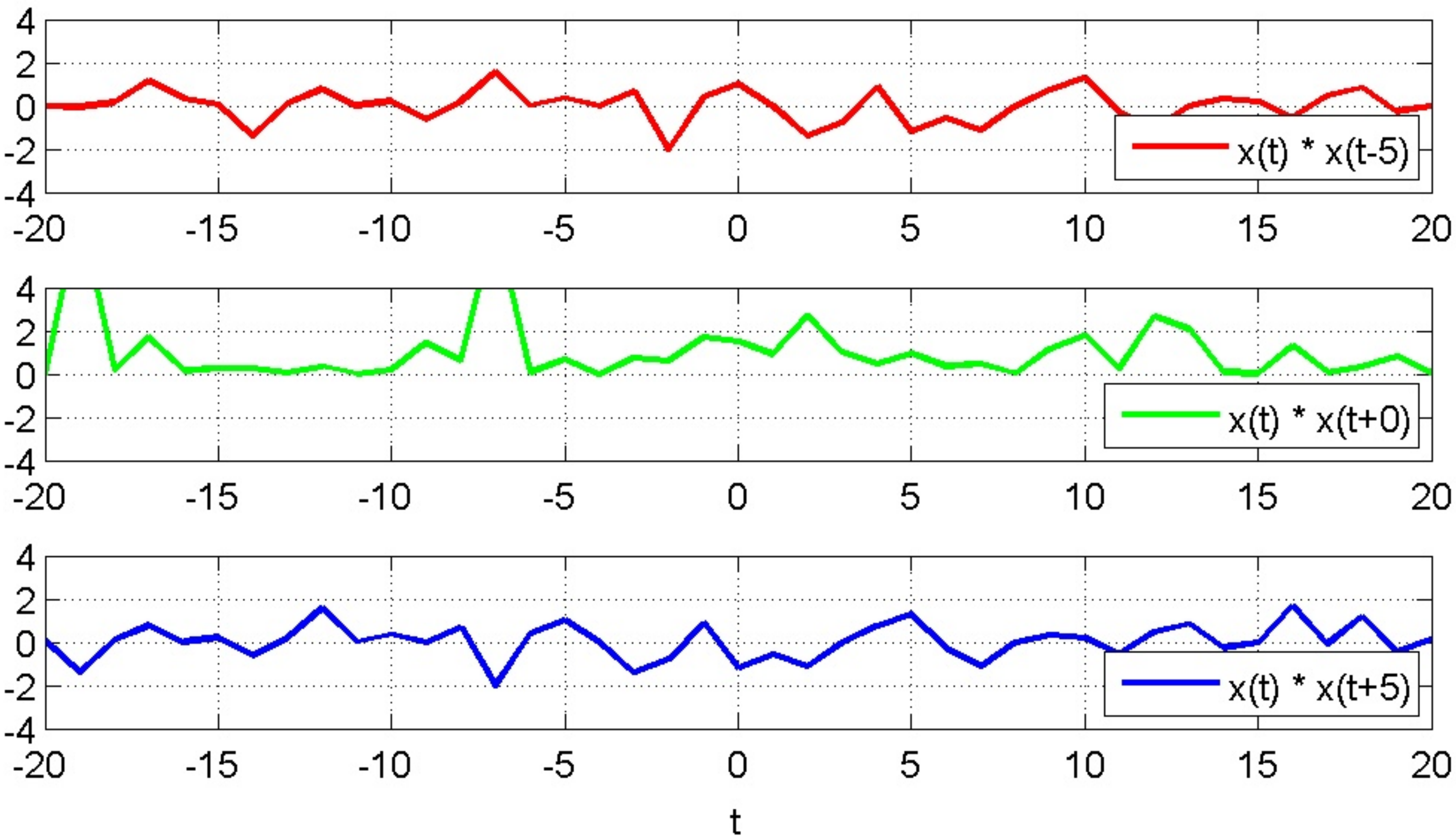
Stochastics Example - I

Let's look at the auto-correlation function a little closer: $\rho_{xx}(\tau) = E\langle x(t)\, x(t+\tau)\rangle$



Stochastics Example - II

The product of $x(t)$ and the delayed version of the same signal $x(t+\tau)$ is:



Stochastics Example - III

The auto-correlation $\rho_{xx}(\tau)$ is the average or mean value over time of this product signal but evaluated over a wide range of positive and negative delays τ.

What is the average of $x(t) \times x(t-5)$?

What is the average of $x(t) \times x(t-0)$?

What is the average of $x(t) \times x(t+5)$?

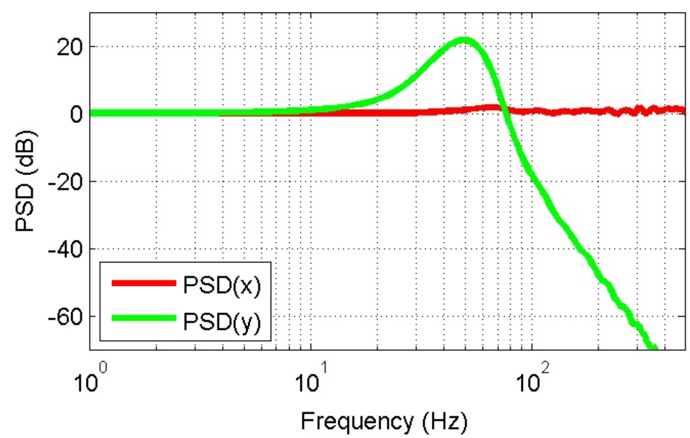
Stochastics Example - IV

Let's input a zero-mean, unity variance, normally distributed (i.e. Gaussian) noise signal $x(t)$ into a 2nd order SISO system and evaluate the normalized auto- and cross-correlations.

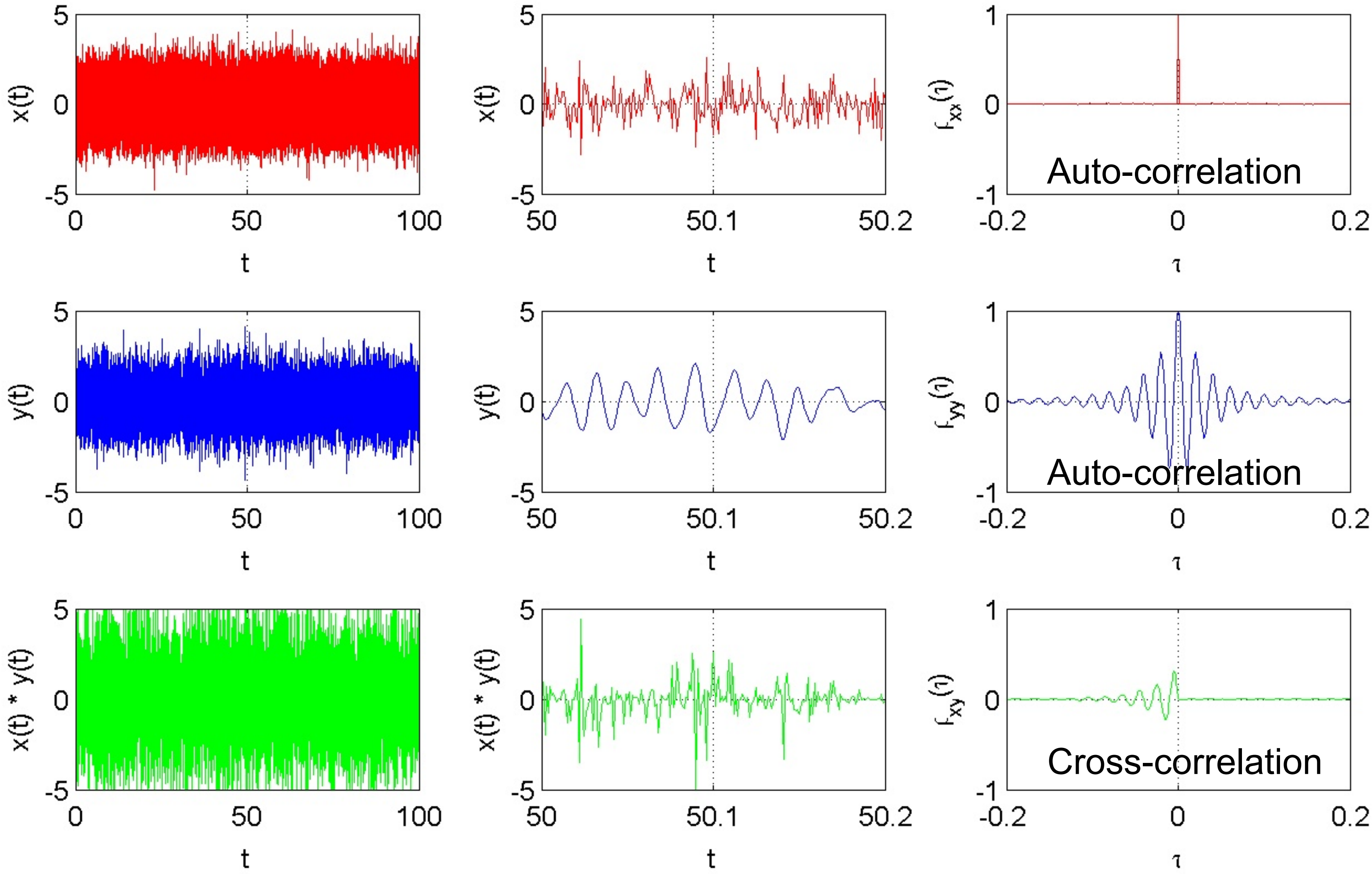
$$x(t) \longrightarrow \boxed{\frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}} \longrightarrow y(t)$$

$$\bar{x} = 0 \qquad \omega = 2\pi \times 50$$

$$\sigma_x^2 = 1 \qquad \zeta = 0.1$$



Stochastics Example - V



Stochastics Example - VI

Conclusions:

- The auto-correlation of the input signal is zero everywhere except at $\tau = 0$
- The auto-correlation of the output signal is NOT zero near $\tau = 0$, but it approaches zero for delays larger than a certain value
- The output auto-correlation is characteristic of a lightly damped band-limited process
- The cross-correlation between the input and output is zero for all positive delays

Vector Stochastics - I

The expectation operator is linear therefore expectations of vectors and matrices are evaluated element-wise.

Let $\mathbf{x}(t)$ be an N-dimensional random vector process, then the mean is:

$$\bar{\mathbf{x}} = E\left\langle \mathbf{x}(t) \right\rangle = \begin{bmatrix} E\left\langle x_1(t) \right\rangle \\ \vdots \\ E\left\langle x_N(t) \right\rangle \end{bmatrix} = \begin{bmatrix} \bar{x}_1 \\ \vdots \\ \bar{x}_N \end{bmatrix}$$

Vector Stochastics - II

A correlation matrix associated with the random vector $\mathbf{x}(t)$ is defined by the following matrix of cross-correlations:

$$\underset{[N\times N]}{\mathbf{R}(\tau)} = E\left\langle \underset{[N\times 1]}{\mathbf{x}(t)}\,\underset{[1\times N]}{\mathbf{x}^T(t+\tau)} \right\rangle$$

$$= \begin{bmatrix} E\langle x_1(t)x_1(t+\tau)\rangle & \cdots & E\langle x_1(t)x_N(t+\tau)\rangle \\ \vdots & \ddots & \vdots \\ E\langle x_N(t)x_1(t+\tau)\rangle & \cdots & E\langle x_N(t)x_N(t+\tau)\rangle \end{bmatrix}$$

Vector Stochastics - III

Similarly, a <u>covariance matrix</u> associated with the random vector $\mathbf{x}(t)$ is defined element-wise by the following matrix of covariances:

$$\underset{[N\times N]}{\mathbf{C}(\tau)} = E\left\langle \left(\mathbf{x}(t)-\overline{\mathbf{x}}\right)\left(\mathbf{x}(t+\tau)-\overline{\mathbf{x}}\right)^T \right\rangle$$

$$= \mathbf{R}(\tau) - \overline{\mathbf{x}}\,\overline{\mathbf{x}}^T$$

Notice that when the vector process $\mathbf{x}(t)$ is zero-mean, the the correlation and covariance matrices are equal!

Vector Stochastics - IV

Also notice that all the elements on the main diagonal of $\mathbf{R}(\tau)$ are auto-correlations and all off-diagonal elements are cross-correlations

$$\mathbf{R}(\tau) = \begin{bmatrix} \rho_{11}(\tau) & \cdots & \rho_{1N}(\tau) \\ \vdots & \ddots & \vdots \\ \rho_{N1}(\tau) & \cdots & \rho_{NN}(\tau) \end{bmatrix}$$

where

$$\rho_{ij}(\tau) = E\left\langle x_i(t)\, x_j(t+\tau) \right\rangle$$

Vector Stochastics - V

If all off-diagonal elements of $\mathbf{R}(\tau)$ are zero, then the components of $\mathbf{x}(t)$ are said to be <u>uncorrelated</u> (but not necessarily independent).

$$\mathbf{R}(\tau) = \begin{bmatrix} \rho_{11}(\tau) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \rho_{NN}(\tau) \end{bmatrix}$$

Vector Stochastics - VI

For the case where the components of $\mathbf{x}(t)$ are zero-mean uncorrelated Gaussian random variables, the auto-correlation and the covariance matrices are the same and simplify to:

$$\mathbf{C}(\tau) = \mathbf{R}(\tau) = \begin{bmatrix} \sigma_{11}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{NN}^2 \end{bmatrix} \delta(\tau) = \boldsymbol{Q}\,\delta(\tau)$$

Summary

- The Kalman filter has the same structure as a Luenberger observer, but with “optimally” chosen observer gains
- Scalar statistical properties such as mean, correlation, and covariance naturally extend to vector stochastic processes