Lecture 10: Stochastic Dynamic System Models for State Estimation

Thus far, we have focused our attention on *deterministic systems* wherein, given the same initial condition $\boldsymbol{x}(t_0)$ and input history (e.g., $\boldsymbol{u}(t)$ in the continuous-time case), the system always follows the same trajectory $\boldsymbol{x}(t)$ and produces the same output $\boldsymbol{y}(t)$ (i.e., there is no randomness). On the other hand, *stochastic systems* explicitly consider uncertainty or randomness. Stochastic differential equations (SDEs) arise in numerous disciplines ranging from quantum physics to finance. A rigorous treatment of SDEs is outside the scope of this course. Instead, we focus on the main forms of stochastic systems that appear in state estimation — that is, we augment our familiar LTI, LTV, and nonlinear system models with additional random inputs called *process noise* and *measurement noise*.

Process Noise. Process noise is used to model uncertainty in the time-evolution of the system (i.e., in the state-rate $\dot{\boldsymbol{x}}$). Process noise can account for model uncertainty that can be due to a lack of precise knowledge of the system parameters, unmodeled effects, or slow changes in the system configuration over time. Process noise can also account for external perturbations (sometimes called *exogenous disturbances*) experienced by the system — for example, an aircraft may experience wind gusts in flight. In essence, we admit that our system model does not perfectly capture the dynamics of the "real" system and we compensate for this fact by using process noise to account for deviations from our idealized model.

Measurement Noise. While process noise concerns the underlying system dynamics, measurement noise models uncertainty or randomness in the outputs of the system (i.e., $\boldsymbol{y}(t)$). Intuitively, even if we have a perfect system model we can still receive erroneous or noisy outputs due to imperfect sensors. All real sensors have limitations such as delays, bias, or noise which can be attributed to random variations of the environment, limited resolution or sensitivity, manufacturing imperfections, etc. The measurement noise term provides one approach to model and account for these effects.

Gaussian and Uncorrelated Noise Assumption. Both the process and measurement noise terms are treated as random variables characterized by a probability distribution—most commonly a zero-mean, Gaussian, and uncorrelated in time (i.e., *white noise* or *independent and identically distributed (iid)*). There are systems that require correlated noise models (i.e., *colored noise*) to account for noise with given frequency content (over time) and that have various (non-Gaussian) distributions.

Continuous-Time Nonlinear Stochastic System. A continuous-time nonlinear system with process and measurement noise can be written as:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), \boldsymbol{w}(t)) \tag{1}$$

$$\boldsymbol{y}(t) = \boldsymbol{h}(\boldsymbol{x}(t), \boldsymbol{v}(t)) \tag{2}$$

$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{3}$$

where as usual, $\boldsymbol{x}(t) \in \mathbb{R}^n$, $\boldsymbol{u}(t) \in \mathbb{R}^m$, $\boldsymbol{y}(t) \in \mathbb{R}^p$, and we've now introduced $\boldsymbol{w}(t) \in \mathbb{R}^w$ as the process noise and $\boldsymbol{v}(t) \in \mathbb{R}^p$ as the measurement noise. Note that the dimension of process noise vectors is not necessarily the same size as the state (i.e., $w \neq n$). For example, we may wish to model the noise as affecting only a subset of the state-rates and then propagating through the dynamics to other states. The measurement noise is commonly assumed to affect the noise of all the output channels. The elements of the covariance matrix describe the relative magnitude of the noise on each channel. For continuous systems process noise that is zero-mean, Gaussian, and white is denoted $\boldsymbol{w}(t) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{Q}_c)$ which is shorthand for

$$\begin{aligned} E[\boldsymbol{w}(t)] &= 0 \\ E[\boldsymbol{w}(t)\boldsymbol{w}^{\mathrm{T}}(\tau)] &= \boldsymbol{Q}_c \delta(t-\tau) \ , \end{aligned}$$

where $\delta(t-\tau)$ is the Dirac delta function and $\boldsymbol{Q}_c \in \mathbb{R}^{w \times w}$ is the continuous-time *process noise covariance* matrix. Similarly, measurement noise that is zero-mean, Gaussian, and white is denoted is denoted $\boldsymbol{v}(t) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{R}_c)$ which is shorthand for

$$\begin{aligned} E[\boldsymbol{v}(t)] &= 0 \\ E[\boldsymbol{v}(t)\boldsymbol{v}^{\mathrm{T}}(\tau)] &= \boldsymbol{R}_c \delta(t-\tau) \ , \end{aligned}$$

where $\boldsymbol{R}_c \in \mathbb{R}^{p \times p}$ is the continuous-time *measurement noise covariance* matrix. The use of the delta function is a consequence of the uncorrelated noise assumption. The covariances $E[\boldsymbol{w}(t)\boldsymbol{w}^{\mathrm{T}}(\tau)]$ and $E[\boldsymbol{v}(t)\boldsymbol{v}^{\mathrm{T}}(\tau)]$ are both zero, respectively, whenever $t \neq \tau$ (i.e., at two different times). When $t = \tau$ the random vectors have covariance $\boldsymbol{Q}_c$ and $\boldsymbol{R}_c$, respectively. If we considered an experiment that consisted of observing system (1)–(3) evolve from the same initial condition and with the same control input history we would find that each experiment had a different outcome due to the presence of random process and measurement noise.

Aside (Dirac delta): The Dirac delta function is defined as

$$\delta(t-\alpha) = \begin{cases} \infty & \text{if } t = \alpha \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

in other words, it has a value of ∞ at $t = \alpha$ and is zero everywhere else. Integrating the Dirac delta function gives unit area, i.e.,

$$\int_{-\infty}^{\infty} \delta(t-\alpha) dt = 1 \ . \tag{5}$$

Later in this lecture we will also use the *sifting property* of the Dirac delta function. That is, for some smooth function $f(t)$

$$\int_{-\infty}^{\infty} f(t)\delta(t-\alpha) dt = f(\alpha) \ . \tag{6}$$

The sifting property can be proven as follows. Since $\delta(t)$ is zero everywhere except at $t = \alpha$ we can change the integration bounds to a small interval around α, for example, $[\alpha - \epsilon, \alpha + \epsilon]$ where ϵ is a small value. On this small interval the function $f(t)$ is approximately constant and equal to $f(\alpha)$. Treating $f(t) = f(\alpha)$ as a constant it follows that:

$$\int_{\alpha-\epsilon}^{\alpha+\epsilon} f(\alpha)\delta(t-\alpha) dt = f(\alpha) \int_{\alpha-\epsilon}^{\alpha+\epsilon} \delta(t-\alpha) dt = f(\alpha) \ . \tag{7}$$

Where the last step follows from the fact that the Dirac delta function integrates to one around any interval containing α.

Continuous-Time Linear Stochastic Systems with Additive Noise. For the nonlinear stochastic system (1)–(3) the noise affect the system in an arbitrary (i.e., nonlinear) way. However, for linear systems it is common to assume that the process and measurement noise are zero-mean, Gaussian, uncorrelated and also *additive*. That is, they enter the system linearly as additive random vectors in the state-rate and measurement equations.

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}(t)\boldsymbol{x}(t) + \boldsymbol{B}(t)\boldsymbol{u}(t) + \underbrace{\boldsymbol{L}(t)\boldsymbol{w}(t)}_{\text{process noise}} \tag{8}$$

$$\boldsymbol{y}(t) = \boldsymbol{C}(t)\boldsymbol{x}(t) + \underbrace{\boldsymbol{v}(t)}_{\text{msmt noise}} \tag{9}$$

$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{10}$$

where, as before, $\boldsymbol{w}(t) \in \mathbb{R}^w$ is the process noise and $\boldsymbol{v}(t) \in \mathbb{R}^p$ is the measurement noise, $\boldsymbol{x}(t) \in \mathbb{R}^n$, $\boldsymbol{u}(t) \in \mathbb{R}^m$, $\boldsymbol{y}(t) \in \mathbb{R}^p$, $\boldsymbol{A}(t) \in \mathbb{R}^{n\times n}$, $\boldsymbol{B}(t) \in \mathbb{R}^{n\times m}$, $\boldsymbol{C}(t) \in \mathbb{R}^{m\times n}$. The matrix $\boldsymbol{L}(t) \in \mathbb{R}^{n\times w}$ is the process noise influence matrix that describes how noise enters the system.

Discrete-Time Linear Stochastic Systems with Additive Noise. For discrete-time linear systems process and measurement noise is described in an analogous way to (8)–(10). That is,

$$\boldsymbol{x}_k = \boldsymbol{F}_{k-1}\boldsymbol{x}_{k-1} + \boldsymbol{G}_{k-1}\boldsymbol{u}_{k-1} + \boldsymbol{L}_{k-1}\boldsymbol{w}_{k-1} \tag{11}$$

$$\boldsymbol{y}_k = \boldsymbol{H}_k\boldsymbol{x}_k + \boldsymbol{v}_k \tag{12}$$

$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0 \tag{13}$$

where, as before, $\boldsymbol{w}_k \in \mathbb{R}^w$ is the process noise and $\boldsymbol{v}_k \in \mathbb{R}^v$ is the measurement noise and $\boldsymbol{F}_k \in \mathbb{R}^{n\times n}$, $\boldsymbol{G}_k \in \mathbb{R}^{n\times m}$, $\boldsymbol{H}_k \in \mathbb{R}^{m\times n}$. The matrix $\boldsymbol{L}_k \in \mathbb{R}^{n\times w}$ is, again, the process noise influence matrix. For discrete-time systems process noise that is zero-mean, Gaussian, and white is denoted $\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q})$ which is shorthand for

$$E[\boldsymbol{w}_k] = 0 \tag{14}$$

$$E[\boldsymbol{w}_k\boldsymbol{w}_k^{\mathrm{T}}] = \boldsymbol{Q} \tag{15}$$

$$E[\boldsymbol{w}_k\boldsymbol{w}_q^{\mathrm{T}}] = \boldsymbol{0} \qquad \forall q \neq k\,, \tag{16}$$

where $\boldsymbol{Q} \in \mathbb{R}^{w\times w}$ is the discrete-time process noise covariance matrix. Similarly, measurement noise that is zero-mean, Gaussian, and white is denoted is denoted $\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R})$ which is shorthand for

$$E[\boldsymbol{v}_k] = 0 \tag{17}$$

$$E[\boldsymbol{v}_k\boldsymbol{v}_k^{\mathrm{T}}] = \boldsymbol{R} \tag{18}$$

$$E[\boldsymbol{v}_k\boldsymbol{v}_q^{\mathrm{T}}] = \boldsymbol{0} \qquad \forall q \neq k\,, \tag{19}$$

where $\boldsymbol{R} \in \mathbb{R}^{p\times p}$ is the discrete-time measurement noise covariance matrix. Notice that the continuous-time Dirac delta function is not necessary to define the noise covariance matrices in the discrete-time case. Also, the continuous-time and discrete-time covariances $\boldsymbol{Q}_c$ and $\boldsymbol{Q}$ or $\boldsymbol{R}_c$ and $\boldsymbol{R}$ are not necessarily equal. In the following we will describe how these two matrices are related when the discrete-system is a discretized version of the continuous one.

Relating Discrete and Continuous-Time Additive Noise [Simon, 2006, Sec. 8.1]. Our goal in this section is establish a relationship between the description of process and measurement noise for both discrete and continuous-time systems. Consider the following discrete-time system that has a sample time Δt.

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} + \boldsymbol{w}_{k-1} \tag{20}$$
$$\boldsymbol{w}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}) \tag{21}$$
$$\boldsymbol{x}_0 = \boldsymbol{0} \tag{22}$$

where $\boldsymbol{w}_k$ is a discrete time-white noise process. At an arbitrary time step k the state is found to be:

$$\boldsymbol{x}_k = \boldsymbol{w}_0 + \boldsymbol{w}_1 + \boldsymbol{w}_2 + \cdots + \boldsymbol{w}_{k-1} \tag{23}$$

and the covariance of the state is therefore

$$\begin{aligned} E[\boldsymbol{x}_k\boldsymbol{x}_k^{\mathrm{T}}] &= E[(\boldsymbol{w}_0 + \boldsymbol{w}_1 + \cdots + \boldsymbol{w}_{k-1})(\boldsymbol{w}_0 + \boldsymbol{w}_1 + \cdots + \boldsymbol{w}_{k-1})^{\mathrm{T}}] && (24)\\ &= E[\boldsymbol{w}_0\boldsymbol{w}_0^{\mathrm{T}}] + E[\boldsymbol{w}_0\boldsymbol{w}_1^{\mathrm{T}}] + \cdots + E[\boldsymbol{w}_0\boldsymbol{w}_{k-1}^{\mathrm{T}}] + \cdots + E[\boldsymbol{w}_{k-1}\boldsymbol{w}_{k-1}^{\mathrm{T}}] && (25)\end{aligned}$$

Recall from (15)–(16) that at any arbitrary time q the white process noise is uncorrelated with the noise at any other time not equal to q. Thus, all of the terms of the form $E[\boldsymbol{w}_q\boldsymbol{w}_r] = 0$ are zero above with $r \neq q$ and the expression becomes

$$\begin{aligned} E[\boldsymbol{x}_k\boldsymbol{x}_k^{\mathrm{T}}] &= E[\boldsymbol{w}_0\boldsymbol{w}_0^{\mathrm{T}}] + E[\boldsymbol{w}_1\boldsymbol{w}_1^{\mathrm{T}}] + \cdots + E[\boldsymbol{w}_{k-1}\boldsymbol{w}_{k-1}^{\mathrm{T}}] && (26)\\ &= \boldsymbol{Q} + \boldsymbol{Q} + \cdots \boldsymbol{Q} && (27)\\ &= k\boldsymbol{Q}\ . && (28)\end{aligned}$$

We've found that the covariance of the state increases linearly with time step k for a given sample time Δt. Now, consider a the continuous-time version of this system

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{w}(t) \tag{29}$$

with zero initial condition $\boldsymbol{x}(t_0) = 0$ at time $t_0 = 0$ and where $\boldsymbol{w}(t)$ is a continuous-time white noise with covariance

$$E[\boldsymbol{w}(t)\boldsymbol{w}^{\mathrm{T}}(\tau)] = \boldsymbol{Q}_c\delta(t-\tau)\ . \tag{30}$$

The system solution is given by

$$\boldsymbol{x}(t) = \int_0^t \boldsymbol{w}(\alpha)d\alpha\ , \tag{31}$$

and the covariance of $\boldsymbol{x}(t)$ is

$$\begin{aligned} E[\boldsymbol{x}(t)\boldsymbol{x}(t)^{\mathrm{T}}] &= E\left[\int_0^t \boldsymbol{w}(\alpha)d\alpha \int_0^t \boldsymbol{w}^{\mathrm{T}}(\beta)d\beta\right] && (32)\\ &= \int_0^t\int_0^t E\left[\boldsymbol{w}(\alpha)\boldsymbol{w}^{\mathrm{T}}(\beta)\right]d\alpha d\beta\ . && (33)\end{aligned}$$

Then, substituting (30)

$$\begin{aligned} E[\boldsymbol{x}(t)\boldsymbol{x}(t)^{\mathrm{T}}] &= \int_0^t\int_0^t \boldsymbol{Q}_c\delta(\alpha-\beta)d\alpha d\beta && (34)\\ &= \boldsymbol{Q}_c\int_0^t\left[\int_0^t \delta(\alpha-\beta)d\alpha\right]d\beta && (35)\end{aligned}$$

For each fixed value of $\beta \in [0, t]$ the inner integral is the dirac delta function multiplied by a smooth (unit-valued function). By the sifting property it follows that for each fixed value of β the inner integral is equal to one. Thus,

$$E[\boldsymbol{x}(t)\boldsymbol{x}(t)^{\mathrm{T}}] = \boldsymbol{Q}_c \int_0^t 1 d\beta \tag{36}$$

$$= \boldsymbol{Q}_c t \; . \tag{37}$$

Now, compare (37) with (28). If we assume the discrete system had a fixed sampling time of Δt then the kth step corresponds to a continuous time of $t = k\Delta t$. For the covariance to grow at the same rate in both cases we require, from (37) with (28),

$$k\boldsymbol{Q} = t\boldsymbol{Q}_c \tag{38}$$

$$k\boldsymbol{Q} = k\Delta t \boldsymbol{Q}_c \tag{39}$$

$$\implies \boldsymbol{Q} = \Delta t \boldsymbol{Q}_c \tag{40}$$

It follows that the discrete-time white noise with covariance $\boldsymbol{Q}$ in a system with a sample period of Δt is equivalent to continuous-time white noise with covariance $\boldsymbol{Q}_c \delta(t)$ where $\boldsymbol{Q}_c = \boldsymbol{Q}/\Delta t$. By a related argument, for the system,

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} \tag{41}$$

$$\boldsymbol{y}_k = \boldsymbol{x}_k + \boldsymbol{v}_k \tag{42}$$

$$\boldsymbol{v}_k \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}) \tag{43}$$

$$\boldsymbol{x}_0 = \boldsymbol{0} \tag{44}$$

it can be shown (see [Simon, 2006, 8.1.2]) that a similar relationship holds between the continuous and discrete-time measurement covariance. That is, choosing

$$\boldsymbol{R} = \boldsymbol{R}_c / \Delta t \tag{45}$$

produces the same effect.

Simulating Stochastic Discrete-Time Systems. A stochastic discrete-time system, such as (11)–(13), can be simulated by using a random number/vector generator at each time step to produce the desired process and/or measurement noise. In MATLAB the multi-variate norman random number generator can be called using `w = mvnrnd(mu,Sigma)` where `mu` is the mean vector and `Sigma` is the covariance matrix.

Example. The code below illustrates simulating the linear system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & -0.5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \tag{46}$$

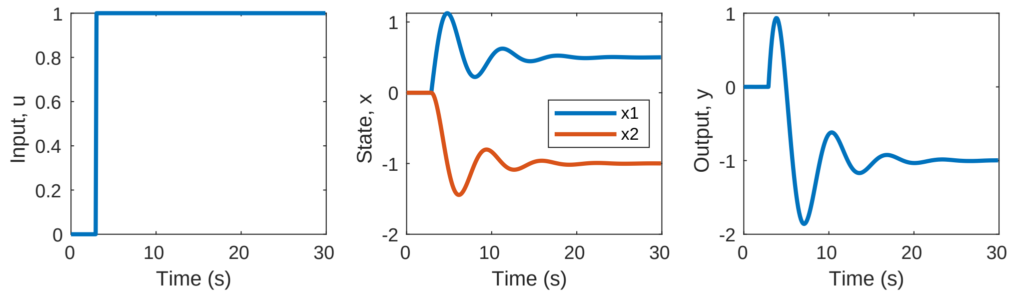
$$y = \begin{bmatrix} 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + v \tag{47}$$

subject to a step input $u(t) = H(t - 0.3)$ from $x(t_0) = [0, 0]^{\mathrm{T}}$ (where $H(\cdot)$ is the Heaviside function) at time $t_0 = 0$ to time $t = 30$ sec. with data sampled and control changes applied at time-intervals of $\Delta t = 0.1$ seconds where $\boldsymbol{Q}_c = 0.005\boldsymbol{I}_{2\times 2}$ and $\boldsymbol{R}_c = 0.001$.

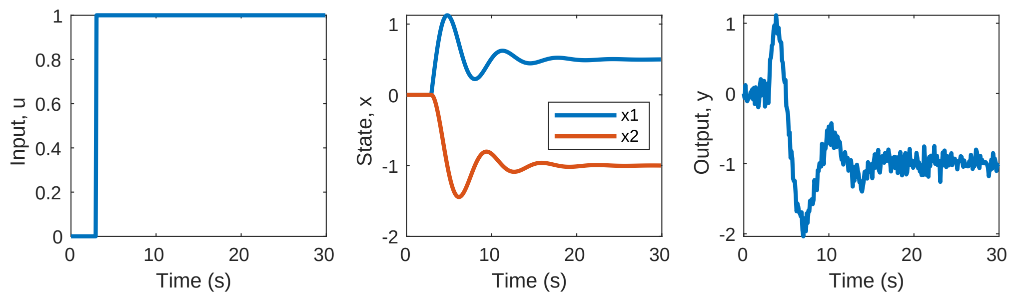
```
rng(3); % fix random number generator seed
A = [0 1;-1 -0.5]; % define LTI system matrices, (A,B,C,D)
B = [1; 0];
C = [2 2];
D = 0;
x0 = [0; 0]; % initial condition
sys = ss(A,B,C,D); % create matlab system object

% continuous-time noise properties
mu_w = zeros(2,1);
Qc = 0.005*[1 0;0 1];
mu_v = 0;
Rc = 0.001;

h = 0.1; % time-step
sys2 = c2d(sys,h); % convert to discrete system
F = sys2.A;
G = sys2.B;
H = sys2.C;
T = 30;

% simulate
N = T/h;
thist(1) = 0;
xhist(:,1) = x0;
yhist(1) = H*x0;
uhist(1) = sin(0);
for k = 2:1:N
    % time/state at previous step
    xkm1 = xhist(:,k-1); % x_{k-1}
    tkm1 = thist(k-1); % t_{k-1}
    % step input
    if ( k > N/10)
        ukm1 = 1;
    else
        ukm1 = 0;
    end
    % Discrete-time dynamics
    wkm1 = mvnrnd(mu_w,Qc*h)';
    xhist(:,k) = F*xkm1 + G*ukm1 + wkm1;
    thist(k) = thist(k-1) + h;
    vk = mvnrnd(mu_v,Rc/h);
    yhist(k) = C*xhist(:,k) + vk;
```

```
    uhist(k) = ukm1;
end
```

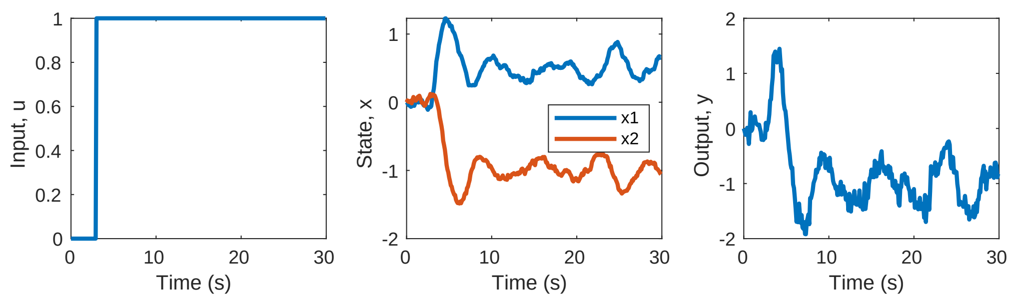
The solution with no measurement or process noise is

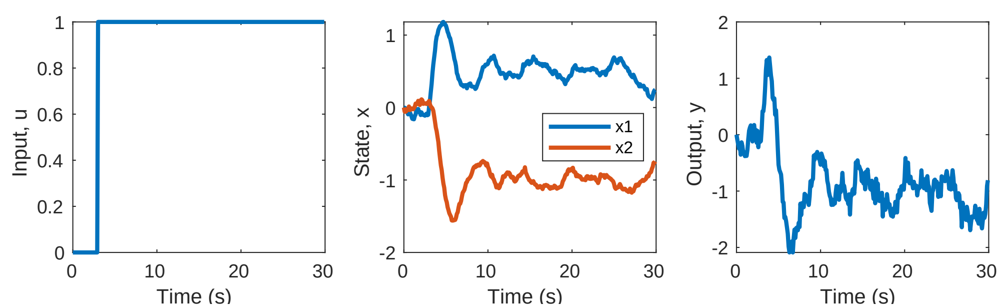


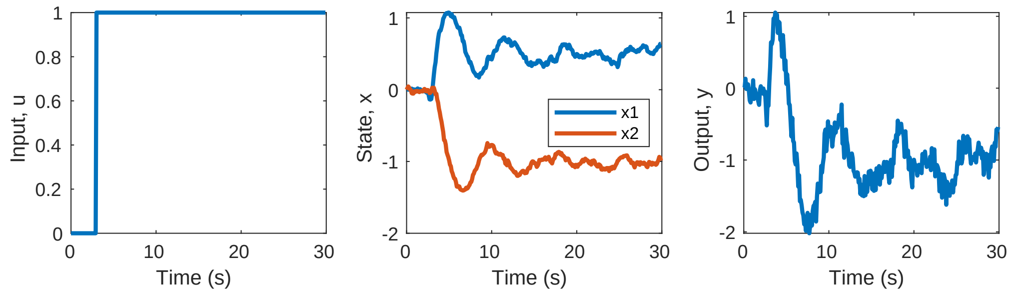
Now adding measurement noise



and simulating three trials with both process and measurement noise:







Notice that each trial above produced a slightly different result due to the stochastic nature of the simulation.

Final Remarks

In this lecture we described a specific class of stochastic differential equation (SDEs) that are commonly used in state estimation wherein the process or measurement noise is zero-mean, Gaussian, and uncorrelated in time. Other differential equations include terms which are stochastic processes such as Brownian motion or a Wiener process. A formal approach to extend calculus to stochastic processes is so called *Itô calculus*.

References

[Simon, 2006] Simon, D. (2006). *Optimal State Estimation: Kalman, H infinity, and Nonlinear Approaches*. John Wiley & Sons.