ASEN 5044, Fall 2024

Statistical Estimation for Dynamical Systems

Lecture 16: Linear Transformations of Gaussian Random Vectors; Intro to Stochastic Processes

Prof. Nisar Ahmed (Nisar.Ahmed@Colorado.edu)

Thursday 10/10/2024

Ann and H.J. Smead Aerospace Engineering Sciences
UNIVERSITY OF COLORADO BOULDER

Announcements

- **Homework 5: Posted on Canvas/Gradescope, due Fri 10/18**
- **Quiz 5: to be posted tomorrow, due Tues 10/15**
- **Advanced topic (optional) lecture slides posted tomorrow**
- **Office hours resume per usual time/place next week**

Overview

Last Time:

- Multivariate Gaussian pdfs for jointly Gaussian random vectors
- Marginals of multivariate Gaussian pdfs

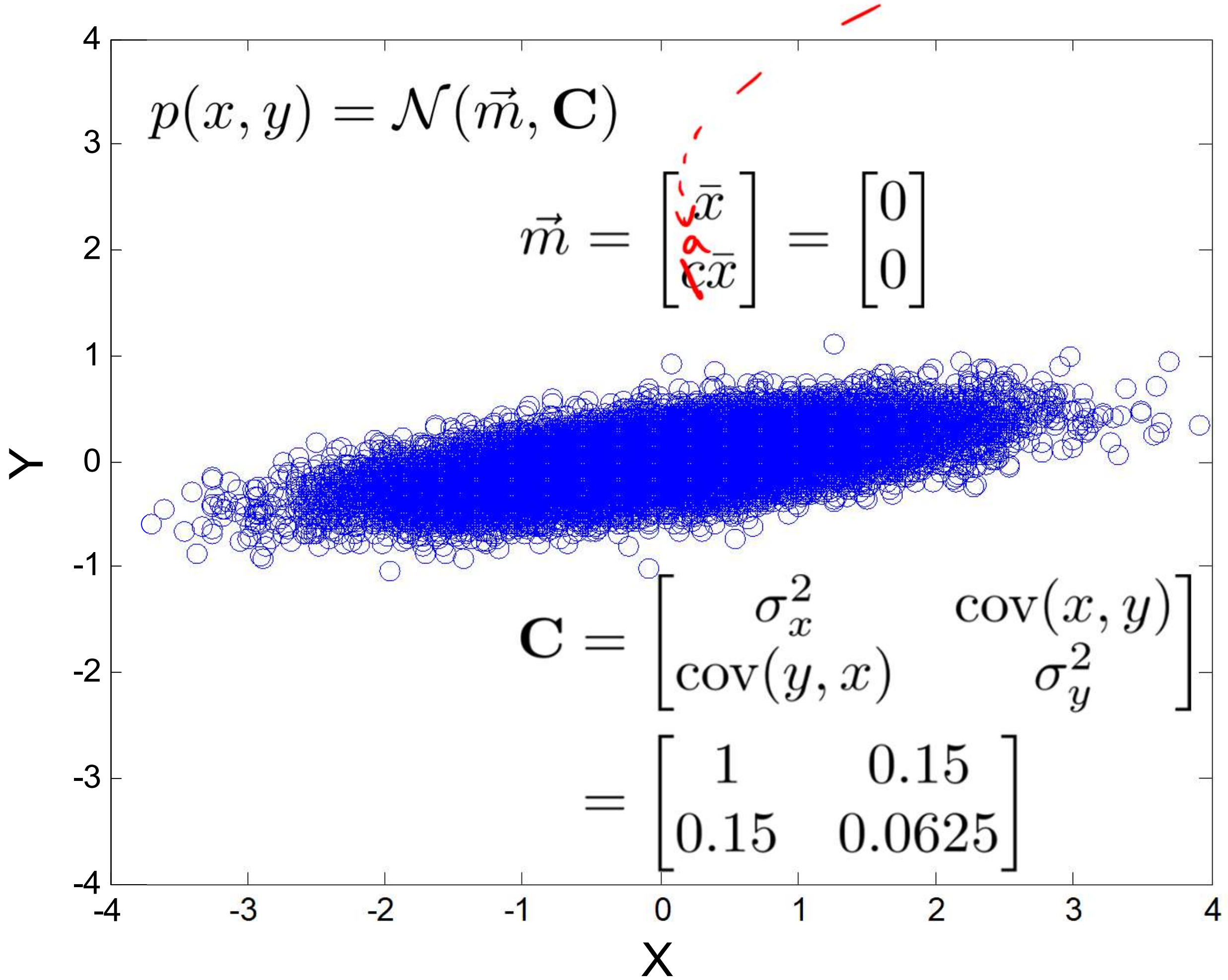
Today:

- Linear transformations of Gaussian scalars and Gaussian random vectors
  - i.e. find $p(y)$ if $p(x)$ is (multivariate) Gaussian and if $y = g(x)$ [$h(y) = x$], where $g(.)$ is linear invertible mapping $y=Ax + b$
- Intro to stochastic processes: pdfs over functions

READ SIMON BOOK, CHAPTER 3.1

Linear Transformations of Gaussian Random Vectors

- In example from last time, x and y were implied to be jointly Gaussian
- Connecting the dots: if p(x,y) = jointly Gaussian, then p(x) = Gaussian and p(y) = Gaussian
- If only given p(x)=Gaussian and y = a*x + d*e, can we verify that p(y) is indeed Gaussian?



{ one approach: could do via convolution:

if $z = ax$ & $w = d \cdot e$

then $y = z + w \rightarrow p(y) = p(z) * p(w)$

[Fact: Convolution of 2 Gaussian RV pdfs also results in a Gaussian pdf, e.g. see CLT]

→ So: what are $p(z)$ & $p(w)$? Are they Gaussian?

→ since $x \sim N(0,1)$, $e \sim N(0,1)$ } need pdfs of $z = ax$ & $w = de$

→ (*) How to get pdf of linearly transformed RVs?

Transformations of Random Variables

- Consider scalar one-to-one mapping y = g(x) (where g(.) is some invertible fxn)
- If we know p(x), then how to get p(y)?

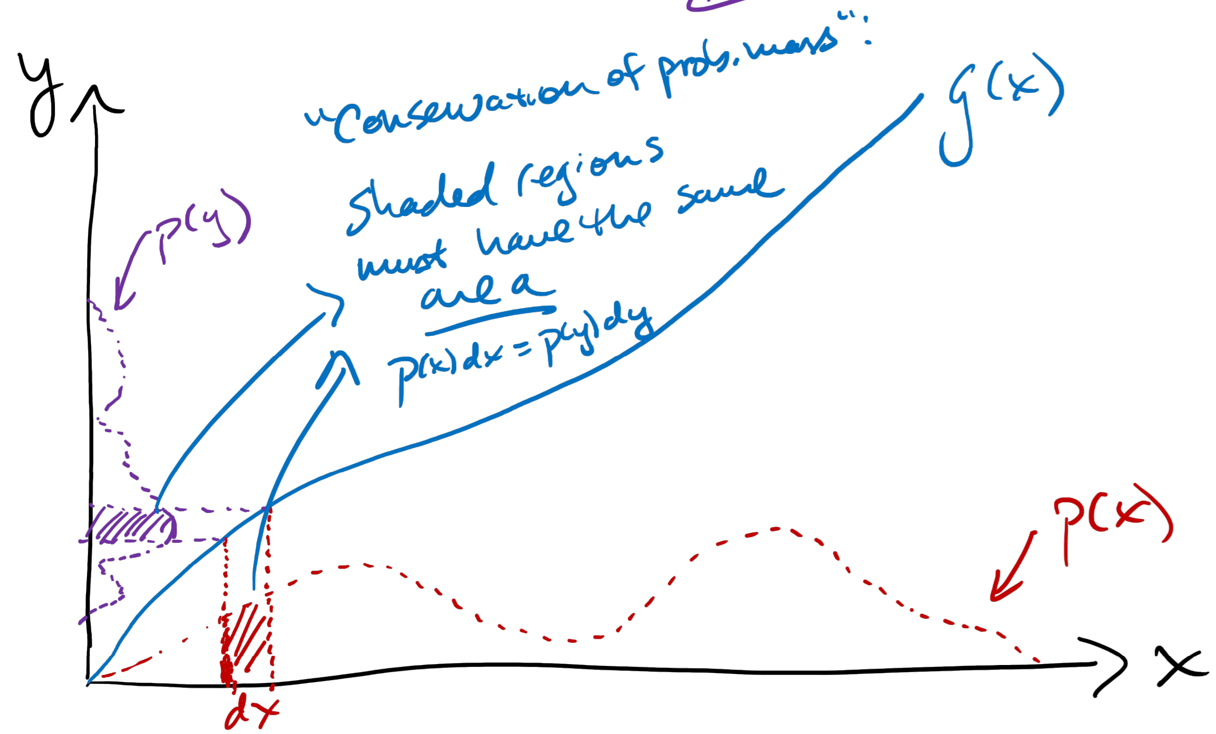
Consider $x \xmapsto{g(\cdot)} y$ where $g(x) = y$ & $g(x)$ is invertible $\left(\begin{matrix}\text{1 to 1} \\ \text{\& onto}\end{matrix}\right)$

$\rightarrow \therefore \exists$ inverse mapping $h(\cdot)$ s.t. $x = h(y)$

$\rightarrow$ Follows that probabilities that x & y lie in corresponding differential regions must be the same:

$$P(X \in [x, x+dx]) = P(Y \in [y, y+dy])$$

$$\Rightarrow \int_x^{x+dx} p_x(u)\,du = \begin{cases} \int_y^{y+dy} p_Y(\xi)\,d\xi, & \text{if } dy > 0 \\ -\int_y^{y+dy} p_Y(\xi)\,d\xi, & \text{if } dy < 0 \end{cases}$$



$$\Rightarrow p_x(x)\,dx = p_Y(y)\,|dy| \quad (\text{for } dx > 0)$$

$\rightarrow$ we know: $x = h(y) \rightarrow p_x(h(y))\,dx = p_Y(y)\,|dy| \Rightarrow \boxed{\begin{aligned} p_Y(y) &= p_x(h(y)) \cdot \left|\frac{dx}{dy}\right| \\ &= p_x(h(y)) \cdot |h'(y)| \end{aligned}}$ (*)

Example: Scaled Univariate Gaussian

- Suppose $X \sim \mathcal{N}_x(0, \sigma_x^2) = p(X)$ and $y = kx$, i.e. $g(x) = kx$ (for some constant $k \neq 0$)

$\rightarrow$ what is $p(Y) =??$

$\rightarrow$ using formula for invertible $g(\cdot)$:

$$p(Y) = p_x(h(y)) \cdot |h'(y)|, \text{ where } h(y) = x = \text{inverse of } g(\cdot) \text{ given } y$$

$$\rightarrow h(y) = \frac{y}{k} = x \qquad \rightarrow h'(y) = \frac{dh}{dy} = \frac{1}{k} \qquad \rightarrow |h'(y)| = \left|\frac{1}{k}\right| = \frac{1}{|k|}$$

$$\rightarrow \text{so } p_y(Y) = p_x\left(\frac{y}{k}\right) \cdot \frac{1}{|k|} \;\rightarrow \text{plug in def. of } p_x(X)\text{: } p_y(Y) = \mathcal{N}_x(0, \sigma_x^2)|_{x=\frac{y}{k}} \cdot \frac{1}{|k|}$$

$$= \frac{1}{|k|} \cdot \frac{1}{\sqrt{2\pi}\sigma_x} \cdot \exp\left\{-\frac{(\frac{y}{k})^2}{2\sigma_x^2}\right\}$$

$$\rightarrow p(Y) = \frac{1}{\sqrt{2\pi}\sigma_x |k|} \cdot \exp\left\{-\frac{y^2}{2(k\sigma_x)^2}\right\} = \mathcal{N}_y(0, k^2\sigma_x^2) = \mathcal{N}_y(0, \sigma_y^2) \text{ where } \sigma_y^2 = k^2 \cdot \sigma_x^2$$

***Gaussian RV x scaled by constant k $\rightarrow$ Gaussian RV y with variance scaled by k^2!! ***

Linear Transforms of Gaussian Random Vectors

- Given $\vec{Y} = A\vec{X} + \vec{b}$ for some invertible $A \in \mathbb{R}^{n\times n}$, constant $\vec{b} \in \mathbb{R}^n$, $\vec{X} = \vec{x} \in \mathbb{R}^n$ such that $\vec{X} \sim \mathcal{N}_{\vec{x}}(\vec{m}_x, C_x)$ (Gaussian random vector)

$\rightarrow \vec{Y} = \vec{y} \in \mathbb{R}^n$ should be a random vector – but what is $p_{\vec{y}}(\vec{Y}) =$??

$\rightarrow$ since we know transformation $\vec{X} \rightarrow \vec{Y}$ is invertible, we can use a generalized Jacobian formula that extends previous scalar $g(\cdot)/h(\cdot)$ transform formula to vector setting:

$$p_{\vec{y}}(\vec{Y}) = p_{\vec{x}}(h(\vec{Y})) \cdot |J_d(\tfrac{x}{y})|,$$

← abs value of determinant

where $h(\vec{Y}) = g^{-1}(\vec{Y}) = \vec{X}$ (if $\vec{Y} = g(\vec{X})$),

$J_d(\frac{x}{y}) = \det\left(\left[\frac{\partial h}{\partial y}\right]\right)$ is det of Jacobian matrix of $\vec{x} = h(\vec{y})$ w.r.t. $\vec{y}$

$\rightarrow$ in this case: to get $h(\vec{y})$, solve $\vec{y} = A\vec{x} + \vec{b}$ to get $\vec{x}$ in terms of $\vec{y}$:

$$h(\vec{y}) = \vec{x} = A^{-1}(\vec{y} - \vec{b}) = A^{-1}\vec{y} - A^{-1}\vec{b}$$

Linear Transformations of Gaussian Random Vectors

- Now need to find $J(\frac{x}{y}) = \left[\frac{\partial h}{\partial y}\right]$ from $\vec{x} = \vec{h}(\vec{y}) = A^{-1}\vec{y} - A^{-1}\vec{b}$

$\rightarrow$ write $\vec{x}$ out component-wise by expanding $\vec{x} = \vec{h}(\vec{y}) = A^{-1}\vec{y} - A^{-1}\vec{b}$, where

$$\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix},$$

Denote $A^{-1} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix}$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_{11}y_1 + d_{12}y_2 + \cdots + d_{1n}y_n - (d_{11}b_1 + d_{12}b_2 + \cdots + d_{1n}b_n) \\ d_{21}y_1 + d_{22}y_2 + \cdots + d_{2n}y_n - (d_{21}b_1 + d_{22}b_2 + \cdots + d_{2n}b_n) \\ \vdots \\ d_{n1}y_1 + d_{n2}y_2 + \cdots + d_{nn}y_n - (d_{n1}b_1 + d_{n2}b_2 + \cdots + d_{nn}b_n) \end{bmatrix} = \begin{bmatrix} h_1(\vec{y}) \\ h_2(\vec{y}) \\ \vdots \\ h_n(\vec{y}) \end{bmatrix}$$

$$= h(\vec{y})$$

Linear Transformations of Gaussian Random Vectors

- So, since:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} d_{11}y_1 + d_{12}y_2 + \cdots + d_{1n}y_n - (d_{11}b_1 + d_{12}b_2 + \cdots + d_{1n}b_n) \\ d_{21}y_1 + d_{22}y_2 + \cdots + d_{2n}y_n - (d_{21}b_1 + d_{22}b_2 + \cdots + d_{2n}b_n) \\ \vdots \\ d_{n1}y_1 + d_{n2}y_2 + \cdots + d_{nn}y_n - (d_{n1}b_1 + d_{n2}b_2 + \cdots + d_{nn}b_n) \end{bmatrix} = \begin{bmatrix} h_1(\vec{y}) \\ h_2(\vec{y}) \\ \vdots \\ h_n(\vec{y}) \end{bmatrix}$$

It follows that Jacobian matrix is:

$$J(\frac{x}{y}) = \left[\frac{\partial h}{\partial y}\right] = \begin{bmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} & \cdots & \frac{\partial h_1}{\partial y_n} \\ \vdots & & & \\ \frac{\partial h_n}{\partial y_1} & \frac{\partial h_n}{\partial y_2} & \cdots & \frac{\partial h_n}{\partial y_n} \end{bmatrix} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix} = A^{-1}$$

$$\Rightarrow J(\tfrac{x}{y}) = A^{-1}, \quad \text{and so } J_d(\tfrac{x}{y}) = \det[A^{-1}]$$

Linear Transformations of Gaussian Random Vectors

- Now substitute $\vec{h}(\vec{y}) = A^{-1}\vec{y} - A^{-1}\vec{b}$ and $J_d(\frac{x}{y}) = \det[A^{-1}]$ into $p_{\vec{y}}(\vec{Y})$ formula:

$$p_{\vec{y}}(\vec{Y}) = p_{\vec{x}}(h(\vec{y})) \cdot |J_d(\tfrac{x}{y})| = \mathcal{N}_{\vec{x}}(\vec{m}_x, C_x)|_{\vec{x}=h(\vec{y})} \cdot |\det[A^{-1}]|$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}|C_x|^{\frac{1}{2}}} \cdot \exp\left\{-\frac{1}{2}(A^{-1}\vec{y} - A^{-1}\vec{b} - \vec{m}_x)^T C_x^{-1}(\cdots)\right\} \cdot |\det[A^{-1}]|$$

Note: $\vec{m}_y = \mathbb{E}[\vec{Y}] = \mathbb{E}[A\vec{X} + \vec{b}] = A\mathbb{E}[\vec{X}] + \vec{b} = A\vec{m}_x + \vec{b} \quad \Rightarrow \vec{m}_y = A\vec{m}_x + \vec{b}$

How to get $\vec{m}_y = A\vec{m}_x + \vec{b}$ into $\exp(\cdot)$ argument for Gaussian pdf over $\vec{Y}$?

Trick: note: $A^{-1}\vec{y} - A^{-1}\vec{b} - \vec{m}_x = A^{-1}\vec{y} - A^{-1}\vec{b} - I \cdot \vec{m}_x = A^{-1}\vec{y} - A^{-1}\vec{b} - \overset{= I!}{A^{-1}A} \cdot \vec{m}_x$

$$= A^{-1}(\vec{y} - \vec{b} - A \cdot \vec{m}_x) = A^{-1}(\vec{y} - (A\vec{m}_x + \vec{b})) = A^{-1}(\vec{y} - \vec{m}_y)$$

$\rightarrow$ So Mahalanobis dist. term in $\exp(\cdot)$ is: $-\frac{1}{2}[A^{-1}(\vec{y} - \vec{m}_y)]^T C_x^{-1}[\cdots]$

$$\rightarrow \; -\frac{1}{2}[(\vec{y}-\vec{m}_y)]^T A^{-T} C_x^{-1} A^{-1}[(\vec{y}-\vec{m}_y)] = -\frac{1}{2}[(\vec{y}-\vec{m}_y)]^T (AC_xA^T)^{-1}[(\vec{y}-\vec{m}_y)]$$

Linear Transformations of Gaussian Random Vectors

- Finally, use the facts that $|\det(A^{-1})| = |\frac{1}{\det(A)}| = \frac{1}{|\det(A)|} = \frac{1}{|A|^{\frac{1}{2}}|A|^{\frac{1}{2}}}$

$$|\det(A)|^{\frac{1}{2}} \cdot |\det(A^T)|^{\frac{1}{2}} \cdot |\det(C_x)|^{\frac{1}{2}} = |AC_xA^T|^{\frac{1}{2}}$$

$\rightarrow$ plug everything into $p_{\vec{y}}(\vec{Y})$ formula – **AND BEHOLD!**:

$$p_{\vec{y}}(\vec{Y}) = \frac{1}{(2\pi)^{\frac{n}{2}}|AC_xA^T|^{\frac{1}{2}}} \cdot \exp\left\{-\frac{1}{2}(\vec{y}-\vec{m}_y)^T(AC_xA^T)^{-1}(\vec{y}-\vec{m}_y)\right\}$$

$$= \mathcal{N}_{\vec{y}}(\vec{m}_y, C_y)$$

where: mean of $\vec{Y} = E[\vec{Y}] = \vec{m}_y = A\vec{m}_x + \vec{b}$,

covariance matrix of $\vec{Y} = E[(\vec{Y}-\vec{m}_y)(\cdots)^T] = C_y = AC_xA^T$,

if $\vec{Y} = A\vec{X} + \vec{b}$ for $\vec{X} \sim \mathcal{N}_{\vec{x}}(\vec{m}_x, C_x)$

(*) **Linear transformations of Gaussian random vector X → another Gaussian random vector Y!**

Summary:
Multivariate Gaussian PDFs and Linear Transformations

We have covered a few key concepts and facts thus far:

- Multivariate Gaussian pdfs for random vectors (Gaussian random vectors)
- Marginal pdfs for Gaussian random vectors
- Linear transform of Gaussian random vector x → Gaussian random vector y
  - Gaussian random vectors x and y also have a joint multivariate Gaussian pdf
  - Properties extend to case where linear mapping from x→y is not invertible

Modeling Random Noise in Dynamical Systems

- How to account for all of the <u>random physical disturbances</u> that robot encounters?

?

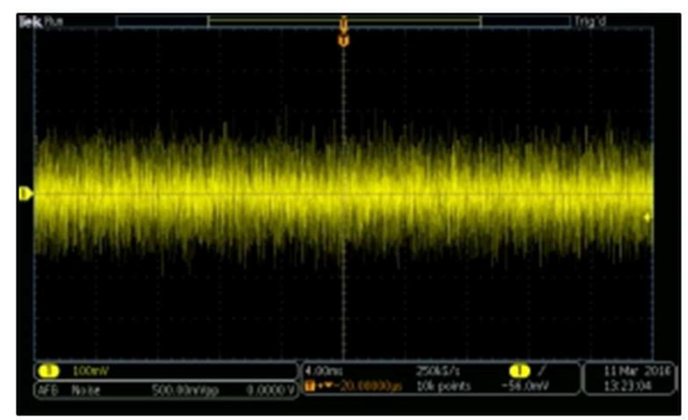
$$x(k+1) = Fx(k) + Gu(k) + w(k)$$

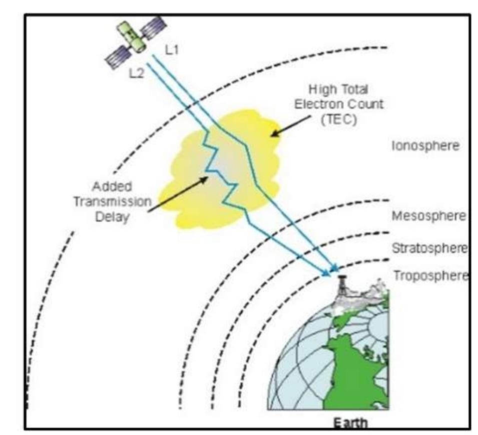
dyn. (process noise)

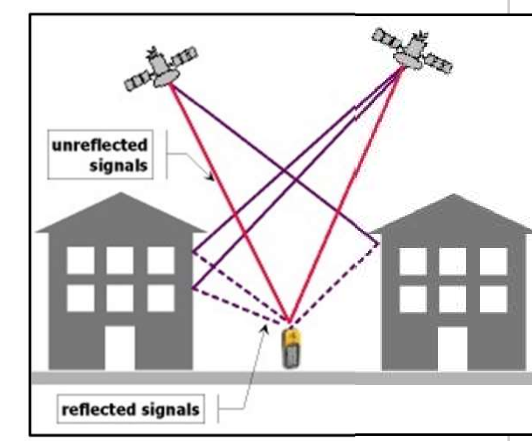
noise inputs and resulting states/sensor data change <u>randomly</u> over <u>time</u>!!!

(sensor/observation noise)

$$y(k+1) = Hx(k+1) + v(k+1)$$







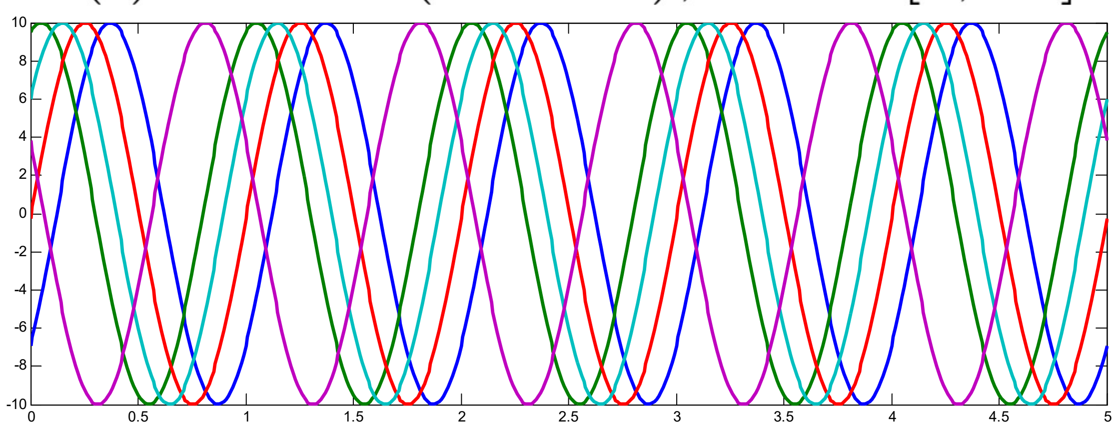
Stochastic Process Models (Continuous Time)

(CT) Stochastic Process = a time sequence X(t) representing evolution of a variable X subject to unpredictable/random variations (i.e. a "random function")

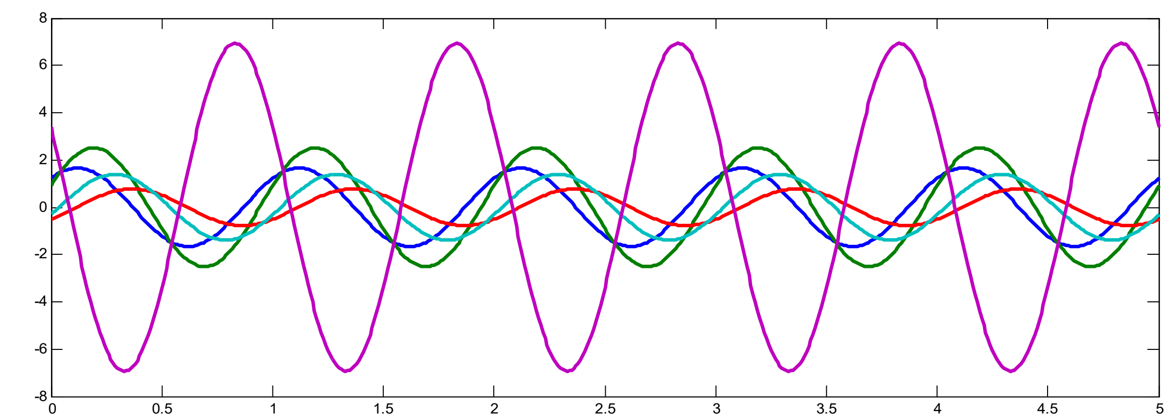
- "random tweaks" of deterministic dynamics (i.e. corrupted signal carrying information)
- purely random noise (i.e. signal with absolutely no information)

Example **realization ensembles (i.e. random function samples)** of stochastic processes:

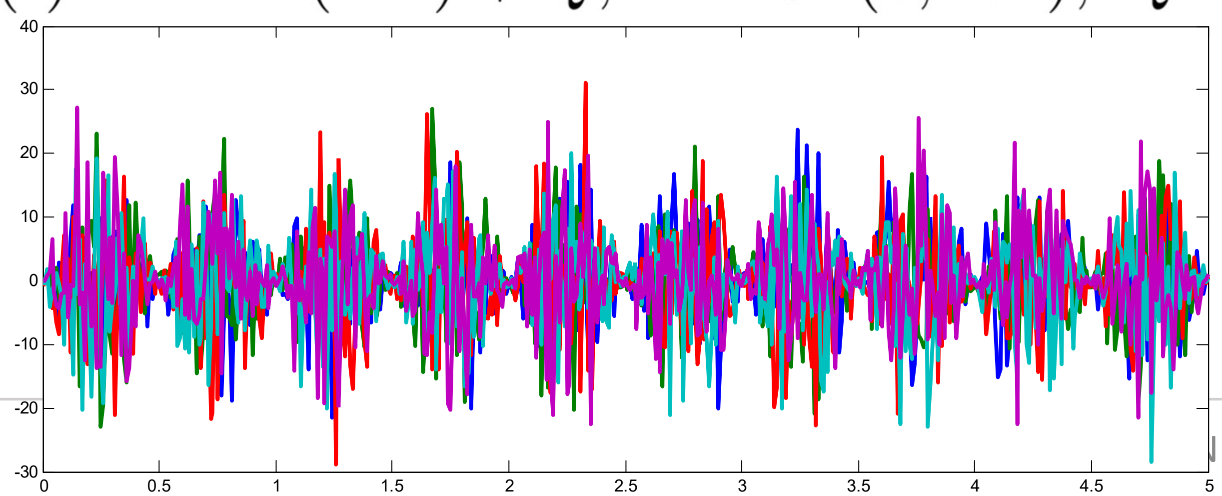
$$X(t) = 10\sin(2\pi t + \theta),\ \theta \sim \mathcal{U}[0, 2\pi]$$



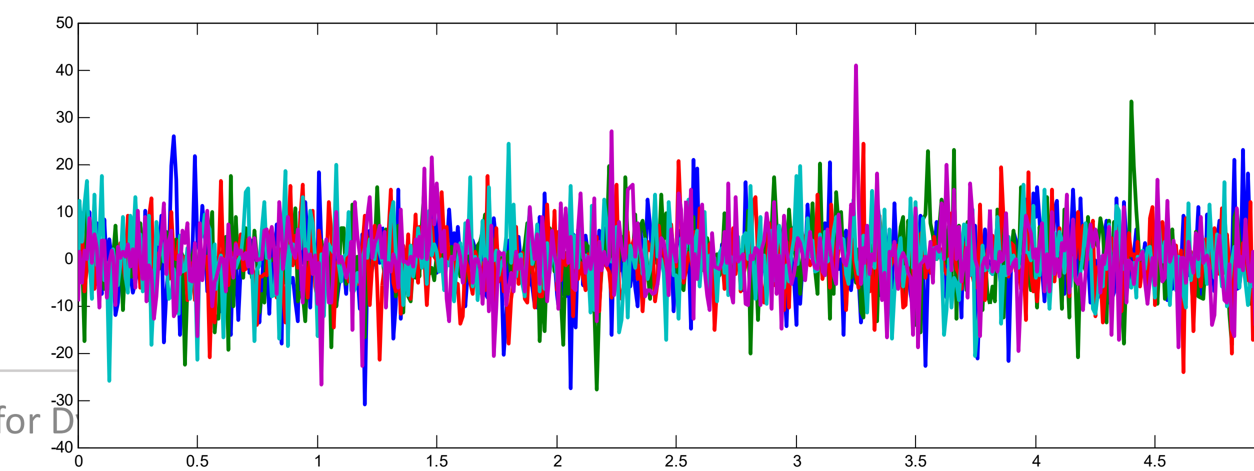
$$X(t) = A\sin(2\pi t + \theta),\ A \sim \mathcal{N}(0, 1),\ \theta \sim \mathcal{U}[0, 2\pi]$$



$$X(t) = A\sin(2\pi t) + Q,\ A \sim \mathcal{N}(0, 100),\ Q \sim \mathcal{U}[-1, 1]$$



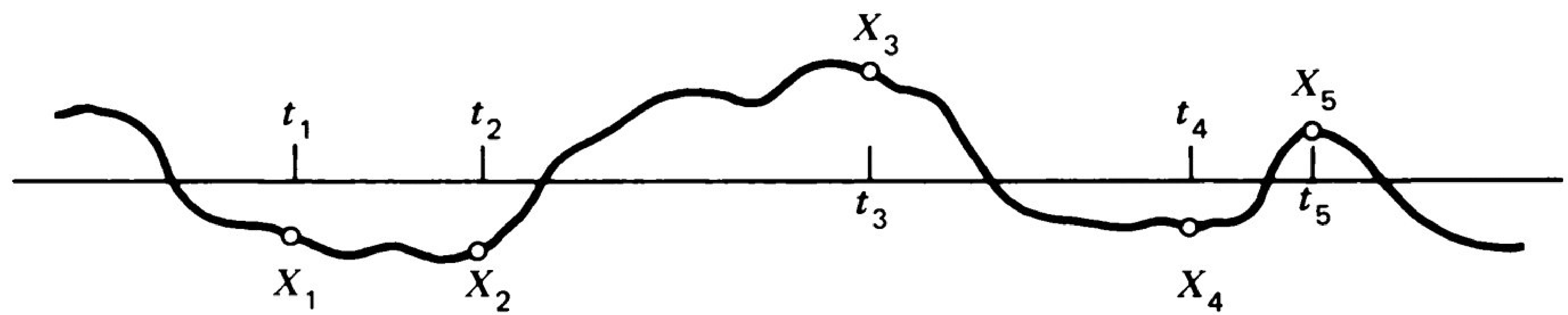
$$X(t) = A\sin(\theta),\ A \sim \mathcal{N}(0, 100),\ \theta \sim \mathcal{U}[0, 2\pi]$$



Stochastic Process Models (Continuous Time)

- Realization of random signal X(t) could be sampled at arbitrary points in time

$$X(t_1) \triangleq X_1$$
$$X(t_2) \triangleq X_2$$
$$\vdots$$
$$X(t_n) \triangleq X_n$$



- **A stochastic process is fully defined by pdfs for all possible sets of samples of X(t)**

instantaneous pdf values: $p(X_1), p(X_2), ..., p(X_n)$

pairwise joint pdfs (rates of change): $p(X_1, X_2), p(X_2, X_3), p(X_1, X_3)...$

$$\vdots$$

in general: would need infinite dimensional pdfs!: $p(..., X_1, X_2, ..., X_n, ...)$

(tricky to exactly specify/obtain the full joint pdf in many practical cases)

Easier Ways to Specify/Characterize a Random Process

- Rather than specifying or writing out full joint pdfs, can describe “moments” of a signal
- Mean at time i: $\bar{X}(t_i) = \bar{X}_i = E[X(t_i)] = \int_{-\infty}^{\infty} X(t_i)p(X(t_i))dX(t_i)$
- **Autocorrelation function:**

$$R_X(t_i, t_j) = E[X(t_i)X(t_j)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} X(t_i)X(t_j)p(X(t_i),X(t_j))dX(t_i)dX(t_j)$$

- Autocovariance function $\triangleq E[(X(t_i) - \bar{X}_i)(X(t_j) - \bar{X}_j)]$