# ASEN 5044, Fall 2018

## Statistical Estimation for Dynamical Systems

## Lecture 13: Multivariate Expectations, Covariance, Correlation, and Multivariate Gaussian PDFs

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

Thurs 10/04/2018

University of Colorado Boulder

Aerospace Engineering

# Announcements

- **HW 4 Due Now, Solutions to be posted later today…**

- **<u>Midterm 1: out today</u>**
  - **One week long take home exam posted to Canvas**
  - **Due Thurs 10/11/2017 on Canvas by 11 am**
  - Open book/notes – must complete by yourself (honor code applies)

- Prof. Ahmed out of town Tues afternoon 10/9 thru Thurs morning 10/11
  - No office hours Tuesday

- Advanced Topic Lecture #2: tomorrow (Fri 10/5/18), 1 pm -1:50 pm
  - Will be recorded/posted

# Last Time…

- Sums of independent random variables
  - Central limit theorem

- Gaussian (Normal) random variables and PDFs/distributions

- PDFs for multiple random variables
  - Joint pdfs
  - Marginal and conditional pdfs

# Today…

- Multivariate expected values
  - mean, covariance, correlation

- Random vectors

- Multivariate Normal (Gaussian) PDFs for random vectors

**READ SIMON BOOK, CHAPTER 2.7**

# From last time: Marginal pdfs and Conditional pdfs

- As with joint probability tables discussed earlier, joint pdfs tell "the whole story"
- Analogous expressions exist for marginalization, conditioning, Bayes' rule, independence
- **Basically just need to replace summations with integrals**

– Marginal pdf: $p(x) = \int_{-\infty}^{\infty} p(x,y)\,dy$ [marginal pdf for x: "averaging" $p(x,y)$ w.r.t. y]

$p(y) = \int_{-\infty}^{\infty} p(x,y)\,dx$ [marginal pdf for y: "averaging" $p(x,y)$ w.r.t. x]

– Conditional pdf: $p(x \mid y=\underset{\substack{\text{some fixed} \\ \text{value for } y}}{\bar{y}}) \triangleq \dfrac{p(x, y=\bar{y})}{p(y=\bar{y})} = \dfrac{p(x, y=\bar{y})}{\int_{-\infty}^{\infty} p(x, y=\bar{y})\,dx}$ [likewise: $p(y \mid x=\bar{x}) = \dfrac{p(x=\bar{x}, y)}{\int_{-\infty}^{\infty} p(x=\bar{x}, y)\,dy}$]

– Bayes' Rule for pdfs: (invert cond. pdfs): $p(x \mid y=\bar{y}) = \dfrac{p(x)\cdot p(y=\bar{y} \mid x)}{p(y=\bar{y})} = \dfrac{p(x)\cdot p(y=\bar{y} \mid x)}{\int_{-\infty}^{\infty} p(x)\, p(y=\bar{y} \mid x)\,dx}$

– Independence: cont. RVs x & y are $\perp\!\!\!\perp$ if & only if $p(x,y) = p(x)\cdot p(y)$ $\forall$ x & y values

(*) All of the above extend to n-dimensional pdfs $p(x_1, x_2, \ldots, x_n)$ in obvious ways

# Expectation Operators and Expected Values for Multivariate pdfs

- Also extend in obvious way – use n-dimensional integrals, etc.
- Key trick is that functions being integrated could take in multiple arguments

Suppose you RVs $x_1, x_2, \ldots, x_n$ w/ joint pdf $p(x_1, x_2, \ldots, x_n)$

then $E[g(x_1, x_2, \ldots, x_n)] = \int_{x_1}\int_{x_2}\cdots\int_{x_n} g(x_1, x_2, \ldots, x_n) \cdot p(x_1, x_2, \ldots, x_n)\, dx_n\, dx_{n-1} \cdots dx_2\, dx_1$

= some (scalar) constant

$\longrightarrow$ can also take expectation w.r.t only <u>some</u> subset of the n RVs:

e.g. $E[g(x_1, x_2, \ldots, x_n)]_{(x_2, x_3, \ldots, x_n)}$

$$\triangleq \int_{x_2}\int_{x_3}\cdots\int_{x_n} g(x_1, x_2, \ldots, x_n)\, p(x_1, x_2, \ldots, x_n)\, dx_n \cdots dx_2$$

= some function of $x_1$ <u>only</u> = marginal expectation of $g(\cdot)$ w.r.t. $x_2 \cdots x_n$

# Useful (Scalar-Valued) Multivariate Moments and Expectations

- Expectation of XY (product of X and Y, i.e. "cross-moment"):

if $E[XY] = 0$, then we say $X$ & $Y$ are orthogonal RVs

$$E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \cdot p(x,y)\,dx\,dy$$

$$\text{If } X \perp\!\!\!\perp Y: \quad E[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \cdot p(x)\cdot p(y)\,dx\,dy = \left(\int_{-\infty}^{\infty} x\cdot p(x)\,dx\right)\left(\int_{-\infty}^{\infty} y\cdot p(y)\,dy\right)$$

$$= \boxed{E[x]\cdot E[y] = E[XY] \text{ if \& only if } X \perp\!\!\!\perp Y}$$

- Covariance of X and Y: how **linearly related** X and Y are to each other

$$\text{cov}(x,y) \triangleq E[(x-\underbrace{\bar{x}}_{E[x]})(y-\underbrace{\bar{y}}_{E[y]})] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\bar{x})(y-\bar{y})\cdot p(x,y)\,dx\,dy$$

- Correlation coefficient: covariance normalized by product of standard devs

$$\text{corr}(x,y) = \rho(x,y) \triangleq \frac{\text{cov}(x,y)}{\text{stdev}(x)\cdot\text{stdev}(y)} = \frac{E[(x-\bar{x})(y-\bar{y})]}{\sqrt{\text{var}(x)}\sqrt{\text{var}(y)}}$$

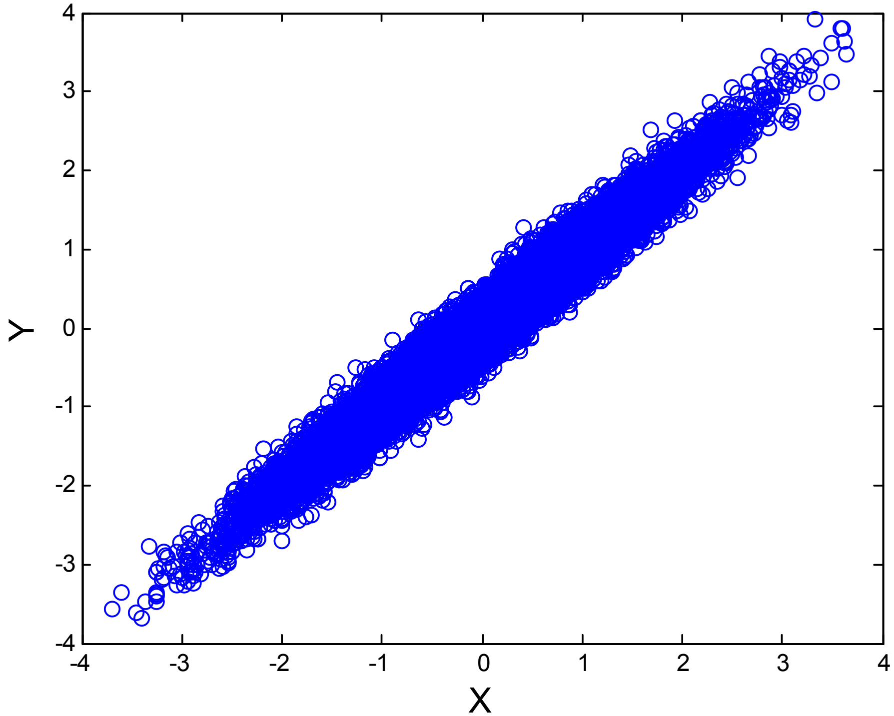
$$\boxed{\rho \in [-1, 1]}$$

where $\text{var}(x) = E[(x-\bar{x})(x-\bar{x})]$

$\text{var}(y) = E[(y-\bar{y})(y-\bar{y})]$

# Example: Sample Covariances and Correlations

- Consider x ~ N(0,1) with 20,000 samples
- Evaluate y = x + 0.2*e, where error e ~ N(0,1)




Sample expected values (Matlab)

$cov(x, y) = 0.9916$

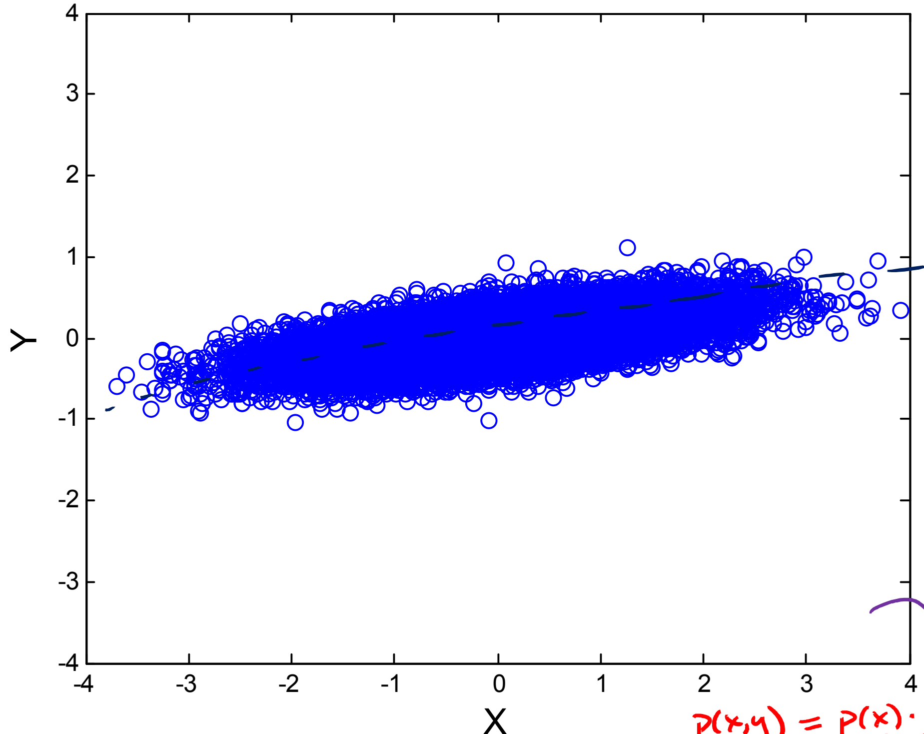
$std(x) = \sqrt{0.9911}$

$std(y) = \sqrt{1.329}$

$corr(x, y) = \rho = \frac{0.9916}{\sqrt{0.9911}\sqrt{1.329}} = 0.9801$

# Example: Sample Covariances and Correlations

- Consider $x \sim N(0,1)$ with 20,000 samples
- Evaluate $y = 0.15*x + 0.2*e$, where error $e \sim N(0,1)$




Sample values (Matlab)

$corr(x, y) = \rho = 0.6051$

→ Not a perfect fit b/w x & y

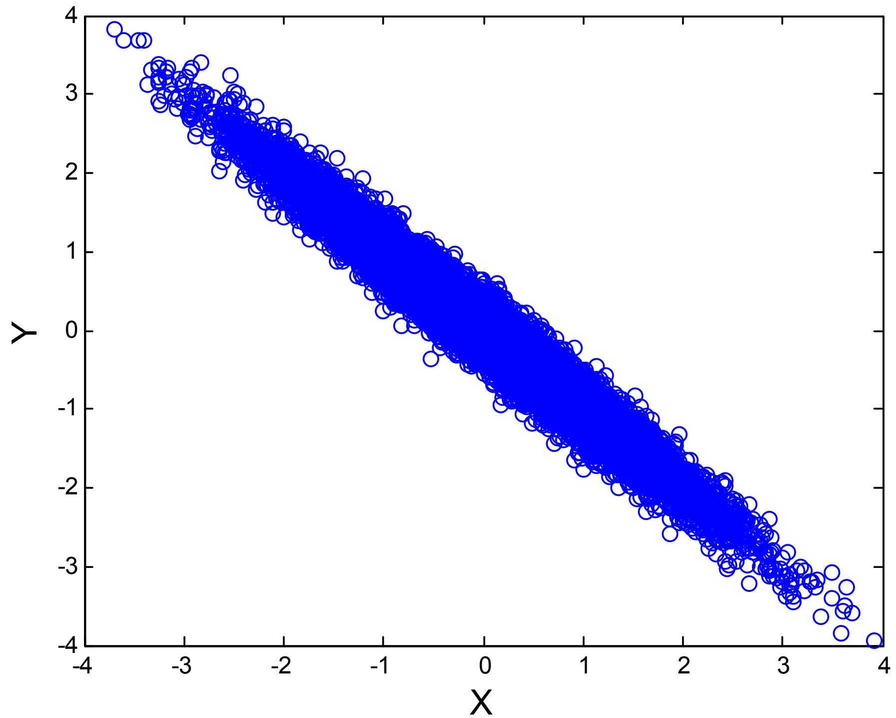
but definitely some information about x from y & vice versa

→ ( like samples from $p(x,y)$ though we don't know what $p(x,y)$ is → all you have is $p(x) \sim N(0,1)$, $p(e) \sim N(0,1)$, $y = ax + de$ )

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

# Example: Sample Covariances and Correlations

- Consider x ~ N(0,1) with 20,000 samples
- Evaluate y = -1*x + 0.2*e, where error e ~ N(0,1)



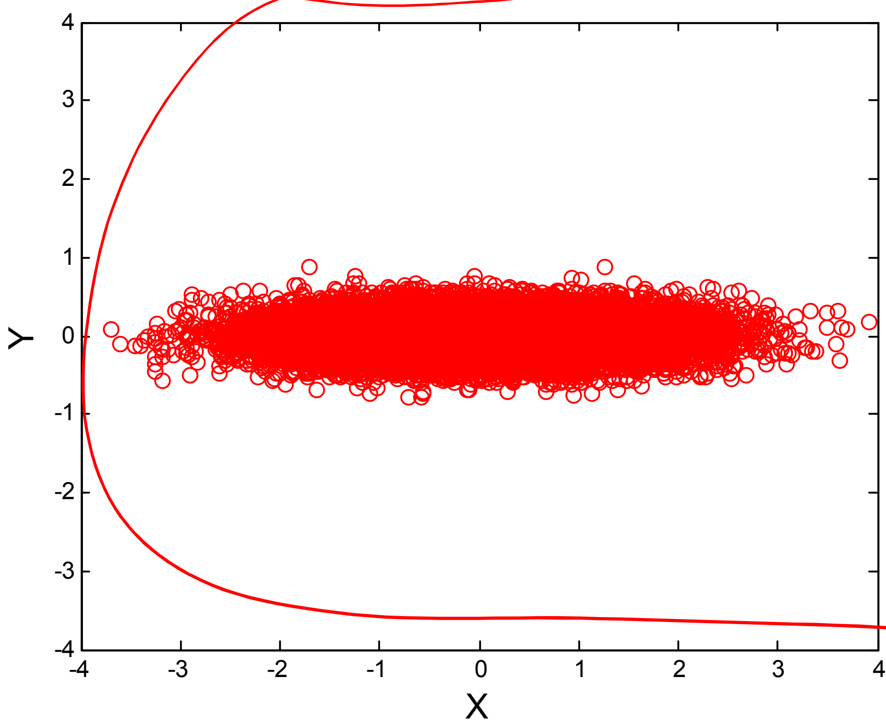

Sample values (Matlab)

$cov(x, y) = -0.9899$

$corr(x, y) = \rho = -0.9807$

# Example: Sample Covariances and Correlations

- Consider x ~ N(0,1) with 20,000 samples
- Evaluate y = 0*x + 0.2*e, where error e ~ N(0,1)




Sample values (Matlab)

$cov(x, y) = 0.0016$

$corr(x, y) = \rho = 0.0082$

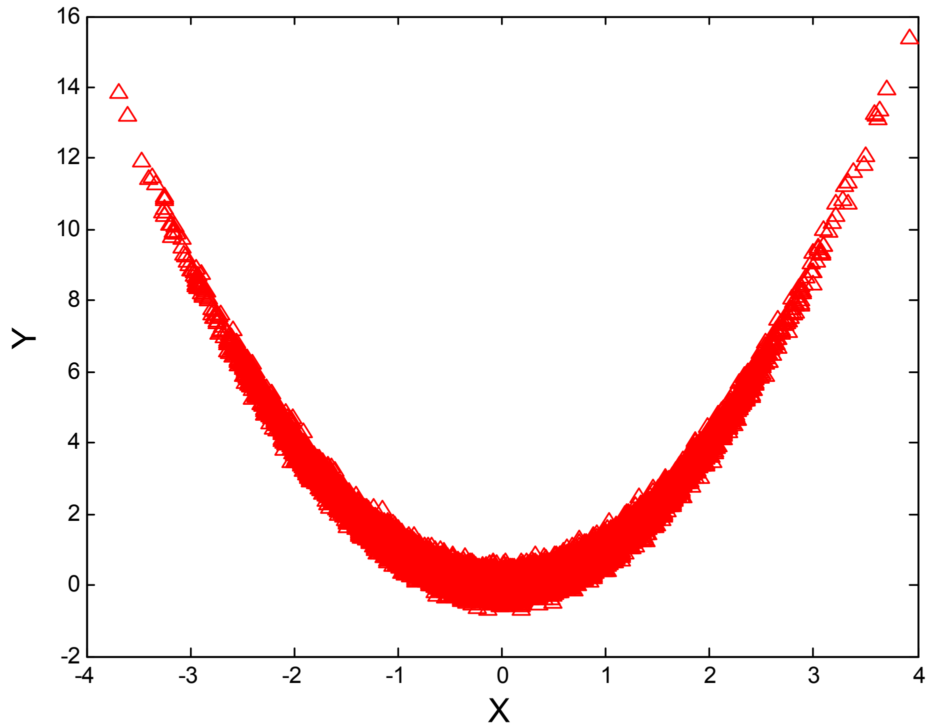
cov & corr → 0 as # samples → ∞

⇒ No info b/w x & y

there is no linear dependence

# Example: Sample Covariances and Correlations

- Consider $x \sim N(0,1)$ with 20,000 samples
- Evaluate $y = x^2 + 0.2 * e$, where error $e \sim N(0,1)$




Sample values (Matlab)

$cov(x, y) = 0.0095$

$std(x) = \sqrt{0.9911}$

$std(y) = \sqrt{2.0119}$

$corr(x, y) = \rho = 0.0067$

→ cov & corr → 0 as # samples → ∞

→ x & y are not linearly related but they are not $\perp$ !

# Difference Between Dependence and Correlation

**VERY IMPORTANT:**

- If X and Y are independent $\Rightarrow$ X and Y are uncorrelated

- If X and Y are uncorrelated $\not\Rightarrow$ X and Y are independent

# N-dimensional Random Vectors

- To deal with n-dimensional dynamic state vectors and state space models with noise, need to work with **random vectors**
- Stack continuous scalar random vars on top of each other with a well defined joint pdf

Given n scalar random vars $X_1, X_2, \ldots, X_n$, define the random vector

(vector RV) $\vec{X} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}$ w/ realizations $\vec{X} = \vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$

& (joint) pdf $p(\vec{X}) = p(X_1, X_2, \ldots, X_n)$

→ Define the mean vector as $\underset{[n\times 1]}{\vec{m}} = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix} = E[\vec{X}] = \int_{-\infty}^{\infty} \vec{x} \cdot p(\vec{x})\, d\vec{x}$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \cdot p(x_1, x_2, \ldots, x_n)\, dx_n \cdots dx_1$$

marginal mean of $X_i$ → where $m_i = E[X_i] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i \cdot p(x_1, x_2, \ldots, x_n)\, dx_n \cdots dx_1 = \bar{x}_i$

# Covariance Matrices

- Consider n-dimensional set of RVs $X_1, X_2, \ldots, X_n$ with joint pdf $p(X_1, X_2, \ldots, X_n)$
- Can assemble all variances and covariances into the nxn ***covariance matrix***

$$\underset{[n\times n]}{\underline{C}} \triangleq \begin{bmatrix} E[(x_1-m_1)^2] & E[(x_1-m_1)(x_2-m_2)] & \cdots & E[(x_1-m_1)(x_n-m_n)] \\ E[(x_2-m_2)(x_1-m_1)] & E[(x_2-m_2)^2] & \cdots & E[(x_2-m_2)(x_n-m_n)] \\ \vdots & \vdots & \ddots & \vdots \\ & & & E[(x_n-m_n)^2] \end{bmatrix}$$

$$= \begin{bmatrix} \sigma_{x_1}^2 & c_{12} & \cdots & c_{1n} \\ c_{21} & \sigma_{x_2}^2 & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ & & & \sigma_{x_n}^2 \end{bmatrix}$$

where $\sigma_{x_i}^2 = \text{var}(x_i) = E[(x_i - m_i)^2]$ (= scalar)

$c_{ij} = c_{ji} = E[(x_i - \underbrace{\bar{x}_i}_{m_i})(x_j - \underbrace{\bar{x}_j}_{m_j})]$ (= scalar)

$= \rho_{ij} \cdot \sigma_{x_i} \cdot \sigma_{x_j}$

symmetric square positive (semi)-definite matrix

(*) in vector form: $\underline{C} \triangleq E[(\vec{x}-\vec{m})(\vec{x}-\vec{m})^T] = \int_{-\infty}^{\infty} (\vec{x}-\vec{m})(\ldots)^T d\vec{x} = E[\vec{x}\vec{x}^T] - \vec{m}\vec{m}^T$

# The Multivariate Normal (Gaussian) PDF

- Random vars $X_1, \ldots, X_n$ are said to be ***jointly normal (Gaussian)*** if their joint pdf is

Special case of a quadratic form (i.e. weighted 2-norm $\|(\vec{x}-\vec{m})\|^2_{\Sigma^{-1}}$ i.e. weighted by $\Sigma^{-1}$)

$$p(\vec{X}) = \frac{1}{(2\pi)^{\frac{n}{2}}|C|^{\frac{1}{2}}} \exp\left\{ \underbrace{-\frac{1}{2}\underbrace{(\vec{x}-\vec{m})^T}_{1\times n}\underbrace{C^{-1}}_{n\times n}\underbrace{(\vec{x}-\vec{m})}_{n\times 1}}_{1\times 1 = \text{scalar}} \right\}$$

$\boxed{\text{Mahalanobis distance}}$

scalar output for vector $\vec{x}, \vec{m}$ & matrix $\Sigma$!

$$= \mathcal{N}(\vec{m}, C)$$

$\longrightarrow$ defined completely by mean $\vec{m} \in \mathbb{R}^n$ & covar matrix $\Sigma \in \mathbb{R}^{n\times n}$

$\longrightarrow$ If $\vec{x} \sim \mathcal{N}(\vec{m}, \Sigma)$, then $\vec{x}$ is said to be a Gaussian random vector

- Matlab: can compute $p(\vec{x})$ for any given $\vec{x}$, $\vec{m}$, & $\Sigma$ using "mvnpdf.m"