TAM 598 Lecture 6 :

RANDOM VECTORS AND MULTIVARIATE GAUSSIANS

Announcements:

- HW 1 covers lectures 1-4 ; due on Feb 12 ; to be submitted via CANVAS
- HW2 covers lectures 4-8 ; due on Feb 26

Random Vectors – take N random variables $X_1, X_2, \dots, X_N$ and put them in a vector

$$\underset{\sim}{X} = (X_1, \dots, X_N)$$

Random vectors are used to model uncertain quantities such as

- → state of a multibody system : vector of coordinates & velocities
- → an unknown function : vector of its function values at N test points
- → an image : vector of pixel values

PDF of a Random Vector is the joint PDF of its components.

$$p(\underset{\sim}{x}) = p(x_1, x_2, \dots, x_N)$$

- can marginalize $p(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1, x_2, \dots x_N)\, dx_2\, dx_3 \dots dx_N$
- can integrate out subsets $p(x_1, x_2) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1, x_2, \dots x_N)\, dx_3 \dots dx_N$

Expectation of a Random Vector is the vector of expectations of each component

$$\mathbb{E}[\underset{\sim}{X}] = \begin{pmatrix} \mathbb{E}[X_1] \\ \vdots \\ \mathbb{E}[X_2] \end{pmatrix}$$

This expectation is linear:

$$\mathbb{E}[a\underset{\sim}{X} + b\underset{\sim}{Y}] = a\,\mathbb{E}[\underset{\sim}{X}] + b\,\mathbb{E}[\underset{\sim}{Y}]$$

Covariance Matrix of Two Random Vectors

Let $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ be random vectors of dimension N and M respectively. The covariance of $\underset{\sim}{X}$ and $\underset{\sim}{Y}$ is the $N \times M$ matrix consisting of all covariances between components of $\underset{\sim}{X}$ and $\underset{\sim}{Y}$, ie

$$\mathbb{C}[\underset{\sim}{X}, \underset{\sim}{Y}] = \left(\mathbb{C}[X_i, Y_j] \right)$$

- sometimes $\mathbb{C}[\underset{\sim}{X}, \underset{\sim}{Y}]$ is called Σ ("sigma")

- we can rewrite the covariance matrix as

$$\mathbb{C}[\underset{\sim}{X}, \underset{\sim}{Y}] = \mathbb{E}\left[(\underset{\sim}{X} - \mathbb{E}[\underset{\sim}{X}])(\underset{\sim}{Y} - \mathbb{E}[\underset{\sim}{Y}])^T \right]$$

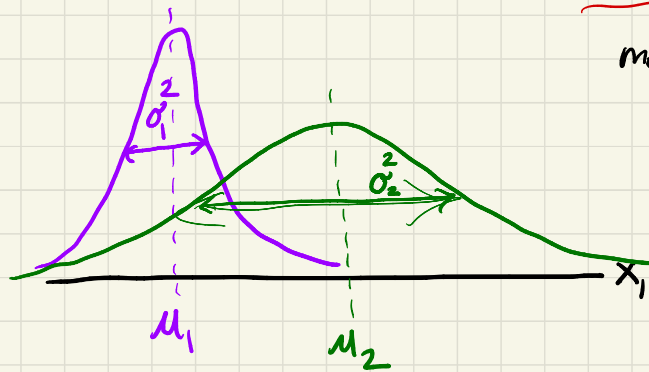
- the $N \times N$ matrix $\mathbb{C}[\underset{\sim}{X}, \underset{\sim}{X}]$ is the self covariance or just the covariance matrix of $\underset{\sim}{X}$.

  diagonal: variance of each component of $\underset{\sim}{X}$

  off-diagonal: covariance of X_i, X_j for $i \neq j$

Multivariate Gaussian Distributions

recall univariate Gaussian distributions



"drawn from"

$$X \sim N(\mu, \sigma^2)$$

mean

variance

PDF :

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

properties of univariate Gaussian distributions :

① closed under linear transformation

corrected typo in class lecture

$$X \sim N(\mu, \sigma^2) \qquad Y = aX + b \qquad Y \sim N(\cancel{\mu + b}\ a\mu + b,\ a^2\sigma^2)$$

② adding / subtracting Gaussian RVs

$$X \sim N(\mu_x, \sigma_x^2)$$
$$Y \sim N(\mu_y, \sigma_y^2)$$

$$X \pm Y = N(\mu_x \pm \mu_y,\ \sigma_x^2 + \sigma_y^2)$$

$\boxed{N \sim (\mu, \sigma^2)}$

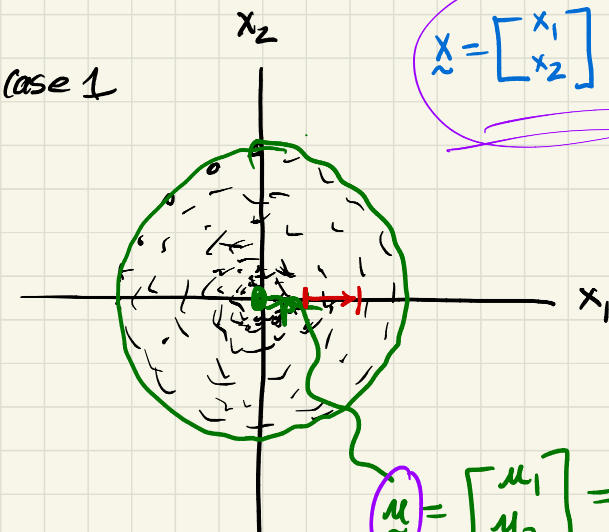
Bivariate Gaussian Distributions

– two random variables from which we construct a two-component random vector

$$\boxed{N \sim (\underset{\sim}{\mu}, \underset{\approx}{\Sigma})}$$

Case 1

$$\underset{\sim}{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$$



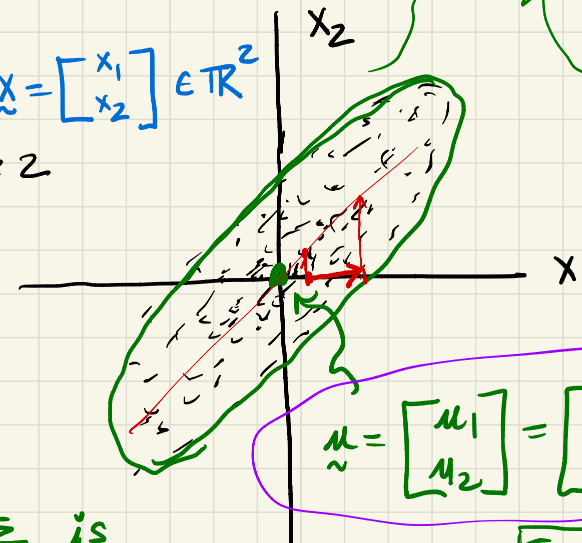
$$\underset{\sim}{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\underset{\approx}{\Sigma} = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$

$\underset{\approx}{\Sigma}$ is "positive definite"

Case 2

$$\underset{\sim}{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$$



$$\underset{\sim}{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\underset{\approx}{\Sigma} = \begin{bmatrix} \Sigma_{11} & \boxed{\Sigma_{12}} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} = \begin{bmatrix} 5 & 2{,}5 \\ 2{,}5 & 5 \end{bmatrix}$$

recall how we defined covariance of two random variables :

$$\operatorname{cov}(x_1, x_2) = \mathbb{E}\left((x_1 - \mathbb{E}[x_1])(x_2 - \mathbb{E}[x_2]) \right)$$

- if $\mathbb{E}[X_1] = 0$, $\mathbb{E}[X_2] = 0$ then

$$\operatorname{cov}(x_1, x_2) = \mathbb{E}[X_1 X_2] \qquad \Sigma_{12}$$

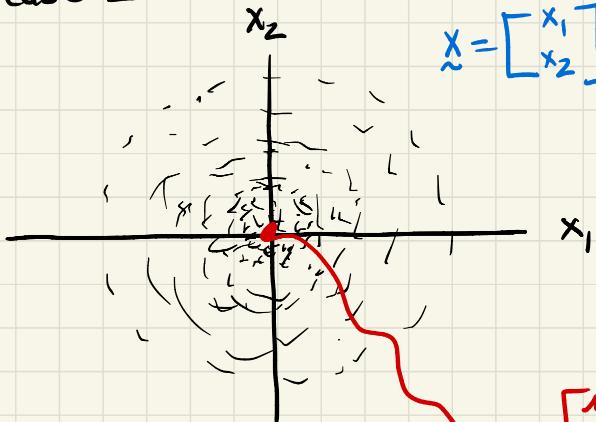
- $$\operatorname{cov}(X_1, X_1) = \mathbb{E}\left((x_1 - \mathbb{E}[x_1])(x_1 - \mathbb{E}[x_1]) \right)$$

$$= \mathbb{E}[x_1^2] = \mathbb{V}(X_1)$$

- another way to think of $\operatorname{cov}(x_1, x_2)$: a measure of similarity of x_1 and x_2. It is a lot like a dot product

$$\begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \qquad \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0$$

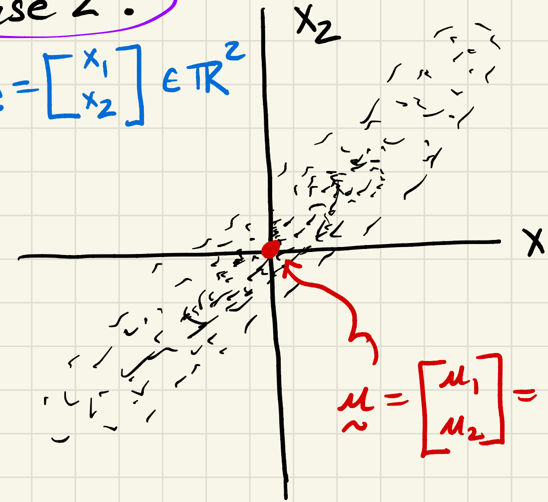
case 1:



$\underset{\sim}{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$

$\underset{\sim}{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Case 2:



$\underset{\sim}{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2$

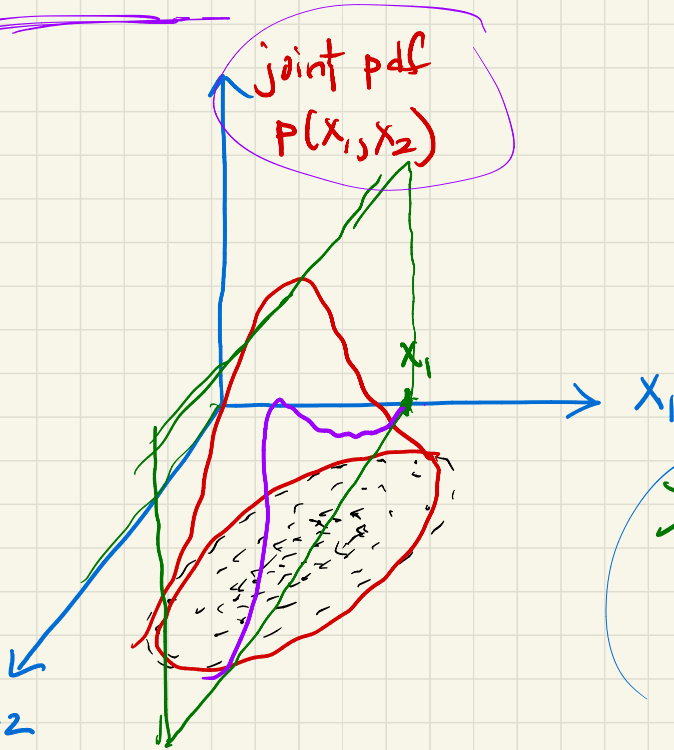
$\underset{\sim}{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

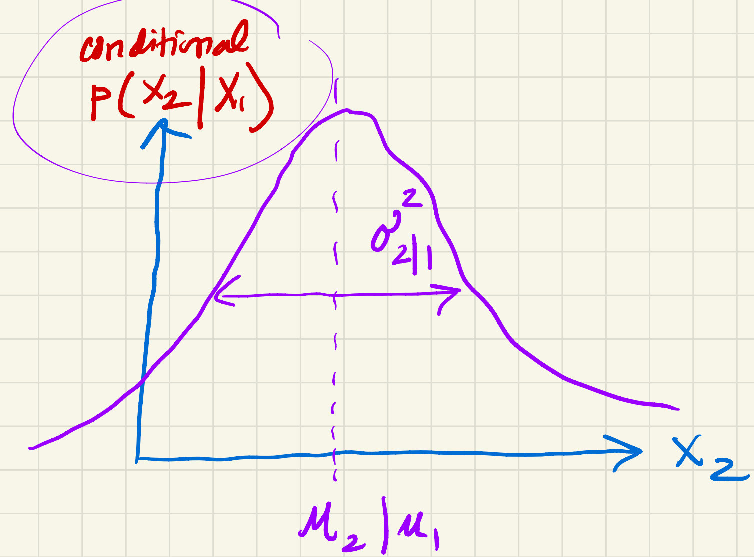
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$$

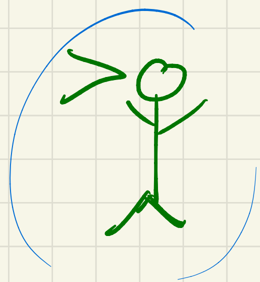
covariance matrix

$$\boxed{\Sigma_{ij} = \mathrm{cov}(X_i, X_j)}$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$$







MULTIVARIATE GAUSSIAN THEOREM

text: Kevin Murphy Ch 4
Bishop ch 2.3

(allows us to reason about the marginals & conditionals of a multivariate Gaussian.)

Suppose $\underset{\sim}{X} = (\underset{\sim}{X}_1, \underset{\sim}{X}_2)$ is jointly Gaussian with parameters

$$\underset{\sim}{\mu} = \begin{bmatrix} \underset{\sim}{\mu}_1 \\ \underset{\sim}{\mu}_2 \end{bmatrix} \qquad \underset{\sim}{\Sigma} = \begin{pmatrix} \underset{\sim}{\Sigma}_{11} & \underset{\sim}{\Sigma}_{12} \\ \underset{\sim}{\Sigma}_{21} & \underset{\sim}{\Sigma}_{22} \end{pmatrix} \qquad \underset{\sim}{\Lambda} = \underset{\sim}{\Sigma}^{-1} = \begin{pmatrix} \underset{\sim}{\Lambda}_{11} & \underset{\sim}{\Lambda}_{12} \\ \underset{\sim}{\Lambda}_{21} & \underset{\sim}{\Lambda}_{22} \end{pmatrix}$$

Then:

the marginals are

$$p(\underset{\sim}{x}_1) = N(\underset{\sim}{x}_1 \mid \underset{\sim}{\mu}_1, \underset{\sim}{\Sigma}_{11})$$

$$p(\underset{\sim}{x}_2) = N(\underset{\sim}{x}_2 \mid \underset{\sim}{\mu}_2, \underset{\sim}{\Sigma}_{22})$$

the conditionals are

$$p(\underset{\sim}{x}_1 \mid \underset{\sim}{x}_2) = N(\underset{\sim}{x}_1 \mid \underset{\sim}{\mu}_{1|2}, \underset{\sim}{\Sigma}_{1|2})$$

where

$$\mu_{1|2} = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$$
$$= \mu_1 - \Lambda_{11}^{-1}\Lambda_{12}(x_2 - \mu_2)$$

$$\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$
$$= \Lambda_{11}^{-1}$$

So if **Multivariate Gaussians on $\mathbb{R}^n$**

$\underline{\underline{\Sigma}}$ pos def, $\underline{\underline{\Sigma}}^{-1}$ also pos def.

$\underline{v}^T \underline{\underline{\Sigma}}\, \underline{v} > 0$

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \sim N\left(\underbrace{\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \vdots \\ \mu_n \end{bmatrix}}_{\substack{\text{mean} \\ \underset{\sim}{\mu} \in \mathbb{R}^n}}, \underbrace{\begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \cdots & \Sigma_{2n} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \cdots & \Sigma_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ \Sigma_{n1} & \Sigma_{n2} & \Sigma_{n3} & \cdots & \Sigma_{nn} \end{bmatrix}}_{\substack{\text{covariance } \underline{\underline{\Sigma}} \in \mathbb{R}^n \times \mathbb{R}^n \\ \text{positive definite}}} \right)$$

joint PDF $\quad p(\underset{\sim}{x}) = \dfrac{1}{\sqrt{(2\pi)^n \det(\underline{\underline{\Sigma}})}} \exp\left\{ -\frac{1}{2} \underbrace{(\underset{\sim}{x} - \underset{\sim}{\mu})^T \underline{\underline{\Sigma}}^{-1} (\underset{\sim}{x} - \underline{\mu})} \right\}$

⊙ closure under linear transf:

$\underset{\sim}{x} \sim N(\underset{\sim}{\mu}, \underline{\underline{\Sigma}}) \qquad \underline{\underline{A}}\,\underline{x} + \underline{b} \sim N(\underline{\underline{A}}\,\underline{\mu} + \underline{b},\ \underline{\underline{A}}\,\underline{\underline{\Sigma}}\,\underline{\underline{A}}^T)$

So if **Multivariate Gaussians** on $\mathbb{R}^n$

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \sim N\left(\underbrace{\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \vdots \\ \mu_N \end{bmatrix}}_{\text{mean}}, \underbrace{\begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \cdots & \Sigma_{2n} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \cdots & \Sigma_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ \Sigma_{n1} & \Sigma_{n2} & \Sigma_{n3} & \cdots & \Sigma_{nn} \end{bmatrix}}_{\text{covariance}} \right)$$

mean

$\underset{\sim}{\mu} \in \mathbb{R}^n$

covariance

$\underset{\sim}{\Sigma} \in \mathbb{R}^n \times \mathbb{R}^n$ pos. def.

joint PDF $$p(\underset{\sim}{x}) = \frac{1}{\sqrt{(2\pi)^n \det(\underset{\sim}{\Sigma})}} \exp\left\{ -\frac{1}{2} (\underset{\sim}{x} - \underset{\sim}{\mu})^T \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{x} - \underset{\sim}{\mu}) \right\}$$

⊙ addition & subtraction : $\underset{\sim}{X} \sim N(\underset{\sim}{\mu}_X, \Sigma_X)$ $\quad \underset{\sim}{Y} \sim N(\underset{\sim}{\mu}_Y, \Sigma_Y)$

$$\underset{\sim}{X} \pm \underset{\sim}{Y} \sim N(\underset{\sim}{\mu}_X \pm \underset{\sim}{\mu}_Y, \Sigma_X + \Sigma_Y)$$

So if Multivariate Gaussians on $\mathbb{R}^n$ MVGT

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \vdots \\ \mu_N \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \cdots & \Sigma_{2n} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \cdots & \Sigma_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ \Sigma_{n1} & \Sigma_{n2} & \Sigma_{n3} & \cdots & \Sigma_{nn} \end{bmatrix} \right)$$

marginals are easy

$$p(x_1) = \int \cdots \int p(x_1, \dots, x_n)\, dx_2 \dots dx_n = N(\mu_1, \Sigma_{11})$$

So if Multivariate Gaussians on $\mathbb{R}^n$

$$\underset{\sim}{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \sim N\left(\overset{\mu_A}{\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \vdots \\ \mu_n \end{bmatrix}} , \overset{\Sigma_A}{\begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \cdots & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \cdots & \Sigma_{2n} \\ \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \cdots & \Sigma_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ \Sigma_{n1} & \Sigma_{n2} & \Sigma_{n3} & \cdots & \Sigma_{nn} \end{bmatrix}} \right)$$

marginals are easy : of sub-vector (x_1, x_2)

$$p(x_1, x_2) = \int \cdots \int p(x_1, x_2, \ldots, x_n)\, dx_3 \ldots dx_n = N(\mu_A, \Sigma_A)$$

So if

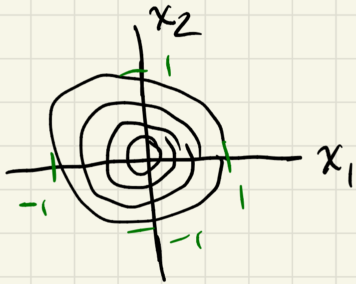
$$\underset{\sim}{x} = \left[\begin{array}{c} x_1 \\ x_2 \\ \hline x_3 \\ \vdots \\ x_n \end{array}\right] \sim N\left(\left[\begin{array}{c} \overbrace{\begin{matrix}\mu_1 \\ \mu_2\end{matrix}}^{\mu_A} \\ \hline \underbrace{\begin{matrix}\mu_3 \\ \vdots \\ \mu_N\end{matrix}}_{\mu_B} \end{array}\right], \left[\begin{array}{cc|cccc} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} & \cdots & & \Sigma_{1n} \\ \Sigma_{21} & \Sigma_{22} & \Sigma_{23} & \cdots & & \Sigma_{2n} \\ \hline \Sigma_{31} & \Sigma_{32} & \Sigma_{33} & \cdots & & \Sigma_{3n} \\ \vdots & \vdots & \vdots & & & \vdots \\ \Sigma_{n1} & \Sigma_{n2} & \Sigma_{n3} & \cdots & & \Sigma_{nn} \end{array}\right] \right)$$

$$\begin{array}{cc} \Sigma_A & \Sigma_{AB} \\ \Sigma_{BA} & \Sigma_B \end{array}$$

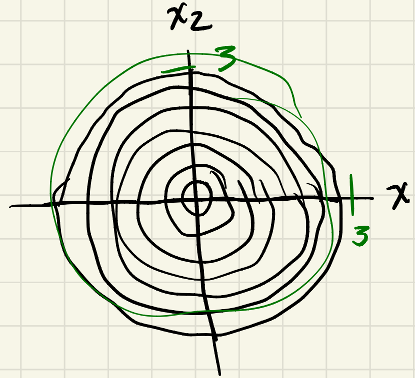
conditionals are also not so bad

$$p(x_1, x_2 \mid \underbrace{x_3, \ldots, x_N}_{y}) = N\left(\mu_A + \Sigma_{AB}\Sigma_B^{-1}(y - \mu_B),\ \Sigma_A - \Sigma_{BA}\Sigma_B^{-1}\Sigma_{AB}\right)$$

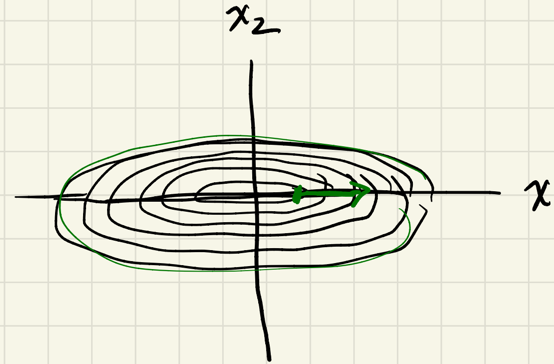
Back to bivariate case



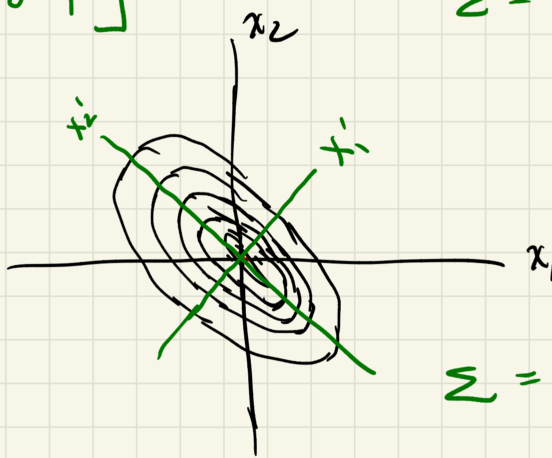
$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



$$\Sigma = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}$$



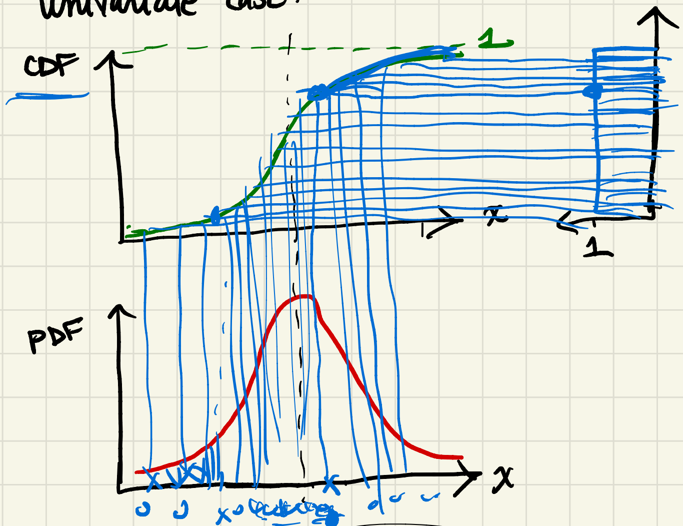
$$\Sigma = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$$



$$\Sigma = \begin{bmatrix} 1 & -0.6 \\ -0.6 & 1 \end{bmatrix}$$

How do we draw samples from an arbitrary multivariate Gaussian?

univariate case:



$$\boxed{X \sim N(0, 1)}$$

univariate $\quad x' = N(\mu, \sigma^2)$

$$X' = \mu + \sigma X$$

$$= \boxed{\mu + \sigma N(0, 1)}$$

$\sigma \rightarrow$ square root of variance

Multivariate

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

draw two gaussians, $N(0,1)$

$$X' \sim N\left(\underbrace{\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}}_{\mu}, \underbrace{\begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}}_{\Sigma}\right)$$

$$X' = \mu + L\, N(0, I)$$

Cholesky decomposition $\quad \Sigma = LL^T$