Lecture 6:

pdf of $X \in \mathbb{R}^n$ is a func $f: \mathbb{R}^n \to \mathbb{R}_{\geq 0}$

$$\int_{\mathbb{R}^n} f(x)\,dx = 1$$

$R \subseteq \mathbb{R}^n$; $\mathbb{P}(X \in R) = \int_R f(x)\,dx$

$$R = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_n, b_n]$$

$$\{(x_1, \ldots, x_n) : a_i \leq x_i \leq b_i \text{ for } i = 1, \ldots, n\}$$

$$\int_R f = \int_{a_n}^{b_n} \ldots \int_{a_1}^{b_1} f(x_1, \ldots, x_n)\,dx_1 \ldots dx_n$$

Case 1: indep. case

$(x_1, \ldots, x_n$ are indep$)$

$$f(x_1, \ldots, x_n) = f_1(x_1) \ldots f_n(x_n) = \prod_{i=1}^{n} f_i(x_i)$$

f_i is the pdf of X_i

esp if X_i are i.i.d, $f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_n(x_i)$

$\to f(x)$

Conditional Density:

$X = (X_1, \dots, X_k)$ with pdf f

conditional pdf of X_k
given $X_1 = x_1, \dots, X_{k-1} = x_{k-1}$

is $f(x_k \mid X_1 = x_1, \dots, X_{k-1} = x_{k-1})$

$$= \frac{f(x_1, \dots, x_{k-1}, x_k)}{\int_{\mathbb{R}} f(x_1, \dots, x_{k-1}, x_k)\,dx_k}$$

↓

Conditional density func

$(x_1, \dots, x_{k-1}) \mapsto$ (cond) pdf

$f(\cdot \mid X_1 = x_1, \dots, X_{k-1} = x_k)$

Marginalization:

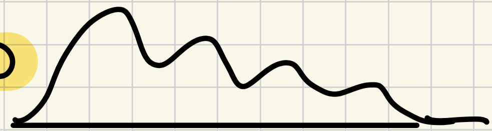
$$f_1(x_1) = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f(x_1, \dots, x_k)\,dx_2\,dx_3 \dots dx_k$$

↓

marginal density func

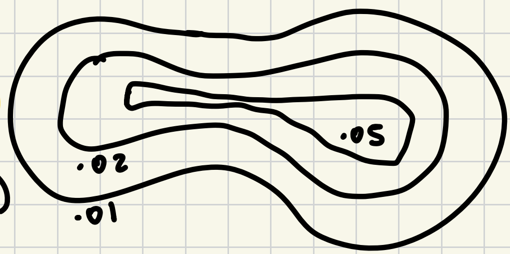
1-D

(x_1)



2-D

(x_1, x_2)



3-D

(x_1, x_2, x_3)

	1	2	3
1	X		
2		X	
3			X

Case 2: Gaussian Case

Call $X \in \mathbb{R}^k$ a Gaussian if it's pdf is

$$f(x) = \frac{1}{(2\pi)^{k/2} \det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right)$$

here $\mu \in \mathbb{R}^k$ is mean vector

and $\Sigma \in \mathbb{R}^{k \times k}$ in p.d. matrix and

Σ: covariance mtx

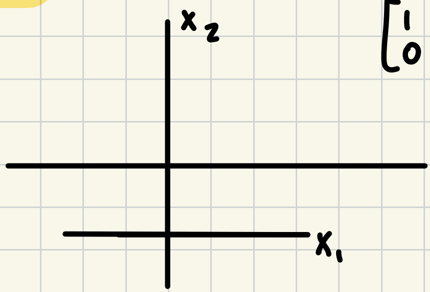
If A is a matrix, b is a vector

$Ax + b$ is still a gaussian. → normal dist.

$$X \sim N(\underbrace{A}_{m \times k}\mu + b, \underbrace{A}_{m \times k} \underbrace{\Sigma}_{k \times k} \underbrace{A^T}_{k \times m})$$

$X \in \mathbb{R}^k$, A is $m \times k$, $b \in \mathbb{R}^m$

$Ax + b \in \mathbb{R}^m$



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

Standarization:

$X \sim N(\mu, \Sigma)$

$X - \mu \sim N(0, \Sigma)$

$N(0, I_k)$

Σ is p.d.

→ U is orthogonal $(UU^T = I_k)$

Λ is diagonall

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_k \end{bmatrix}$$

$\Sigma^{1/2} = U \Lambda^{1/2} U^T$

$\Sigma^{1/2} \Sigma^{1/2} = \Sigma$

|

$\downarrow$

$$\left(\begin{array}{l}\text{another way:}\\ \Sigma^{1/2} = U\Lambda^{1/2} \\ \Sigma^{1/2}(\Sigma^{1/2})^T = \Sigma\end{array}\right)$$

take $\Sigma^{-1/2} = (\Sigma^{-1})^{1/2} = (U\Lambda^{-1}U^T)^{1/2}$

$$\Sigma^{-1/2}(X-\mu) \sim N\left(0, \Sigma^{-1/2}\Sigma(\Sigma^{-1/2})^T\right)$$

$$\underbrace{U\Lambda^{-1/2}U^T}\,\underbrace{\cancel{U}\Lambda U^T}\,\underbrace{\cancel{U}\Lambda^{-1/2}U^T}$$

I_K

$U I_n U^T = I_n$

$X_1, X_2, X_3, \ldots$ iid w/ mean μ and cov. Σ.

each X_i is a rand vec. in $\mathbb{R}^n$

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \sqrt{n}\,(\overline{X}_n - \mu) \rightsquigarrow N(0, \Sigma)$$

as $n \to \infty$.