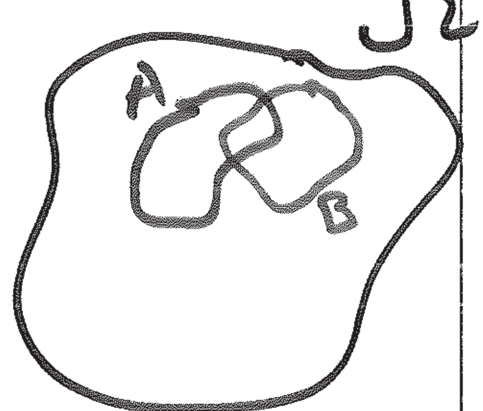
Rev. of prob + stats

Ω : set of all events

Pr : assigns a scalar likelihood (prob) to each event in Ω

event $A \subseteq \Omega$



$0 \leq Pr(A) \leq 1$

$Pr(\Omega) = 1$

$$Pr(A \text{ or } B) = Pr(A \cup B)$$

$$= Pr(A) + Pr(B) - \underbrace{Pr(A \cap B)}_{=0}$$

iff A, B mut. excl.

$$Pr(A \text{ and } B) = Pr(A \cap B)$$

$$= Pr(A) \cdot Pr(B) \text{ independence.}$$

Conditional Prob.

$$Pr(A \text{ given } B) = Pr(A \mid B)$$

$$= \frac{Pr(A \cap B)}{Pr(B)}$$

$$P(B \mid A) = \frac{Pr(A \cap B)}{Pr(A)}$$

$$Pr(B \mid A) \cdot Pr(A) = Pr(A \mid B) \cdot Pr(B)$$

$$Pr(A \mid B) = \frac{Pr(B \mid A) \cdot Pr(A)}{Pr(B)}$$

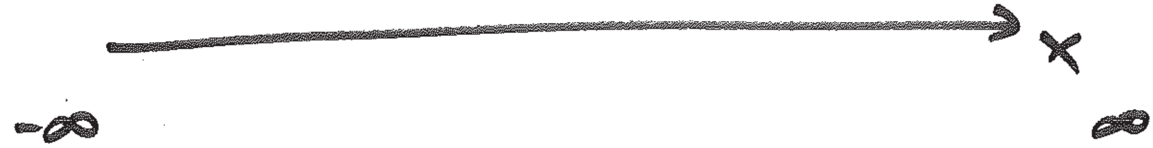
If A, B indep.

$$Pr(A \mid B) = Pr(A)$$

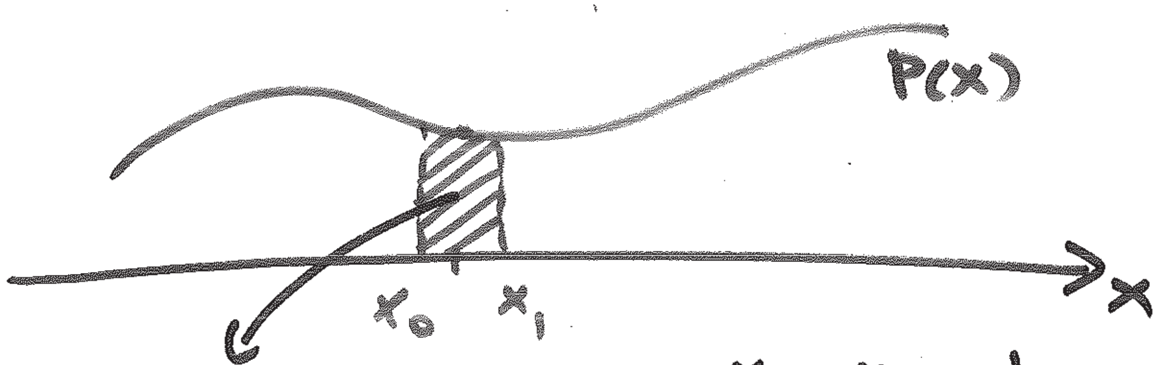
$$Pr(B \mid A) = Pr(B)$$

Cont. Rand. variables

X cont. val. in $\mathbb{R}$



$p(x)$ prob. density fncn.



$x_1 - x_0 = dx$

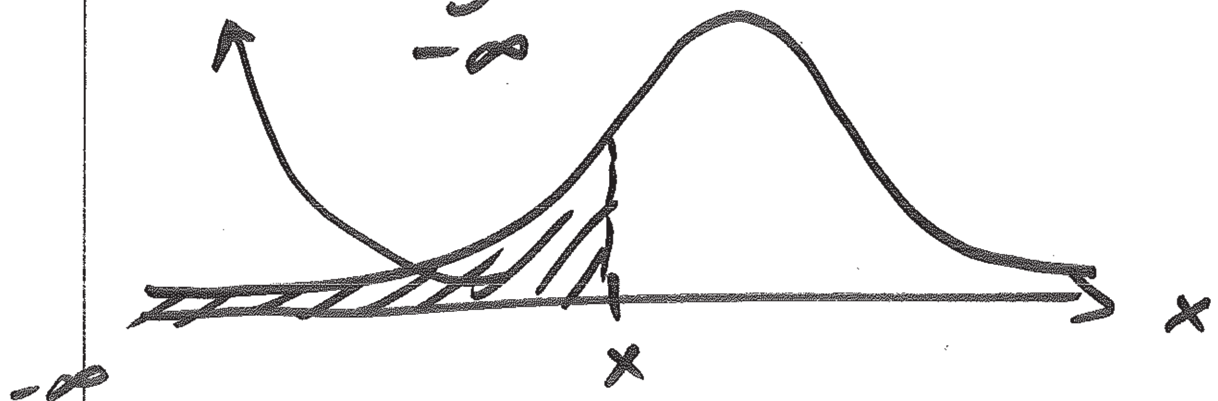
$Pr(x_0 \le x \le x_1)$

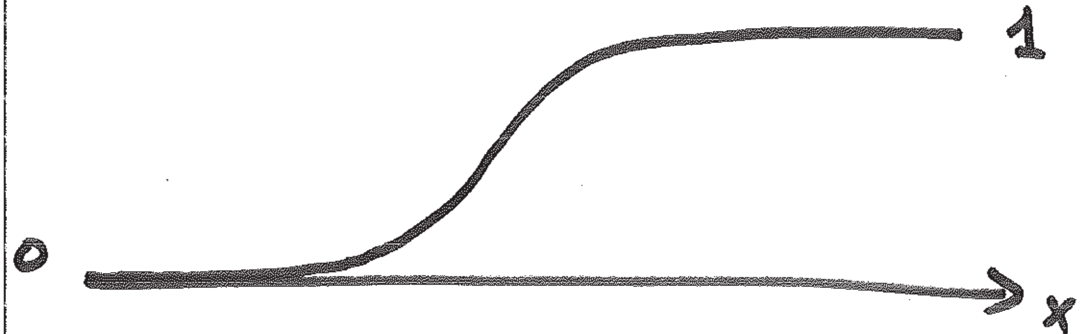
$$\approx p(x_0) \cdot dx = Pr(x_0 \le x \le x_0 + dx)$$

$$\int_{-\infty}^{\infty} p(x)\, dx = 1 \quad \underline{\text{and}} \quad p(x) \ge 0$$

Cumulative Distr. fncn: (CDF)

$$P(x) = \int_{-\infty}^{x} p(\tau)\, d\tau$$





$$\underbrace{\frac{d}{dx} P(x)}_{\text{CDF}} = \underbrace{p(x)}_{\text{PDF}}$$

Discrete Random variables:

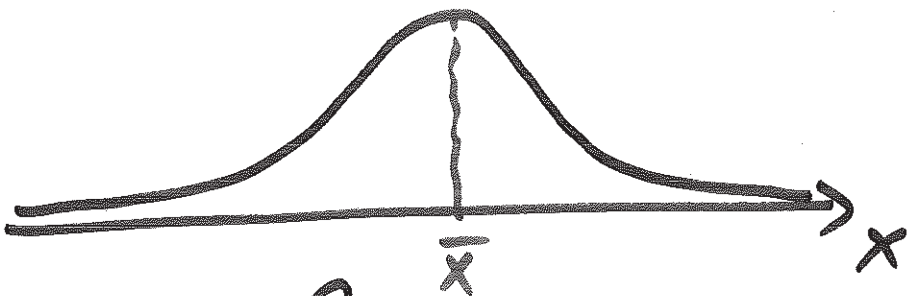
$x \in \{x_i\}_{i=1}^{M}$ $\quad Pr(x = x_i) = p_i$

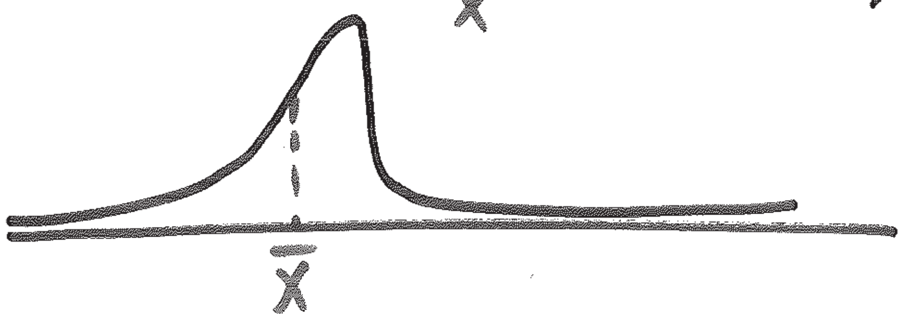
$p_i \geq 0$

$p_i \longrightarrow \sum_{i=1}^{M} p_i = 1$

Moments

$$E(x) = \int x\,p(x)\,dx = \bar{x},\ \mu_x,\ m$$





$$E(f(x)) = \int f(x)\,p(x)\,dx$$

$$E(x^2) = \int x^2\,p(x)\,dx$$: mean-sq. value of x

$$\text{var} = \int (x-m)^2\,p(x)\,dx = \sigma^2$$

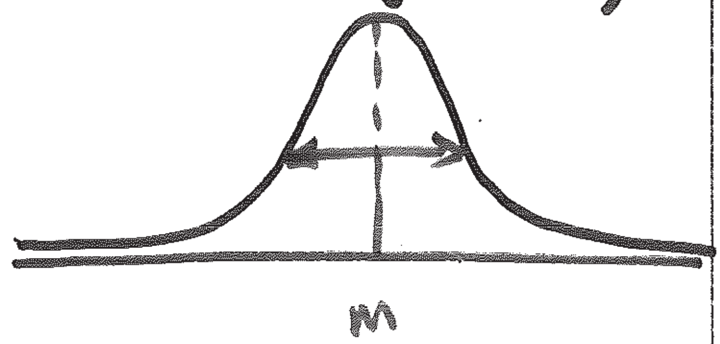
$$= E(x^2) - m^2$$

σ : standard deviation

Gaussian r.v. (normal r.v.)

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

$x \sim \mathcal{N}(m, \sigma^2)$



Two r.v.'s a, b

$$E(\alpha a + \beta b) = \alpha E(a) + \beta E(b)$$

$$\mathrm{Var}(\alpha a + \beta b) = \alpha^2 \mathrm{var}(a) + \beta^2 \mathrm{var}(b) + 2\alpha\beta\,\mathrm{cov}(a,b)$$

$$\mathrm{Cov}(a,b) = E\left((a - \bar{a})(b - \bar{b})\right)$$

Two r.v. x_1, x_2 : $\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

Joint PDF

$$p(x_1, x_2) = p(\underline{x})$$

$$p(x_1, x_2) \longrightarrow p(x_1) = \int p(x_1, x_2)\, dx_2$$

$$p(x_2) = \int p(x_1, x_2)\, dx_1$$

marginals

$$p(x_1, x_2) = p(x_1 | x_2) \cdot p(x_2)$$

$$= p(x_2 | x_1) \cdot p(x_1)$$

$$p(x_1 | x_2) = \frac{p(x_2 | x_1)\, p(x_1)}{p(x_2)}$$

<u>corr. coeff:</u>

$$\rho = \frac{\mathrm{cov}(x_1, x_2)}{\sqrt{\mathrm{var}(x_1) \cdot \mathrm{var}(x_2)}} = \frac{\mathrm{cov}(x_1, x_2)}{\sigma_1 \cdot \sigma_2}$$

$$-1 \leq \rho \leq 1$$

$$\rho = 1 \Rightarrow x_2 = \alpha x_1 + \beta$$

$$\alpha > 0$$

$\underline{X} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ $\quad Cov(\underline{X}) = \begin{bmatrix} var(x_1) & cov(x_1, x_2) \\ cov(x_1, x_2) & var(x_2) \end{bmatrix}$

$$Cov(\underline{X}) = E\left[(X-m)(X-m)^T\right]$$

$$C_X = Cov(\underline{X}) = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

uncorrelated $\rho = 0$ $\quad Cov(X) = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$

$$\boxed{C_X = C_X^T}$$

uncorrelated

$$E\left[(x_1 - \bar{x}_1)(x_2 - \bar{x}_2)\right] = 0$$

$$\Rightarrow \quad E(x_1 x_2) = E(x_1)E(x_2)$$

Indep. $\quad P(x_1, x_2) = p(x_1) \cdot p(x_2)$

Indep $\Longrightarrow$ Uncorrelated

$\not\Longleftarrow$

$$\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix} \qquad E(X) = \begin{bmatrix} m_1 \\ \vdots \\ m_N \end{bmatrix} = \underline{m}$$

$$\mathrm{Cov}(x) = \underbrace{C_x}_{N \times N} = E\left[(x-m)(x-m)^T\right]$$

$$= \begin{array}{c|c} & j \\ \hline i & \begin{bmatrix} \sigma_1^2 & & & & \\ & \sigma_2^2 & & & \\ & & \ddots & & \\ & & * & & \\ & & & & \sigma_N^2 \end{bmatrix} \end{array} \rightarrow \mathrm{Cov}(x_i, x_j)$$

$$x \sim N(\underline{m}, C_x)$$

$$P(\underline{x}) = \frac{1}{(2\pi)^{N/2}\sqrt{\det(C_x)}} \exp\left(-\frac{1}{2}(x-m)^T C^{-1} (x-m)\right)$$

$$C = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \ldots \sigma_N^2)$$

$$= \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_N^2 \end{bmatrix}$$

$$\exp\Big(\underbrace{-\frac{1}{2}(x-m)^T C^{-1}(x-m)}\Big) =$$

$$-\frac{1}{2}\begin{bmatrix} x_1 - m_1, & \ldots & x_N - m_N \end{bmatrix} \begin{bmatrix} \sigma_1^{-2} & & \\ & \ddots & \\ & & \sigma_N^{-2} \end{bmatrix} \begin{bmatrix} x_1 - m_1 \\ \vdots \\ x_N - m_N \end{bmatrix}$$

$$= -\frac{1}{2}\frac{(x_1-m_1)^2}{\sigma_1^2} - \frac{1}{2}\frac{(x_2-m_2)^2}{\sigma_2^2} + \cdots - \frac{1}{2}\frac{(x_N-m_N)^2}{\sigma_N^2}$$

$$= \exp\left(-\frac{1}{2}\frac{(x_1-m_1)^2}{\sigma_1^2}\right) \cdots \exp\left(-\frac{1}{2}\frac{(x_N-m_N)^2}{\sigma_N^2}\right)$$

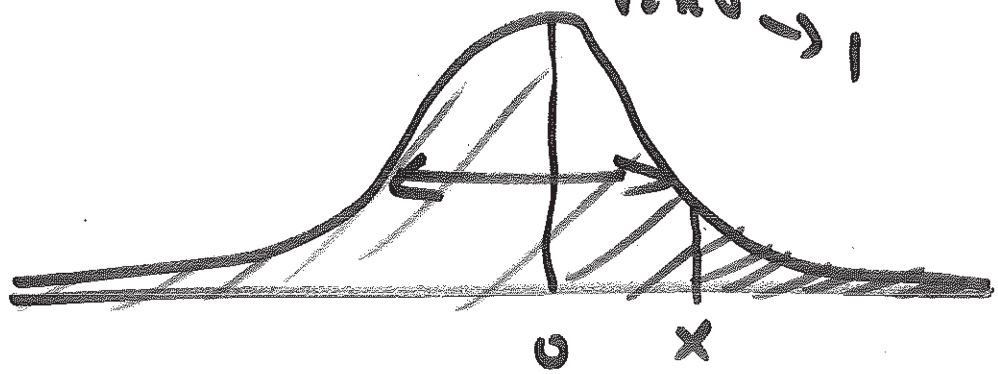
$$(2\pi)^{N/2}\sqrt{\det(C_x)} = \underbrace{\sqrt{2\pi}\cdot\sqrt{2\pi}\cdots\sqrt{2\pi}}_{N}\cdot\sigma_1\cdot\sigma_2\cdots\sigma_N$$

$$p(x_1, x_2, \ldots, x_N) = p(x_1)\cdot p(x_2)\cdots p(x_N)$$

indep.

A little digression to $\mathcal{N}(0,\sigma^2)$

$\mathcal{N}(0,1)$ $\quad p(x) = \frac{1}{\sqrt{2\pi}\,\underbrace{\sigma}_{\to 1}} \exp\left(\frac{x^2}{-2\underbrace{\sigma^2}_{\to 1}}\right)$



Cumulative dist. fncn.

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$$

$$Q(x) = 1 - \Phi(x)$$

$$Q(x) \approx \frac{1}{\sqrt{2\pi}\, x} e^{-x^2/2} \qquad x > 4$$

Lec. 3
W '12

chi-squared distribution

Given $x(n) \sim N(0,1)$ i.i.d.

$\chi = \sum_{n=1}^{N} x^2(n)$ $\quad \chi_N^2$: chi-sq. r.v

$\downarrow$ $\quad$ N degrees

central chi-sq. dist. of free

$$E(\chi) = N$$

$$Var(\chi) = 2N$$

$$p(x) = \frac{1}{2^{N/2}\Gamma(\frac{N}{2})} x^{\frac{N}{2}-1} e^{-x/2}$$

$\Gamma(n) = (n-1)!$

$$N = 2 \Rightarrow p(x) = \frac{1}{2} e^{-\frac{x}{2}}$$

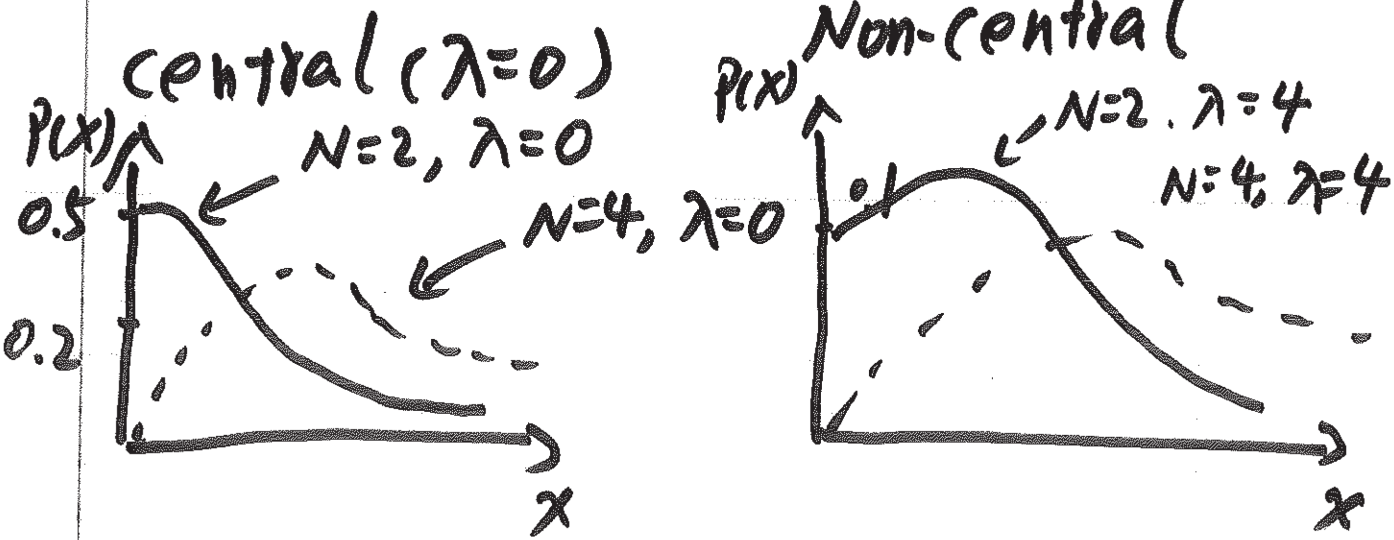
$\Rightarrow$ exp. r.v

if $x(n) \sim N(m_n, 1)$

$\chi = \sum_{n=1}^{N} x^2(n)$: Non-central chi-sq.

$\chi_N^2(\lambda)$

$$\lambda = \sum_{n=1}^{N} m_n^2 \quad \begin{cases} E(x) = N + \lambda \\ Var(x) = 2N + 4\lambda \end{cases}$$



F - density

Given $x_1 \sim \chi^2_{N_1}(0)$

$x_2 \sim \chi^2_{N_2}(0)$

$$\frac{x_1/N_1}{x_2/N_2} \sim F_{N_1, N_2}$$

test statistic

$$T = \frac{1}{N}\sum_{n=1}^{N} x(n)$$

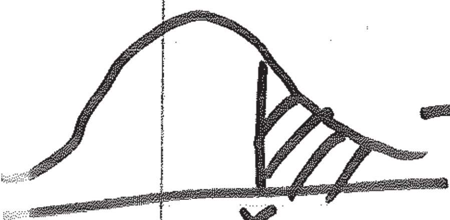
$x(n) \sim N(0, \sigma^2)$ i.i.d.

$P_r(T > \gamma)$

$$\Pr(T>\gamma) = \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}\,\sigma/\sqrt{N}} e^{-\frac{x^2}{2\sigma^2/N}}\,dx$$

$$x' = \frac{x}{\sqrt{\sigma^2/N}}$$

$$= \int_{\frac{\gamma}{\sqrt{\sigma^2/N}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x'^2}{2}}\,dx'$$



$\frac{\gamma}{\sqrt{\sigma^2/N}}$

$$= Q\left(\frac{\gamma}{\sqrt{\sigma^2/N}}\right)$$

Estimate $\Pr(T>\gamma)$ using simulation (Monte-Carlo)

1. Generate $\{X(n)\}_{n=1}^{N}$ (randn $N(0,\sigma^2)$)
2. compute $T = \frac{1}{N}\sum x(n)$
3. repeat ①→② ⇒ Get $T_1 \cdots T_M$
4. look at how many $\# T_k > \gamma \rightarrow M_\gamma$
5. Estimate $\Pr(T>\gamma) \approx \frac{M_\gamma}{M} = \hat{P}$

Review of Lin. Algebra

$$\underline{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} \qquad \underline{x}^T = [x_1, \cdots x_N]$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1N} \\ \vdots & \ddots & & \\ a_{M1} & & \ddots & a_{MN} \end{bmatrix}$$

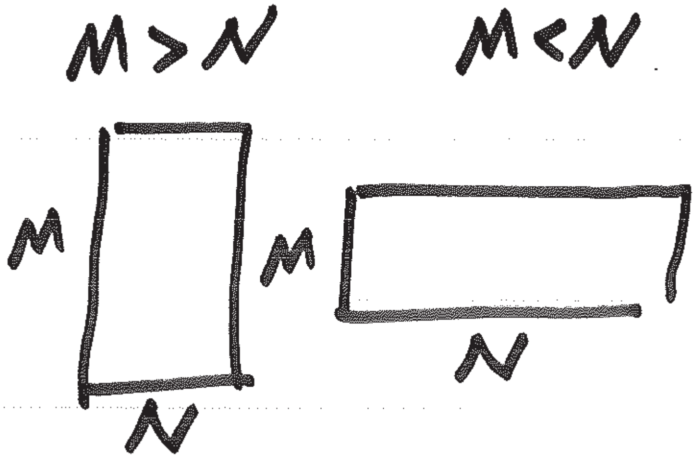
$M \times N$

$$A = [\underline{a}_1 \,|\, \underline{a}_2 \,|\, \cdots \ \underline{a}_N]$$

$M = N$

A: square

A: $M \neq N$



rank(A) = # of linear indep. rows or columns (whichever is less)

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \text{rank} = 2 \qquad \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 4 \\ 3 & 6 \end{bmatrix} \text{rank} = 1$$

$AA^{-1} = A^{-1}A = I$ square invertible mat.

$A^T = A$ symmetric

A positive semi-definite

if $\forall x \neq 0$ $\quad x^T A x \geq 0$

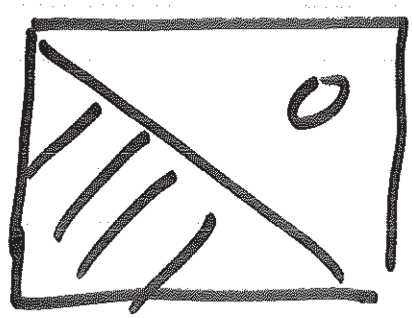
$\downarrow$

$x^T A x > 0$

positive definite mat.
existence of a real square root mat.

$A = CC^T$

C



Cholesky

~~sqrtcm~~ sqrtm(A)

sqrt(A) → element by element

$A^{-1} = A^T \Leftrightarrow A^TA = AA^T = I$

orthognal (unitary) mat.

$A = [\underline{a}_1 | \underline{a}_2 \cdots \underline{a}_N]$

$$\underline{a}_i^T \underline{a}_j = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$A : N \times N$

$A^2 = A \Rightarrow AA^2 = A^2 = A^3 = A$

$\cdots \quad A^K = A$ for all $K \geq 2$

idempotence

$(AB)^T = B^T A^T$

$\det(A^T) = \det(A)$

$\det(A^{-1}) = \dfrac{1}{\det(A)}$

$\det(AB) = \det(A) \cdot \det(B)$

$tr(A) = \sum_{i=1}^{N} a_{ii}$

$tr(A \cdot B) = tr(B \cdot A)$

$tr(\underline{x}\,\underline{y}^T) = tr(\underline{y}^T \underline{x}) = \underline{y}^T \underline{x}$

scalar

inner product.

scalar function of vector var.

$$f(\underline{x}) = f\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}\right) \quad f: \mathbb{R}^N \to \mathbb{R}$$

Grad. of $f = \nabla f = \frac{\partial f}{\partial \underline{x}} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_N} \end{bmatrix}$

$f(x) = b^T x \Rightarrow \nabla f = b$

$f(x) = x^T A x$

$\nabla f = 2Ax$ if A sym.

$\nabla f = (A^T + A)x$ asym.

Eigen decomp.

$$A: N \times N$$

$$\underbrace{A \underline{v}}_{\text{eigen vector}} = \underbrace{\lambda}_{\text{scalar / eigenvalue}} \underline{v}$$

$\underline{v}^T \underline{v} = \|\underline{v}\|^2 = 1$

in general A is not necessarily symm.

$$A = SDS^{-1} \qquad S: \text{orthognal mat.}$$
$$= SDS^T$$

if A is symm.

$$A = VDV^T \qquad V^TV = VV^T = I$$

$$D = \mathrm{diag}(\lambda_1 \cdots \lambda_N)$$

λ_i's real

if $\lambda_i > 0$ positive definite

$$V = [\underline{v}_1, \underline{v}_2, \cdots \underline{v}_N]$$

$$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{bmatrix}$$

$$A = VDV^T = [\underline{v}_1 \cdots \underline{v}_N] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_N^T \end{bmatrix}$$

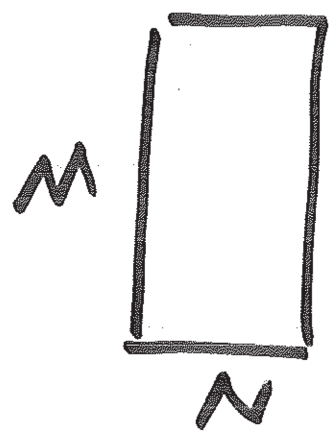
$$A = \lambda_1 \underline{v}_1 \underline{v}_1^T + \lambda_2 \underline{v}_2 \underline{v}_2^T \cdots + \lambda_n \underline{v}_n \underline{v}_n^T$$

$$= \sum_{i=1}^{N} \lambda_i (\underline{v}_i \underline{v}_i^T)$$

$[V, D] = eig(A)$

Singular value Decom.
(SVD)

$A : M \times N$ (real)



$M > N$

$A = U S V^T$

U: $M \times M$, $\overline{\text{orth.}}$

S: $M \times N$

V^T: $N \times N$, $\overline{\text{orth.}}$

$U U^T = U^T U = I_{M \times M}$

$V V^T = V^T V = I_{N \times N}$

$$S = \left[\begin{array}{c} \begin{matrix} s_1 & & 0 \\ & s_2 \ddots & \\ 0 & & s_N \end{matrix} \\ \hline O \end{array}\right]$$

$s_1 \geq s_2 \geq \cdots \geq s_N \geq 0$

singular values

$$A = USV^T \qquad A^TA = VS^T\underbrace{U^TU}_{I}SV^T$$

$$= V\underbrace{S^TS}_{\begin{bmatrix} s_1^2 & & & 0 \\ & s_2^2 & & \\ & & \ddots & \\ 0 & & & s_N^2 \end{bmatrix}}V^T$$

s_i^2 are eig. of $\dfrac{A^TA}{\lambda_i}$

$$= VDV^T$$

$$s.v(A) = \sqrt{\text{eig}(A^TA)}$$

$$[U, S, V] = \text{svd}(A)$$

condition number

$$A \doteq SVD \Rightarrow \underbrace{s_1}_{\text{largest}} \geq s_2 \cdots \underbrace{s_N}_{\text{smallest}} \geq 0$$

$$K(A) = \frac{s_1}{s_N} \qquad \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix}$$

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ well-conditioned $K(A) \to 1$

ill-conditioned $K(A) \to \infty$

Given r. vector $\underline{x}$ $\quad E(\underline{x}) = 0$

$Cov(\underline{x}) = C_x$

$$y = Ax + b \rightsquigarrow E(y) = ?$$

$$y - b = Ax \qquad Cov(y) = ?$$

$$
\begin{aligned}
E(y) &= E(Ax + b) \\
&= E(Ax) + b \\
&= A\underbrace{E(x)}_{=0} + b \\
&= b
\end{aligned}
$$

$$
\begin{aligned}
Cov(y) &= E\left((y-b)(y-b)^T\right) \\
&= E\left(Ax\,(Ax)^T\right) \\
&= E(Ax\,x^T A^T) \\
&= A\,E(xx^T)\,A^T \\
&= A C_x A^T
\end{aligned}
$$

$C_x = V D V^T$

$C_y = A V D V^T A^T$

Let $A = V^T$

$$C_y = \underbrace{V^T V}_{I} D \underbrace{V^T V}_{I} = D$$

"whitening"

$x \sim N(0, \underbrace{I}_{100 \times 100})$

$$x = \text{randn}(100, 1) = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \quad N = 100$$

y s.t. $E(y) = b$

$Cov(y) = C$

$y = Ax + b \quad E(y) = b$

$$\mathrm{Cov}(y) = A \underbrace{\mathrm{Cov}(x)}_{I} A^T$$

$$= AA^T = C$$

$$C = V\underline{D}V^T = V D^{\frac{1}{2}} D^{\frac{1}{2}} V^T$$

$$= \underbrace{V D^{\frac{1}{2}}}_{A} \underbrace{(D^{\frac{1}{2}})^T V^T}_{A^T}$$

sqrtm(C)