26 Aug Class Notes:

Cor (Cholesky Factorization): Let A be a symmetric and positive semidefinite matrix in $M_n(\mathbb{R})$. Then there is a lower triangular matrix L such that $A = LL^T$.
(If A is positive definite, the choice of L is unique upto scaling by diag $(\pm 1, \pm 1, \ldots)$)

Pf: Since A is p.s.d., $\exists$ a unique $\sqrt{A}$.
Let $\sqrt{A} = LQ = Q^T L^T$ (by symmetry and Q-R factorization)

$$A = (\sqrt{A})^2 = LQQ^T L^T = LL^T$$

This shows existence.

Uniqueness: Let there be 2 decompositions: $L_1L_1^T$ and $L_2L_2^T$.

$$L_1L_1^T = L_2L_2^T$$
$$\Rightarrow L_2^{-1}L_1 = L_2^T(L_1^{-1})^T$$
$$\Rightarrow (L_1^{-1}L_2)^{-1} = (L_1^{-1}L_2)^T$$

This is lower triangular + orthogonal.

$$\Rightarrow L_1^{-1}L_2 = \text{diag}(\pm 1, \pm 1, \ldots)$$
$$\Rightarrow L_2 = L_1 \,\text{diag}(\pm 1, \pm 1, \ldots)$$ ◻

$T: \mathcal{H} \longrightarrow \mathcal{H}$
w.r.t. an orthonormal basis $\{e_i\}_{i \geq 1}$, $(T)_{ij} = \langle Te_i, e_j \rangle$
Knowing the entries of (T), can you decide if T is a bounded operator?

Cholesky decomposition has analogues in more general settings like Hardy spaces (z-transform gives H^2 spaces), "Jensen's inequalities in finite subdiagonal algebras" - by S. Nayak

Finding the Cholesky decomposition:

- Step 1:

$$A = \left[\begin{array}{c|c} \alpha_{11} & * \\ \hline a_{21} & A_{22} \end{array}\right] = \left[\begin{array}{c|c} \lambda_{11} & 0 \\ \hline l_{21} & L_{22} \end{array}\right] \left[\begin{array}{c|c} \lambda_{11} & 0 \\ \hline l_{21} & L_{22} \end{array}\right]^T$$

$$= \left[\begin{array}{c|c} \lambda_{11}^2 & * \\ \lambda_{11} l_{21} & l_{21} l_{21}^T + L_{22} L_{22}^T \end{array}\right].$$

$\therefore \alpha_{11} = \lambda_{11}^2 \ (\geq 0)$ (consistent with the fact that for a psd matrix, diagonal entries ≥ 0).

$\therefore \lambda_{11} := \sqrt{\alpha_{11}}$

$$l_{21} = \frac{1}{\lambda_{11}} a_{21} \in \mathbb{R}^{n-1}$$

If $\alpha_{11} = 0$, is there an incompatibility issue?
No, since $\alpha_{11} = 0 \Rightarrow a_{21} = \vec{0} \in \mathbb{R}^{n-1}$:

$$|(A)_{21}| = |\langle Ae_1, e_2\rangle|$$
$$= |\langle \sqrt{A}e_1, \sqrt{A}e_2\rangle| \overset{C\text{-}S}{\leq} \left(\langle \sqrt{A}e_1, \sqrt{A}e_1\rangle \langle \sqrt{A}e_2, \sqrt{A}e_2\rangle\right)^{1/2}$$
$$= \left(\langle Ae_1, e_1\rangle \langle Ae_2, e_2\rangle\right)^{1/2}$$
$$= \left((A)_{11} \cdot (A)_{22}\right)^{1/2}$$

- <u>Step 2</u>: $L_{22} = CHOL(A_{22} - l_{21} l_{21}^T)$
  But to proceed with this recursive Cholesky decomposition algorithm, we have to show that ~~A_{22} is indeed~~ $A_{22} - l_{21} l_{21}^T$ is indeed sym. and p.s.d.

<u>Lemma</u>: Let $A \in M_n(\mathbb{R})$ be sym., p.s.d. Then $A_{22} - l_{21} l_{21}^T$ is also sym. and p.s.d.

Pf: Take $\vec{x_1} \in \mathbb{R}^{n-1} \setminus \{\vec{0}\}$.

$$\vec{x} = \begin{bmatrix} -a_{21}^T \vec{x_1} / \alpha_{11} \\ \vec{x_1} \end{bmatrix} \in \mathbb{R}^n.$$

$$0 = \langle \vec{x}, A\vec{x}\rangle = \left\langle \begin{pmatrix} \alpha_0 \\ \hline \vec{x_1} \end{pmatrix}, \left[\begin{array}{c|c} \alpha_{11} & a_{21}^T \\ \hline a_{21} & A_{22} \end{array}\right] \begin{pmatrix} \alpha_0 \\ \hline \vec{x_1} \end{pmatrix} \right\rangle$$
$$= \alpha_{11}\alpha_0^2 + \alpha_0 (a_{21}^T \vec{x_1}) + (\vec{x_1}^T a_{21}) \alpha_0 + \langle \vec{x_1}, A_{22}\vec{x_1}\rangle$$
$$= -\frac{\langle \vec{a_{21}}, \vec{x_1}\rangle^2}{\alpha_{11}} + \langle \vec{x_1}, A_{22}\vec{x_1}\rangle \quad \blacksquare$$

Random Sampling:

Inverse Transform Sampling:

If X is a continuous random variable with distribution function F_X, then the random variable $F_X(X)$ has a uniform distribution on $[0,1]$.

$(\mathbb{P}(F_X(X) \leq a) = a)$.

If Y has a uniform distribution on $[0,1]$ and X has distribution F_X (strictly increasing), then the random variable $F_X^{-1}(Y)$ has the same distribution as X.

Problem: X - r.v. with distn F_X (strictly inc.). We want to generate values of X distributed according to its distn.

Step I: Generate a random number u from the std. uniform distn. on $[0,1]$.

Step II: Compute $F_X^{-1}(u)$.

Step III: Go to Step I.

Example: X - exponential distn.

$$F_X(x) = 1 - e^{-\lambda x}$$

$$\therefore F_X^{-1}(y) = \frac{-1}{\lambda} \ln(1-y)$$

Draw y_0 from $U \sim \text{Unif}(0,1)$ and return $-\frac{1}{\lambda}\ln(1-y)$.

Problem: $F_n = \{(x_1, \ldots, x_n) : x_i \geq 0, \sum_{i=1}^{n} x_i = 1\} \subseteq \mathbb{R}^n$.

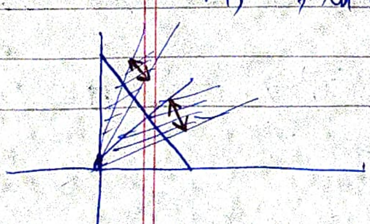
How do we sample uniformly from F_n?

$X_1 \sim \text{Exp}(\lambda), \ldots, X_n \sim \text{Exp}(\lambda)$.

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \lambda^n e^{-\lambda(x_1 + \cdots + x_n)}$$

Now we standardize x_i's:

$$x_1, \ldots, x_n \longmapsto \frac{x_1}{\sum_{i=1}^{n} x_i}, \ldots, \frac{x_n}{\sum_{i=1}^{n} x_i}.$$



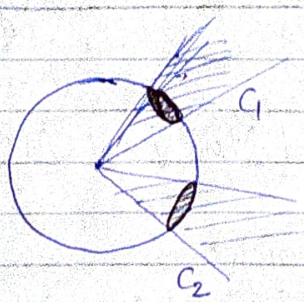
Problem 2: $\{(x_1, \dots, x_n) : x_1^2 + \dots + x_n^2 = 1\} = S^{n-1} \subseteq \mathbb{R}^n$.

$X_1 \sim N(0,1), \dots, X_n \sim N(0,1)$.

$x_1, \dots, x_n$

$\Downarrow$

$$\left(\frac{x_1}{\sqrt{\sum_{i=1}^n x_i^2}}, \dots, \frac{x_n}{\sqrt{\sum_{i=1}^n x_i^2}} \right) \in S^{n-1}$$



$\mathbb{P}((X_1, \dots, X_n) \in C_1)$
$= \mathbb{P}((X_1, \dots, X_n) \in C_2)$.

- Box Muller Transform:

Thm: Let U_1, U_2 be iid $U(0,1)$.
Let $Z_0 = \sqrt{-2 \ln U_1} \cos(2\pi U_2)$
$Z_1 = \sqrt{-2 \ln U_1} \sin(2\pi U_2)$
Then Z_0, Z_1 are iid $N(0,1)$.

Pf: $R^2 = -2 \ln U_1 \sim Exp(\frac{1}{2}) = \chi^2(2)$, $\theta = 2\pi U_2$

If X, Y iid $N(0,1)$
$R^2 = X^2 + Y^2 \sim \chi^2(2)$, $\theta = \tan^{-1}(Y/X)$
$X = R\cos\theta$, $Y = R\sin\theta$
Now use change of variables.

(Given Z_0, Z_1 iid $N(0,1)$, you can get joint normals Z_0', Z_1' with specified covariance matrix Σ by using Cholesky decomposition on $\Sigma = LL^T$. Then $L\begin{pmatrix} Z_0 \\ Z_1 \end{pmatrix} \sim \begin{pmatrix} Z_0' \\ Z_1' \end{pmatrix}$

Cholesky gives a quick decomposition of Σ into matrices of the form B, B^T)

Left-Invariant Haar measure on G: (locally compact groups)

Defn: A non-trivial Radon measure μ on G is said to be a l.i. Haar measure if $\mu(g \cdot E) = \mu(E) \ \forall E \in \sigma_G$ (σ_G: Borel σ algebra of G).

Method I: Step I: Choose a random vector $\vec{u_1} \in \mathbb{R}^n$ uniformly from S^{n-1}.

Step II: Choose a random vector $\vec{u_2} \in \mathbb{R}^n$ uniformly from $\{\vec{u_1}\}^{\perp} \cap S^{n-1} \approx S^{n-2}$

$$M[\vec{u_1} | \dots | \vec{u_n}] = [M\vec{u_1} | \dots | M\vec{u_n}], \quad M \in O_n(\mathbb{R}).$$

Method II (Gauss-Gram-Schmidt Approach):

Generate a random matrix X by filling an $n \times n$ matrix with iid $N(0,1)$ random variables.

$X_{ij} \sim N(0,1) \quad \forall \ 1 \leq i,j \leq n.$

$$\text{Joint density} = \frac{1}{(2\pi)^{n^2/2}} \prod_{i,j=1}^{n} e^{-x_{ij}^2/2}$$

$$= \frac{1}{(2\pi)^{n^2/2}} e^{-\sum_{i,j=1}^{n^2} x_{ij}/2}$$

Note that $tr(X^T X) = \sum_{i,j=1}^{n} x_{ij}^2 = \|X\|_F^2$ (frobenius norm).

$$\therefore \text{Joint density} = \frac{1}{(2\pi)^{n^2/2}} e^{-\|X\|_F^2/2}$$

What is the density of MX, $M \in O_n(\mathbb{R})$?

All entries of MX are iid $N(0,1)$ since: $(X)_{ij}$'s are independent ~~and~~ so taking $m_{i*} X_{*j}$, $m_{i*} X_{*j'}$ where $j \neq j'$, we get different $(X)_{ij}$'s.

For $j = j'$, but $i \neq i'$, the rows of M are orthogonal so calculating covariance gives:

$$[m_{i1} \ \dots \ m_{in}] \begin{bmatrix} m_{i'1} \\ \vdots \\ m_{i'n} \end{bmatrix} = 0$$

Joint density of MX : $\frac{1}{(2\pi)^{n^2/2}} \cdot e^{-\|MX\|_F^2/2}$

$= \frac{1}{(2\pi)^{n^2/2}} e^{-\|X\|_F^2/2}$