Cholesky Factorization

Symmetric positive definite matrix:-

Recall $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$ is symmetric

if $\underset{\sim}{A} = \underset{\sim}{A}^T$

Such a matrix also satisfies

for $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^m$ $\quad \underset{\sim}{x}^T \underset{\sim}{A} \underset{\sim}{y} = (\underset{\sim}{x}^T \underset{\sim}{A} \underset{\sim}{y})^T$

$$= \underset{\sim}{y}^T \underset{\sim}{A}^T \underset{\sim}{x}$$

$$= \underset{\sim}{y}^T \underset{\sim}{A} \underset{\sim}{x}$$

$$\boxed{(\underset{\sim}{x}, \underset{\sim}{A}\underset{\sim}{y}) = (\underset{\sim}{A}\underset{\sim}{x}, \underset{\sim}{y})}$$

If $(\underset{\sim}{x}, \underset{\sim}{A} \underset{\sim}{y}) = (\underset{\sim}{A} \underset{\sim}{x}, \underset{\sim}{y})$ for all $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^m$

then $A = A^T$

→ 'A' is said to be symmetric positive definite matrix if in addition to $\underset{\sim}{A} = A^T$, $\underset{\sim}{x}^T \underset{\sim}{A} \underset{\sim}{x} > 0 \ \forall$ non-zero $\underset{\sim}{x} \in \mathbb{R}^m$

Note:- If $\underset{\sim}{A} \in \mathbb{R}^{m\times m}$ is symmetric positive definite (S.P.D) and $\underset{\sim}{X} \in \mathbb{R}^{m\times n}$ $(m \geq n)$ (X is full rank) then the matrix $\underset{\sim}{X}^T \underset{\sim}{A} \underset{\sim}{X}$ is S.P.D

Proof:- $\Rightarrow (\underset{\sim}{X}^T \underset{\sim}{A} \underset{\sim}{X})^T = \underset{\sim}{X}^T \underset{\sim}{A}^T \underset{\sim}{X} = \underset{n\times m}{\underset{\sim}{X}^T} \underset{m\times m}{\underset{\sim}{A}} \underset{m\times n}{\underset{\sim}{X}}$

A non-zero $\underset{\sim}{y} \in \mathbb{R}^n$, $\underset{\sim}{y} \neq 0$

$$\underset{\sim}{y}^T (\underset{\sim}{X}^T \underset{\sim}{A} \underset{\sim}{X}) \underset{\sim}{y}$$

$$= (\underset{\sim}{X} \underset{\sim}{y})^T A (\underset{\sim}{X} \underset{\sim}{y}) > 0$$

(Since $\underset{\sim}{X} \underset{\sim}{y} = 0$ only for $\underset{\sim}{y} = 0$ and A is S.P.D)

if choose $\underset{\sim}{X}$ such that each column of $\underset{\sim}{X}$ has 1 in each column and zeros elsewhere, we can express any $n\times n$ principal submatrix of $\underset{\sim}{A}$ to be of form $\underset{\sim}{X}^T \underset{\sim}{A} \underset{\sim}{X}$ for this choice of $\underset{\sim}{X}$

(i) For a S.P.D, $a_{ii} > 0$ for all i

(a_{ii} is diagonal entry of the matrix $\underset{\sim}{A}$)

(ii) Eigenvalues of S.P.D matrix are also positive.

$$\underset{\sim}{A}\underset{\sim}{u} = \lambda \underset{\sim}{u} \quad \text{for } \underset{\sim}{u} \neq 0$$

(λ is eigenvalue, $\underset{\sim}{u}$ is eigenvector)

We have $\underset{\sim}{x}^T \underset{\sim}{A} \underset{\sim}{x} > 0$

$\forall \underset{\sim}{x} \neq 0$

If I choose my $\underset{\sim}{x}$ to $\underset{\sim}{u}$ the eigenvector

$$\underset{\sim}{u}^T \underset{\sim}{A} \underset{\sim}{u} > 0$$

$$\Rightarrow \underset{\sim}{u}^T (\lambda \underset{\sim}{u}) > 0$$

$$\Rightarrow \boxed{\lambda > 0}$$

(iii) element with largest modulus lies on the main diagonal!

* Symmetric Gaussian Elimination

$$\underset{\sim}{A} = \begin{bmatrix} 1 & \underset{\sim}{w}^T \\ \underset{\sim}{w} & \underset{\sim}{K} \end{bmatrix}$$

$$\underline{L}_1 \underline{A} = \underbrace{\begin{bmatrix} 1 & \underline{\omega}^T \\ 0 & \underline{K} - \underline{\omega}\underline{\omega}^T \end{bmatrix}}_{\underline{W}}$$

Gaussian elimination would continue by zeroing out second column and so on!

$$\underline{L}_1 \underline{A} = \underline{W}$$

$$\underline{A} = \underline{L}^{-1} \underline{W}$$

$$= \begin{bmatrix} 1 & \underline{0}^T \\ \underline{\omega} & \underline{I} \end{bmatrix} \begin{bmatrix} 1 & \underline{\omega}^T \\ \underline{0} & \underline{K} - \underline{\omega}\underline{\omega}^T \end{bmatrix}$$

In order to maintain symmetry Cholesky factorization zeros out first row to match zeros introduced in first column.

$$L_1 A U_1 = \underbrace{\begin{bmatrix} 1 & \underline{\omega}^T \\ 0 & \underline{K} - \underline{\omega}\underline{\omega}^T \end{bmatrix}}_{\underline{L}_1 \underline{A}} U_1 = \begin{bmatrix} 1 & \underline{0}^T \\ \underline{0} & \underline{K} - \underline{\omega}\underline{\omega}^T \end{bmatrix}$$

$$\begin{bmatrix} 1 & \underset{\sim}{w}^T \\ 0 & \underset{\sim}{K} - \underset{\sim}{w}\underset{\sim}{w}^T \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & \underset{\sim}{K} - \underset{\sim}{w}\underset{\sim}{w}^T \end{bmatrix}}_{W} \begin{bmatrix} 1 & \underset{\sim}{w}^T \\ \underset{\sim}{0} & I \end{bmatrix}$$

$$L_1 A = W \quad U_1^{-1}$$

$$\underset{\sim}{A} = \begin{bmatrix} 1 & \underset{\sim}{w}^T \\ \underset{\sim}{w} & K \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 1 & 0^T \\ \underset{\sim}{w} & I \end{bmatrix} \begin{bmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & \underset{\sim}{K} - \underset{\sim}{w}\underset{\sim}{w}^T \end{bmatrix} \begin{bmatrix} 1 & \underset{\sim}{w}^T \\ \underset{\sim}{0} & I \end{bmatrix}}$$

The idea behind Cholesky factorization is to continue this process, zeroing columns and rows of $\underset{\sim}{A}$ symmetrically until $\underset{\sim}{A}$ is reduced to an identity matrix!

* **Cholesky Factorization!**

Let us consider $a_{11} \neq 1$, $a_{11} > 0$

$$\alpha = \sqrt{a_{11}}$$

$$\underset{\sim}{A} = \begin{bmatrix} a_{11} & \underset{\sim}{w}^T \\ \underset{\sim}{w} & \underset{\sim}{K} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} \alpha & \underset{\sim}{0}^T \\ \frac{1}{\alpha}\underset{\sim}{w} & \underset{\sim}{I} \end{bmatrix}}_{\underset{\sim}{R}_1^T} \underbrace{\begin{bmatrix} 1 & \underset{\sim}{0}^T \\ \underset{\sim}{0} & \underset{\sim}{K} - \frac{1}{a_{11}}\underset{\sim}{w}\underset{\sim}{w}^T \end{bmatrix}}_{\underset{\sim}{A}_1} \underbrace{\begin{bmatrix} \alpha & \frac{1}{\alpha}\underset{\sim}{w}^T \\ \underset{\sim}{0} & \underset{\sim}{I} \end{bmatrix}}_{\underset{\sim}{R}_1}$$

$$\underset{\sim}{A} = \underset{\sim}{R}_1^T \underset{\sim}{A}_1 \underset{\sim}{R}_1$$

If $\boxed{\left(\underset{\sim}{K} - \frac{1}{a_{11}}\underset{\sim}{w}\underset{\sim}{w}^T\right)_{11}} > 0$, we can again factorize $\underset{\sim}{A}_1 = \underset{\sim}{R}_2^T \underset{\sim}{A}_2 \underset{\sim}{R}_2$

we can repeat this process,

$$\underset{\sim}{A} = \underbrace{\underset{\sim}{R}_1^T \underset{\sim}{R}_2^T \dots \underset{\sim}{R}_m^T}_{\underset{\sim}{R}^T} \underset{\sim}{I} \underbrace{\underset{\sim}{R}_m \dots \underset{\sim}{R}_2 \underset{\sim}{R}_1}_{\underset{\sim}{R}}$$

We get factorization of the form

$\boxed{\underset{\sim}{A} = \underset{\sim}{R}^T \underset{\sim}{R}}$ where $\underset{\sim}{R}$ is upper triangular and $r_{jj} > 0$

Note: How do we know (1,1) entry of $K - \frac{1}{a_{11}} w w^T$ is positive?

Since A is symmetric positive definite matrix

$\underbrace{K - \frac{1}{a_{11}} w w^T}$ is lower right principal submatrix of $\underbrace{R_1^{-T} A R^{-1}}$ (S.P.D)

$$A = R_1^T A_1 R_1$$

$$A_1 = R_1^{-T} A R^{-1}$$

By induction, the same argument shows that all matrices A_j that appear during factorization are S.P.D and this process does not break down!

Thm: Every S.P.D matrix $A \in \mathbb{R}^{m \times m}$ has a unique Cholesky factorization $A = R^T R$, $r_{jj} > 0$

$\underset{\sim}{R}$ is upper triangular

<u>Algo</u>

$$\underset{\sim}{R} = \underset{\sim}{A}$$

for $k = 1:m$

for $j = k+1:m$

$$R(j, j:m) = R(j, j:m) - \frac{R(k, k:m)}{R(k,k)}$$

end

$$R(k, k:m) = \frac{R(k, k:m)}{\sqrt{R(k,k)}}$$

end

The operation count of cholesky factorization is $\sim \frac{1}{3} m^3$

<u>Thm:-</u> Let $\underset{\sim}{A} \in \mathbb{R}^{m \times m}$ is S.P.D for ε_M sufficiently small, cholesky algorithm is guaranteed to run to completion (i.e no $r_{kk} \le 0$ will arise)

and generates $\tilde{R}$ satisfies

$$\tilde{\underset{\sim}{R}}^T \tilde{\underset{\sim}{R}} = \underset{\sim}{A} + \underset{\sim}{\delta A} \, ; \quad \frac{\|\delta A\|}{\|A\|} = O(\varepsilon_M)$$

for $\underset{\sim}{\delta A} \in \mathbb{R}^{m \times m}$

If $\underset{\sim}{A}$ is ill-conditioned $\tilde{R}$ will generally not close to R, at best we can have

$$\frac{\|\tilde{R} - R\|}{\|R\|} = O(\kappa(A)\, \varepsilon_M)$$

But product $\tilde{\underset{\sim}{R}}^T \tilde{\underset{\sim}{R}}$ is much more accurate!

Solving $\underset{\sim}{A}\underset{\sim}{x} = \underset{\sim}{b}$ using Cholesky if $\underset{\sim}{A}$ is S.P.D is the standard way!