2.6 Methods for Symmetric Positive-Definite Matrices

Symmetric Matrices (i.e. $A^T = A$) hold a favored position in linear analysis because they only contain about half as many independent entries as general, $n\times n$ systems. This raises the question as to whether symmetric matrices admit of an efficient factorization method. For matrices which are symmetric, positive definite, this goal can be achieved with the Cholesky factorization.

Def An $n\times n$ matrix is called symmetric if $\boxed{A^T = A.}$
The matrix A is $\boxed{\text{positive-definite}}$ if $\boxed{\vec{x}^T A \vec{x} > 0}$ $\forall \vec{x} \neq 0$.

Ex. Show that $A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$ is symmetric, positive-definite.

$A^T = A$.

We show: pos-def. $x^T A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$= 2x_1^2 + 4x_1x_2 + 5x_2^2 = 2(x_1+x_2)^2 + 3x_2^2 > 0$$

unless: $x_1 = x_2 = 0$. $\square$

Ex. Show $A = \begin{bmatrix} 2 & 4 \\ 4 & 5 \end{bmatrix}$ is NOT positive-definite.

$x^T A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 8x_1x_2 + 5x_2^2 =$

complete the square...

$$2(x_1^2 + 4x_1x_2) + 5x_2^2 = 2(x_1 + 2x_2)^2 - 8x_1x_2 + 5x_2^2$$
$$= 2(x_1 + 2x_2)^2 - 3x_2^2 < 0, \text{ is possible for some } \vec{x}.$$

$\boxed{\text{Property 2}}$ Note that a symmetric, positive-definite matrix is automatically non-singular. Why? (If $\vec{x}^T A\vec{x} > 0$, then $A\vec{x} \neq \vec{0}$ unless $\vec{x} = \vec{0}$.)

(*) If $A_{n\times n}$ is symmetric, then A is positive-definite iff all of its eigenvalues are positive.

$\boxed{\text{Pf}}$: Since A is symmetric it follows (by a theorem in Linear Algebra) that the set of unit, eigenvectors of A are orthonormal & span $\mathbb{R}^n$.

($\Rightarrow$) If A is positive-definite & $A\vec{v} = \lambda\vec{v}$, then $0 < v^T A v = v^T \lambda v$
$= \lambda \|v\|_2^2$, so $\lambda > 0$.

($\Leftarrow$) Conversely, suppose that all of the eigenvalues of A are positive, (again, A is symmetric)
then write any nonzero vector: $\vec{x} = c_1\vec{v}_1 + \cdots + c_n\vec{v}_n$ where the $\vec{v}_i$ are orthonormal unit vectors & not all $c_i = 0$.

Then: $x^T A x = (c_1\vec{v}_1 + \cdots + c_n\vec{v}_n)^T(\lambda_1 c_1\vec{v}_1 + \cdots + \lambda_n c_n\vec{v}_n)$
$= \lambda_1 c_1^2 + \cdots + \lambda_n c_n^2 > 0$, so A is positive-definite. $\square$

$\boxed{\text{Property 2}}$ If A is $n\times n$ symmetric, positive definite & X is an $n\times m$ matrix of $\boxed{\text{full rank}}$ with $n \geq m$, Then $\boxed{X^T A X}$ is $m\times m$ $\boxed{\text{symmetric, positive-definite.}}$

$\boxed{\text{Pf}}$: The matrix: $X^T A X$ is symmetric, since:

$$(X^T A X)^T = X^T A^T (X^T)^T = X^T A X.$$

To show positive-definite, let $\vec{v} \neq \vec{0}$ be m-dimensional.

Note: $\vec{v}^T (X^T A X) \vec{v} = (X\vec{v})^T A (X\vec{v}) \geq 0$, due to positive-definiteness of A. Since X has full rank, its columns are $\boxed{\text{LID}}$, so $X\vec{v} = \vec{0}$ implies $\vec{v} = 0$. □

$\boxed{\text{Def.}}$ A $\boxed{\text{principal submatrix}}$ of a square matrix A is a square submatrix whose diagonal entries are diagonal entries of A. (same rows/cols)

$\boxed{\text{Ex.}}$

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix} \to \begin{bmatrix} 6 & 7 \\ 10 & 11 \end{bmatrix}$$

$\boxed{\text{Property 3}}$ Any principal submatrix of a symmetric, positive-definite matrix is also symmetric, pos-def.

Cholesky Factorization $\boxed{A = R^T R}$ (R: upper-$\triangle$)

A 2×2 example

Consider the symmetric, pos-def matrix:

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad \text{(Note } a > 0 \text{ by property 3)}.$$

In addition, we know that $\begin{vmatrix} a & b \\ b & c \end{vmatrix} = ac - b^2 > 0$ since $|A|$ = product of the eigenvalues of A. (all $\lambda_i > 0$ by property 1).

Writing $A = R^T R$, for an upper-$\triangle$ matrix R, implies the form:

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = \underbrace{\begin{bmatrix} \sqrt{a} & 0 \\ u & v \end{bmatrix}}_{R^T} \underbrace{\begin{bmatrix} \sqrt{a} & u \\ 0 & v \end{bmatrix}}_{R} = \begin{bmatrix} a & u\sqrt{a} \\ u\sqrt{a} & u^2 + v^2 \end{bmatrix}$$

We want to check whether this is possible. Comparing the LHS/RHS we see: $\boxed{u = \frac{b}{\sqrt{a}}}$ & $\boxed{v^2 = c - u^2}$.

Note that: $v^2 = c - \left(\frac{b}{\sqrt{a}}\right)^2 = c - \frac{b^2}{a} > 0$ since $|A| > 0$.

This verifies that v can be defined as a real-number, & so the Cholesky factorization exists for 2×2 symmetric, positive-definite matrices, where:

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} \sqrt{a} & 0 \\ \frac{b}{\sqrt{a}} & \sqrt{c - b^2/a} \end{bmatrix} \begin{bmatrix} \sqrt{a} & \frac{b}{\sqrt{a}} \\ 0 & \sqrt{c - b^2/a} \end{bmatrix} = R^T R.$$

Note that the Cholesky factorization is NOT unique, as we could have set: $\boxed{v = -\sqrt{c - b^2/a}}$.

$\boxed{\text{Theorem}}$ (Cholesky Factorization Theorem) If A is symmetric, positive-definite $(n \times n)$, then there exists an upper-Δ matrix R, where: $\boxed{A = R^T R}$.

$\boxed{\text{Pf}}$: We construct R by $\boxed{\text{induction}}$ on the size of A: n. ($n=2$ was proved above).

Consider the partition of A:
$$A = \left[\begin{array}{c|c} a & \vec{b}^T \\ \hline \vec{b} & C \end{array}\right]$$
$C: (n-1)\times(n-1)$

As before, set $\boxed{u = \vec{b}/\sqrt{a}}$, let $\boxed{A_1 = C - uu^T}$ & define the invertible matrix S:

$$S = \left[\begin{array}{c|c} \sqrt{a} & \vec{u}^T \\ \hline 0 & \\ \vdots & I \\ 0 & \end{array}\right] \quad \text{yields:}$$

$$S^T \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & A_1 \end{array}\right] S = \left[\begin{array}{c|c} \sqrt{a} & 0 \cdots 0 \\ \hline \vec{u} & I \\ & \end{array}\right] \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & A_1 \\ 0 & \end{array}\right] \left[\begin{array}{c|c} \sqrt{a} & \vec{u}^T \\ \hline 0 & \\ \vdots & I \\ 0 & \end{array}\right]$$

$$= \left[\begin{array}{c|c} a & \vec{b}^T \\ \hline \vec{b} & \vec{u}\vec{u}^T + A_1 \end{array}\right] = A$$

Notice that A_1 is symmetric, positive-definite; this follows from the facts that:

$$\left[\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{array}\right] = (S^T)^{-1} A S^{-1}$$ is symmetric, pos-def. by property 2, and therefore, so is the $(n-1)\times(n-1)$ principal submatrix A_1 by property 3.

By the induction hypothesis, $A_1 = V^T V$ where V is upper-Δ.

Finally, define: $R = \left[\begin{array}{c|c} \sqrt{a} & \vec{u}^T \\ \hline 0 & \\ \vdots & V \\ 0 & \end{array}\right]$ and check that:

$$R^T R = \left[\begin{array}{c|ccc} \sqrt{a} & 0 & \cdots & 0 \\ \hline \vec{u} & & V^T & \end{array}\right] \left[\begin{array}{c|c} \sqrt{a} & \vec{u}^T \\ \hline 0 & \\ \vdots & V \\ 0 & \end{array}\right] = \left[\begin{array}{c|c} a & \vec{b}^T \\ \hline \vec{b} & \vec{u}\vec{u}^T + \underbrace{V^T V}_{=A_1} \end{array}\right]$$

$$= A. \ \square$$

From this proof we can devise the <u>standard algorithm for Cholesky factorization.</u>

The matrix $\boxed{R}$ is built from the outside in. First, we find $r_{11} = \sqrt{a_{11}}$ & set the rest of the top row of R to $\boxed{u^T = b^T / r_{11}}$. Then $\underline{u^T u}$ is subtracted from the <u>lower principal $(n-1)\times(n-1)$ submatrix</u>, and the same steps are repeated until all the rows of R are filled.

According to the theorem, the new principal submatrix is <u>pos-def.</u> at every stage of construction, so the square root operation succeeds.

$\boxed{\text{Ex.}}$ Find the <u>Cholesky factorization</u> of $\begin{bmatrix} 4 & -2 & 2 \\ -2 & 2 & -4 \\ 2 & -4 & 11 \end{bmatrix}$.

Find the top row of R: $r_{11} = \sqrt{a_{11}} = \boxed{2}$.

$R_{12} = \dfrac{a_{12}}{R_{11}} = \dfrac{-2}{2} = \boxed{-1}$ $\qquad R_{13} = \dfrac{a_{13}}{R_{11}} = \dfrac{2}{2} = \boxed{1}$

So the top row of R is: $[2, -1, 1] \rightarrow R = \begin{bmatrix} 2 & -1 & 1 \\ & & \\ & & \end{bmatrix}$ "$\vec{u}^T$"

Subtracting the outer product $\vec{u}\vec{u}^T$ from the <u>lower principal</u> 2×2 submatrix of A gives us: $\vec{u}\vec{u}^T = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

$\rightarrow \begin{bmatrix} & & \\ & 2 & -4 \\ & -4 & 11 \end{bmatrix} - \begin{bmatrix} & & \\ & 1 & -1 \\ & -1 & 1 \end{bmatrix} = \begin{bmatrix} & & \\ & 1 & -3 \\ & -3 & 10 \end{bmatrix}$ "A_1"

Now we repeat the same steps on the 2x2 submatrix

To find $\boxed{R_{22}=1}$, $\boxed{R_{23}=\frac{-3}{1}=-3.}$ $=\left(\frac{A_{23}}{R_{22}}\right)$

$\left(R_{22}=\sqrt{A_{11}}=\sqrt{1}=1\right)$

$$R=\begin{bmatrix} 2 & -1 & 1 \\ & 1 & -3 \\ & & \end{bmatrix}$$

$$A_1=\begin{bmatrix} 1 & -3 \\ -3 & 10 \end{bmatrix}$$

The lower principal submatrix of A is: $10-uu^T=10-(-3)(-3)=1$,

so $R_{33}=\sqrt{1}=1$.

Thus $R=\begin{bmatrix} 2 & -1 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix}$, indeed: $\boxed{A=R^TR}$.

<u>Note:</u> To solve $A\vec{x}=\vec{b}$ → ① $A=R^TR$ → $R^T(R\vec{x})=b$

② Solve: $R^Tc=b$ → $Rx=c$. ✓

Cholesky Pseudo Code

for $k=1,\dots,n$

if $\underline{A_{kk}<0}$, $\boxed{\text{END}}$

$R_{kk}=\sqrt{A_{kk}}$

$\vec{u}^T=\frac{1}{R_{kk}}\cdot A_{ki}$

$R_{k,k+1}=u^T$

$A_{k+1,k+1}=A_{k+1..}-uu^T$

$\boxed{5(a)}$

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} \qquad \left.\begin{aligned} R_{11} &= \sqrt{a_{11}} = 1 \\ R_{12} &= \tfrac{2}{1} = 2 \end{aligned}\right\} \rightarrow R = \begin{bmatrix} 1 & \overset{\text{"}u^T\text{"}}{2} \\ & \end{bmatrix}$$

$$A_1 = 8 - (2)(2) = \boxed{4} \rightarrow R_{22} = \sqrt{A_{11}} = \sqrt{4} = 2$$

$$R = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \rightarrow R^T R = \begin{bmatrix} 1 & 0 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 8 \end{bmatrix} = A. \checkmark$$