ECEn 671: Mathematics of Signals and Systems

Randal W. Beard

Brigham Young University

September 1, 2023

Section 1

Cholesky Factorization

Square Root of a Matrix

- If $B = B^H > 0$ then we can compute the "square root" of B as $B = QQ^H$ where $Q = B^{\frac{1}{2}}$ is the square root of B.
- In general, the square root of a matrix is not unique!

Example

Let $B = \begin{pmatrix} 9 & 0 \\ 0 & 0 \end{pmatrix}$

We can write

$$B = \begin{pmatrix} 3 \\ 0 \end{pmatrix} \begin{pmatrix} 3 & 0 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$$

So both $Q = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ and $Q = \begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$ are square roots of B.

Cholesky Factorization

Definition

The Cholesky factorization of B is a square lower triangular square root $L \in \mathbb{C}^{n \times n}$ of B, where

$$B = LL^H.$$

Note that this can also be written as

$$B = U^H U$$

where $U = L^H$ is upper triangular.

Cholesky Factorization: Numerical Algorithm

Let $B = \begin{pmatrix} \alpha & \mathbf{v}^H \\ \mathbf{v} & B_1 \end{pmatrix}$. Then factor B as

$$\begin{aligned} B &= \begin{pmatrix} \alpha & \mathbf{v}^H \\ \mathbf{v} & B_1 \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{\alpha} & 0 \\ \frac{\mathbf{v}}{\sqrt{\alpha}} & I_{n-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & B_1 - \frac{\mathbf{v}\mathbf{v}^H}{\alpha} \end{pmatrix} \begin{pmatrix} \sqrt{\alpha} & \frac{\mathbf{v}^H}{\sqrt{\alpha}} \\ 0 & I_{n-1} \end{pmatrix} \end{aligned}$$

Cholesky Factorization: Numerical Algorithm, cont.

(RECURSIVE ALGORITHM)

Now find the Cholesky factorization of $B_1 - \frac{\mathbf{v}\mathbf{v}^H}{\alpha} \triangleq G_1 G_1^H$, so that

$$B = \begin{pmatrix} \sqrt{\alpha} & 0 \\ \frac{\mathbf{v}}{\sqrt{\alpha}} & I_{n-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & G_1 G_1^H \end{pmatrix} \begin{pmatrix} \sqrt{\alpha} & \frac{\mathbf{v}^H}{\sqrt{\alpha}} \\ 0 & I_{n-1} \end{pmatrix}$$
$$= \begin{pmatrix} \sqrt{\alpha} & 0 \\ \frac{\mathbf{v}}{\sqrt{\alpha}} & G_1 \end{pmatrix} \begin{pmatrix} \sqrt{\alpha} & \frac{\mathbf{v}^H}{\sqrt{\alpha}} \\ 0 & G_1^H \end{pmatrix}$$

which implies that the Cholesky factor is

$$L = \begin{pmatrix} \sqrt{\alpha} & 0 \\ \frac{\mathbf{v}}{\sqrt{\alpha}} & G_1 \end{pmatrix}.$$

Cholesky Factorization: Example

Let $B = \begin{pmatrix} 1 & 2 & 4 & 1 \\ 2 & 13 & 17 & 8 \\ 4 & 17 & 29 & 16 \\ 1 & 8 & 16 & 30 \end{pmatrix}$. Then

$$B = \begin{pmatrix} \alpha_1 & \mathbf{v}_1^\top \\ \mathbf{v}_1 & B_1 \end{pmatrix},$$

where

$$\begin{aligned} \alpha_1 &= 1 \\ \mathbf{v}_1 &= \begin{pmatrix} 2 & 4 & 1 \end{pmatrix}^\top \\ B_1 &= \begin{pmatrix} 13 & 17 & 8 \\ 17 & 29 & 16 \\ 8 & 16 & 30 \end{pmatrix}. \end{aligned}$$

Cholesky Factorization: Example, cont.

Therefore

$$
\begin{aligned}
B &= \begin{pmatrix} \sqrt{\alpha_1} & \mathbf{0}^\top \\ \frac{\mathbf{v}_1}{\sqrt{\alpha_1}} & G_1 \end{pmatrix} \begin{pmatrix} \sqrt{\alpha_1} & \frac{\mathbf{v}_1^\top}{\sqrt{\alpha_1}} \\ \mathbf{0} & G_1^\top \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & & & \\ 4 & & G_1 & \\ 1 & & & \end{pmatrix} \begin{pmatrix} 1 & 2 & 4 & 1 \\ 0 & & & \\ 0 & & G_1^\top & \\ 0 & & & \end{pmatrix}
\end{aligned}
$$

where

$$
\begin{aligned}
G_1 G_1^\top &= B_1 - \frac{\mathbf{v}_1 \mathbf{v}_1^\top}{\alpha_1} \\
&= \begin{pmatrix} 13 & 17 & 8 \\ 17 & 29 & 16 \\ 8 & 16 & 30 \end{pmatrix} - \frac{1}{1} \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix} \begin{pmatrix} 2 & 4 & 1 \end{pmatrix} \\
&= \begin{pmatrix} 9 & 9 & 6 \\ 9 & 13 & 12 \\ 6 & 12 & 29 \end{pmatrix}.
\end{aligned}
$$

Cholesky Factorization: Example, cont.

Therefore

$$\begin{aligned} G_1 G_1^\top &= \begin{pmatrix} \sqrt{\alpha_2} & \mathbf{0}^\top \\ \frac{\mathbf{v}_2}{\sqrt{\alpha_2}} & G_2 \end{pmatrix} \begin{pmatrix} \sqrt{\alpha_2} & \frac{\mathbf{v}_2^\top}{\sqrt{\alpha_2}} \\ \mathbf{0} & G_2^\top \end{pmatrix} \\ &= \begin{pmatrix} 3 & 0 & 0 \\ 3 & & \\ 2 & & G_2 \end{pmatrix} \begin{pmatrix} 3 & 3 & 2 \\ 0 & & \\ 0 & & G_2^\top \end{pmatrix} \end{aligned}$$

where

$$\begin{aligned} G_2 G_2^\top &= B_2 - \frac{\mathbf{v}_2 \mathbf{v}_2^\top}{\alpha_2} \\ &= \begin{pmatrix} 13 & 12 \\ 12 & 29 \end{pmatrix} - \frac{1}{9} \begin{pmatrix} 9 \\ 6 \end{pmatrix} \begin{pmatrix} 9 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 4 & 6 \\ 6 & 25 \end{pmatrix}. \end{aligned}$$

Cholesky Factorization: Example, cont.

Therefore

$$
\begin{aligned}
G_2 G_2^\top &= \begin{pmatrix} \sqrt{\alpha_3} & 0^\top \\ \frac{\mathbf{v}_3}{\sqrt{\alpha_3}} & G_3 \end{pmatrix} \begin{pmatrix} \sqrt{\alpha_3} & \frac{\mathbf{v}_3^\top}{\sqrt{\alpha_3}} \\ 0 & G_3^\top \end{pmatrix} \\
&= \begin{pmatrix} 2 & 0 \\ 3 & G_3 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 0 & G_3^\top \end{pmatrix}
\end{aligned}
$$

where

$$
\begin{aligned}
G_3 G_3^\top &= B_3 - \frac{\mathbf{v}_3 \mathbf{v}_3^\top}{\alpha_3} \\
&= 25 - \frac{1}{4} 3 \cdot 3 \\
&= 16
\end{aligned}
$$

Therefore $G_3 = 4$.

Cholesky Factorization: Example, cont.

Combining gives

$$L = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & & & \\ 4 & & G_1 & \\ 1 & & & \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 3 & & \\ 1 & 2 & & G_2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 3 & 2 & 0 \\ 1 & 2 & 3 & G_3 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 3 & 2 & 0 \\ 1 & 2 & 3 & 4 \end{pmatrix}.$$

Applications of Cholesky Factorization: Quadratic Forms

The quadratic form

$$x^H Q x = \|x\|_Q^2$$

where $Q = Q^H$, can be written as

$$x^H Q x = x^H U^H U x = \|Ux\|_2^2$$

where $Q = U^H U = LL^H$

In other words, work with the regular 2-norm as opposed to the Q norm.

Applications of Cholesky Factorization: Simulating a random vector

Suppose you want to generate in Matlab/Simulink/Python/etc. a Gaussian random vector with covariance $R = R^T > 0$.

The Matlab `randn([m,1])` command returns an $m \times 1$ random vector which is normally distributed with zero mean and co-variance I ($\mathcal{N}(0, I)$).

To generate $\mathcal{N}(0, R)$ let $R = LL^T$ and let $z = Lx$ where $x \sim \mathcal{N}(0, I)$.

Then

$$E\{zz^T\} = E\{Lxx^T L^T\} = LE\{xx^T\}L^T = LL^T = R$$
$$\Rightarrow z \sim \mathcal{N}(0, R).$$

Applications of Cholesky Factorization: Solving normal equations

Normal equations are given by

$$R\mathbf{c} = \mathbf{b}$$

where $R = R^H$ is the Grammian and full rank if the data vectors are linearly independent.

Let $R = LL^H$, then $LL^H\mathbf{c} = \mathbf{b}$

First solve

$$L\mathbf{y} = \mathbf{b}$$

by forward substitution, and then solve

$$L^H\mathbf{c} = \mathbf{y}$$

by backward substitution.

Applications of Cholesky Factorization: Kalman filtering

In Kalman filtering we propagate two items; The estimate $\hat{x}(k)$ and the error covariance $P(k)$ where $P(k) = P^T(k) > 0$.

If implemented directly, numerical error can cause $P(k)$ to become indefinite introducing large errors into the estimate $\hat{x}(k)$.

To avoid this problem a "square root" Kalman filter is usually implemented where $P(k) = L(k)L^T(k)$ and $L(k)$ is propagated instead of $P(k)$. Then even with numerical errors in $L(k)$, $P(k)$ is still symmetric positive definite.

Cholesky Factorization: cont.

In Matlab:

```
L1 = [2, 0, 0; 3, 4, 0; 5, 6, 7];
A = L1 * L1';
L = chol(A)'
```

In Python:

```
import numpy as np
import scipy.linalg as linalg

L1 = np.array([[2, 0, 0], [3, 4, 0], [5, 6, 7]])
A = L1 @ L1.T
L = linalg.cholesky(A)
```

L should equal L_1.

Note that both Matlab and Python return an upper triangular matrix.

Homework problem: Write your own custom `cholesky` function and compare to the built in `cholesky` function on 100 randomly generated symmetric matrices.