# Algorithms and Data Structures

# Matrix Decomposition
# Cholesky Decomposition

$$\begin{bmatrix}A_{00} & A_{01} & A_{02}\\ A_{10} & A_{11} & A_{12}\\ A_{20} & A_{21} & A_{22}\end{bmatrix} = \overbrace{\begin{bmatrix}L_{00} & 0 & 0\\ L_{10} & L_{11} & 0\\ L_{20} & L_{21} & L_{22}\end{bmatrix}}^{\text{Lower Triangular L}} \underbrace{\begin{bmatrix}L_{00} & L_{10} & L_{20}\\ 0 & L_{11} & L_{21}\\ 0 & 0 & L_{22}\end{bmatrix}}_{\text{Transpose of L}}$$

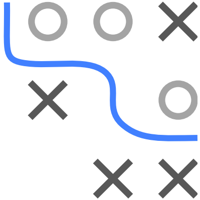
**Learning goals**

- Cholesky decomposition
- Properties of Cholesky decomposition

# CHOLESKY DECOMPOSITION

**Aim:** Solve LES of the form $\mathbf{Ax} = \mathbf{b}$

with $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{A}$ positive-definite



1. Write $\mathbf{A}$ as $\mathbf{A} = \mathbf{L}\mathbf{L}^\top$
2. Solve $\mathbf{Ly} = \mathbf{b}$ by forward substitution
3. Solve $\mathbf{L}^\top\mathbf{x} = \mathbf{y}$ by back substitution

# CHOLESKY DECOMPOSITION / 2

Example: Let $\mathbf{Ax} = \boldsymbol{b}$ be a LES

$$\begin{pmatrix} 4 & 2 & 2 & 2 \\ 2 & 5 & 3 & 3 \\ 2 & 3 & 11 & 5 \\ 2 & 3 & 5 & 19 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 22 \\ 33 \\ 61 \\ 99 \end{pmatrix}$$

# CHOLESKY DECOMPOSITION

1. Write $\mathbf{A}$ as $\mathbf{A} = \mathbf{L}\mathbf{L}^\top$

$$\begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & l_{31} & l_{41} \\ 0 & l_{22} & l_{32} & l_{42} \\ 0 & 0 & l_{33} & l_{43} \\ 0 & 0 & 0 & l_{44} \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 & 2 \\ 2 & 5 & 3 & 3 \\ 2 & 3 & 11 & 5 \\ 2 & 3 & 5 & 19 \end{pmatrix}$$

$$l_{11}^2 = a_{11} \quad \rightarrow \quad l_{11} = \sqrt{a_{11}} = \sqrt{4} = 2$$

# CHOLESKY DECOMPOSITION

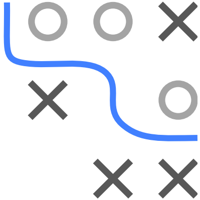
1. Write $\mathbf{A}$ as $\mathbf{A} = \mathbf{L}\mathbf{L}^\top$

$$
\begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix}
\begin{pmatrix} l_{11} & l_{21} & l_{31} & l_{41} \\ 0 & l_{22} & l_{32} & l_{42} \\ 0 & 0 & l_{33} & l_{43} \\ 0 & 0 & 0 & l_{44} \end{pmatrix}
=
\begin{pmatrix} 4 & 2 & 2 & 2 \\ 2 & 5 & 3 & 3 \\ 2 & 3 & 11 & 5 \\ 2 & 3 & 5 & 19 \end{pmatrix}
$$

$$l_{11}^2 = a_{11} \quad \rightarrow \quad l_{11} = \sqrt{a_{11}} = \sqrt{4} = 2$$

$$l_{21} \cdot l_{11} = a_{21} \quad \rightarrow \quad l_{21} = \frac{a_{21}}{l_{11}} = \frac{2}{2} = 2$$

# CHOLESKY DECOMPOSITION



1. Write $\mathbf{A}$ as $\mathbf{A} = \mathbf{L}\mathbf{L}^\top$

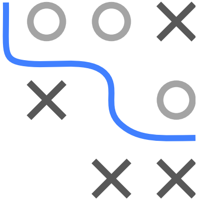
$$\begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & l_{31} & l_{41} \\ 0 & l_{22} & l_{32} & l_{42} \\ 0 & 0 & l_{33} & l_{43} \\ 0 & 0 & 0 & l_{44} \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 & 2 \\ 2 & 5 & 3 & 3 \\ 2 & 3 & 11 & 5 \\ 2 & 3 & 5 & 19 \end{pmatrix}$$

$$l_{11}^2 = a_{11} \quad \rightarrow \quad l_{11} = \sqrt{a_{11}} = \sqrt{4} = 2$$

$$l_{21} \cdot l_{11} = a_{21} \quad \rightarrow \quad l_{21} = \frac{a_{21}}{l_{11}} = \frac{2}{2} = 2$$

$$l_{22}^2 + l_{21}^2 = a_{22} \quad \rightarrow \quad l_{22} = \sqrt{a_{22} - l_{21}^2} = \sqrt{5 - 1^2} = 2$$

# CHOLESKY DECOMPOSITION



1. Write $\mathbf{A}$ as $\mathbf{A} = \mathbf{L}\mathbf{L}^\top$

$$\begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & l_{31} & l_{41} \\ 0 & l_{22} & l_{32} & l_{42} \\ 0 & 0 & l_{33} & l_{43} \\ 0 & 0 & 0 & l_{44} \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 & 2 \\ 2 & 5 & 3 & 3 \\ 2 & 3 & 11 & 5 \\ 2 & 3 & 5 & 19 \end{pmatrix}$$

$$l_{11}^2 = a_{11} \quad \rightarrow \quad l_{11} = \sqrt{a_{11}} = \sqrt{4} = 2$$

$$l_{21} \cdot l_{11} = a_{21} \quad \rightarrow \quad l_{21} = \frac{a_{21}}{l_{11}} = \frac{2}{2} = 2$$

$$l_{22}^2 + l_{21}^2 = a_{22} \quad \rightarrow \quad l_{22} = \sqrt{a_{22} - l_{21}^2} = \sqrt{5 - 1^2} = 2$$

$$l_{31} \cdot l_{11} = a_{31} \quad \rightarrow \quad l_{31} = \frac{a_{31}}{l_{11}} = \frac{2}{2} = 1$$

$$\vdots$$

General formula: $l_{jj} = \left(a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2\right)^{\frac{1}{2}} \qquad l_{ij} = \frac{1}{l_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}\right)$

# CHOLESKY DECOMPOSITION

2. Solve $\mathbf{Ly} = \mathbf{b}$ by forward substitution

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 1 & 4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 22 \\ 33 \\ 61 \\ 99 \end{pmatrix}$$

$$\begin{pmatrix} 2y_1 \\ y_1 + 2y_2 \\ y_1 + y_2 + 3y_3 \\ y_1 + y_2 + y_3 + 4y_4 \end{pmatrix} = \begin{pmatrix} 22 \\ 33 \\ 61 \\ 99 \end{pmatrix}$$

$$\Rightarrow \quad y_1 = 11, y_2 = 11, y_3 = 13, y_4 = 16$$

# CHOLESKY DECOMPOSITION / 2

3. Solve $\mathbf{L}^\top \mathbf{x} = \mathbf{y}$ by back substitution

$$\begin{pmatrix} 2 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 11 \\ 11 \\ 13 \\ 16 \end{pmatrix}$$

$$\Rightarrow \quad x_4 = 4, x_3 = 3, x_2 = 2, x_1 = 1$$

# CHOLESKY DECOMPOSITION / 3

Calculation of the lower triangular matrix (**L**):

$$\begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix} \begin{pmatrix} l_{11} & l_{21} & l_{31} & l_{41} \\ 0 & l_{22} & l_{32} & l_{42} \\ 0 & 0 & l_{33} & l_{43} \\ 0 & 0 & 0 & l_{44} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$$

Thus the entries of **L** (j rows, i columns) result from

$$l_{ij} = \begin{cases} 0 & \text{for } i < j \\ \left(a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2\right)^{\frac{1}{2}} & \text{for } i = j \\ \frac{1}{l_{jj}}\left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}\right) & \text{for } i > j \end{cases}$$

**Important:** Order of calculation (row by row) matters!

$\rightarrow l_{11}, l_{21}, l_{22}, l_{31}, l_{32}, l_{33}, \ldots, l_{nn}$

# CHOLESKY DECOMPOSITION / 4

**Algorithm** Cholesky decomposition

> **for** $j = 1$ to $n$ **do**
>     $l_{jj} = \left(a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2\right)^{\frac{1}{2}}$
>     **for** $i = j + 1$ to $n$ **do**
>         $l_{ij} = \frac{1}{l_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} l_{jk}\right)$
>     **end for**
> **end for**

If we consider only the (dominant) multiplications, we count in each step of the outer loop

- For diagonal elements: $(j - 1)$ multiplications
- For non-diagonal elements: $(n - j)(j - 1)$ multiplications

# CHOLESKY DECOMPOSITION / 5

In total, we estimate the computational effort with

$$
\begin{aligned}
& \sum_{j=1}^{n} [(j-1) + (n-j)(j-1)] \\
= \; & \sum_{j=1}^{n} [j - 1 + nj - n - j^2 + j] = \sum_{j=1}^{n} [(n+2)j - 1 - j^2] \\
= \; & n\frac{(n+2)(n+1)}{2} - n - n\frac{(n+1)(2n+1)}{6} \\
= \; & n \cdot \frac{3(n+2)(n+1) - 6 - (n+1)(2n+1)}{6} \\
= \; & n \cdot \frac{3n^2 + 9n + 6 - 6 - 2n^2 - 2n - n - 1}{6} \\
\approx \; & \frac{1}{6}n^3 + \mathcal{O}(n^2) \quad \text{for large } n
\end{aligned}
$$

# PROPERTIES OF CHOLESKY DECOMPOSITION

- Most important procedure for positive-definite matrices
- Algorithm is always stable (no pivoting necessary)
- **Existence** and **uniqueness**: The Cholesky decomposition exists and is unique for a positive-definite matrix **A**
- Runtime behavior:
  - Decomposition of the matrix: $\frac{n^3}{6} + \mathcal{O}(n^2)$ multiplications
  - Forward and back substitution: $n^2$

# PROPERTIES OF CHOLESKY DECOMPOSITION / 2

```
cholesky = function(a) {
  n = nrow(a)
  l = matrix(0, nrow = n, ncol = n)
  for (j in 1:n) {
    l[j, j] = (a[j, j] - sum(l[j, 1:(j - 1)]^2))^0.5
    if (j < n) {
      for (i in (j + 1):n) {
        l[i, j] = (a[i, j] -
          sum(l[i, 1:(j - 1)] * l[j, 1:(j - 1)])) / l[j, j]
      }
    }
  }
  return(l)
}
```

# PROPERTIES OF CHOLESKY DECOMPOSITION / 3

```
A = crossprod(matrix(runif(16), 4, 4))
cholesky(A)

t(chol(A))

A = crossprod(matrix(runif(1e+06), 1e+03, 1e+03))
system.time(cholesky(A))
system.time(chol(A))
```

# APPLICATION EX.: MULTIVARIATE GAUSSIAN

**Target:** Efficient evaluation of the density of a normal distribution.

The density of the $d$-dimensional multivariate normal distribution is

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{d}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \exp\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\}$$

with $\mathbf{x} \in \mathbb{R}^d$, $\mathrm{Cov}(\mathbf{x}) = \boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}$ **positive-definite**.

With $\mathbf{z} = \mathbf{x} - \boldsymbol{\mu}, \mathbf{z} \in \mathbb{R}^d$ we obtain:

$$(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{\mathbf{-1}} (\mathbf{x} - \boldsymbol{\mu}) = \mathbf{z}^\top \boldsymbol{\Sigma}^{-1} \mathbf{z}$$

**Problem:** Calculation of $\boldsymbol{\Sigma}^{\mathbf{-1}}$ is numerically unstable and requires a long time.

# APPLICATION EX.: MULTIVARIATE GAUSSIAN / 2

**Solution:** Use Cholesky decomposition to avoid inverting $\Sigma^{-1}$. Write $\Sigma$ as $\Sigma = \mathbf{L}\mathbf{L}^\top$, $\mathit{rank}(\mathbf{L}) = d$.

Thus it holds:

$$\begin{aligned}
\mathbf{z}^\top \Sigma^{-1} \mathbf{z} &= \mathbf{z}^\top (\mathbf{L}\mathbf{L}^\top)^{-1} \mathbf{z} \\
&= \mathbf{z}^\top (\mathbf{L}^\top)^{-1} \mathbf{L}^{-1} \mathbf{z} \\
&= (\mathbf{L}^{-1}\mathbf{z})^\top \mathbf{L}^{-1} \mathbf{z} \\
&= \mathbf{v}^\top \mathbf{v}
\end{aligned}$$

with $\mathbf{v} = \mathbf{L}^{-1}\mathbf{z}$, $\mathbf{v} \in \mathbb{R}^d$.

To avoid inverting $\mathbf{L}$ we can calculate $\mathbf{v}$ as a solution of the LES

$$\mathbf{L}\mathbf{v} = \mathbf{z}$$

Then we can calculate $\mathbf{v}^T\mathbf{v}$ as a scalar product of two $d$-dimensional vectors.