Algorithms and Data Structures

Matrix Decomposition

Introduction to Matrix Decomposition and Recap on Matrices

System of Linear Equations – Matrix Form

$$A = \begin{bmatrix} a_{11} + a_{12} + a_{13} + \dots + a_{1n} \\ a_{21} + a_{22} + a_{23} + \dots + a_{2n} \\ a_{31} + a_{32} + a_{33} + \dots + a_{3n} \\ \dots \\ \dots \\ a_{m1} + a_{m2} + a_{m3} + \dots + a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}, \quad B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$

Coefficient matrix — *Variable matrix* — *Constant matrix*

Learning goals

- Systems of linear equations
- Basic knowledge of matrices

SYSTEMS OF LINEAR EQUATIONS

Motivation: Large datasets can be challenging when it comes to data processing. Solving a LES of 10 equations in 10 unknowns might be easy, but an increasing size of the system usually comes along with algorithmic complexity and efficiency which may become critical.

Definition: We consider a system of linear equations

$$\mathbf{A}\mathbf{x} = \boldsymbol{b}$$

with $\mathbf{A} \in \mathbb{R}^{n \times n}$ **regular** (invertible), $\mathbf{x} \in \mathbb{R}^n$ and $\boldsymbol{b} \in \mathbb{R}^n$.

In Chapter 3 - Numerics we have considered the condition of linear systems and have shown that

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|$$

We have seen that a solution of the LES by using $\boldsymbol{x} = \boldsymbol{A}^{-1}\boldsymbol{b}$ is to be avoided from a numerical perspective.

SYSTEMS OF LINEAR EQUATIONS / 2

Reminder: Why should we not calculate $\mathbf{A}^{-1}\mathbf{b}$ (`x = solve(A) %*% b`)?

- **Effort**: Calculation of inverse needs n^3 flops. In addition there are about $2n^2$ flops for matrix-vector multiplication.
- **Memory**: The n^2 entries of the inverted matrix must be stored.
- **Stability**: Two (possibly ill-posed) subproblems are solved:
  1. Calculation of $\mathbf{A}^{-1}$ by solving $\mathbf{Ax} = \mathbf{0}$
     $\rightarrow$ Condition $\kappa_1 = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$
  2. Calculation of matrix-vector product $\mathbf{A}^{-1}\mathbf{b}$
     $\rightarrow$ Condition $\kappa_2 = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$

  The error is amplified by the factor $\kappa = \kappa_1 \cdot \kappa_2 = (\|\mathbf{A}\|\|\mathbf{A}^{-1}\|)^2$.

SYSTEMS OF LINEAR EQUATIONS / 3

Example: Hilbertmatrix (see Chapter 3 - Numerics)

We solve the LES once by matrix inversion and once directly using `solve(H, b)`.

```
n = 10
H = hilbert(n)
x = rnorm(n)
b = H %*% x
```

SYSTEMS OF LINEAR EQUATIONS / 4

```
microbenchmark(
xhat_inverting = solve(H) %*% b, # matrix inversion
xhat_solving = solve(H, b)) # direct solution
## Unit: microseconds
## expr min lq mean median uq max neval
## xhat_inverting 25.7 28.55 31.554 31.3 33.00 59.0 100
## xhat_solving 14.8 16.10 18.890 17.9 19.55 81.2 100
## cld
## b
## a
```

```
norm(xhat_inverting - x) # matrix inversion
## [1] 0.00547865346403784З
```

```
norm(xhat_solving - x) # direct solution
## [1] 0.000375089792123334З
```

SYSTEMS OF LINEAR EQUATIONS / 5

Better: Solve LES directly with `x = solve(A, b)`

In this chapter: How can a LES be solved in a **stable** and **efficient** way? How does `solve(A, b)` work?

Idea: Decompose the matrix **A** into a product of matrices in such a way that the linear system is "easily" solvable.

Important procedures:

- LU decomposition
- Cholesky decomposition
- QR decomposition

REMINDER: ELEMENTARY MATRICES

Elementary row and column transformations of a matrix $\boldsymbol{A}$:

- **type I:** Row switching (Column switching)
- **type II:** Multiplication of row (column) i by a real number $\lambda \neq 0$
- **type III:** Addition of multiples of row (column) j to row (column) i

These transformations are applied when multiplying $\boldsymbol{A}$ by **elementary matrices** from the left (row transformations) or from the right (column transformations).

REMINDER: ELEMENTARY MATRICES / 2

Be $\mathbf{A}$ a (4×4) - matrix: $\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix}$

Example elementary matrix type I:
Switch row 2 and row 4 in $\mathbf{A}$:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 13 & 14 & 15 & 16 \\ 9 & 10 & 11 & 12 \\ 5 & 6 & 7 & 8 \end{pmatrix}$$

The elementary matrix is created by switching the 2nd and 4th row (column) of the identity matrix.

REMINDER: ELEMENTARY MATRICES / 3

Example elementary matrix type II:

Multiply column 3 of **A** with λ.

$$\mathbf{A}\begin{pmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & \lambda & 0\\0 & 0 & 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 2 & \lambda 3 & 4\\5 & 6 & \lambda 7 & 8\\9 & 10 & \lambda 11 & 12\\13 & 14 & \lambda 15 & 16\end{pmatrix}$$

Example elementary matrix type III:

Multiply row 1 with λ and add it to row 3.

$$\begin{pmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\\lambda & 0 & 1 & 0\\0 & 0 & 0 & 1\end{pmatrix}\mathbf{A} = \begin{pmatrix}1 & 2 & 3 & 4\\5 & 6 & 7 & 8\\\lambda 1 + 9 & \lambda 2 + 10 & \lambda 3 + 11 & \lambda 4 + 12\\13 & 14 & 15 & 16\end{pmatrix}$$

REMINDER: ELEMENTARY MATRICES / 4

The elementary matrix of type III results from $\boldsymbol{E} = \boldsymbol{I} + \lambda \boldsymbol{e}_i \boldsymbol{e}_j^\top (i \neq j)$.
For the example above:

$$\boldsymbol{E} = \boldsymbol{I_4} + \lambda \boldsymbol{e}_3 \boldsymbol{e}_1^\top = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$$

$\boldsymbol{E}^{-1}$ results from $\boldsymbol{E}^{-1} = \boldsymbol{I} - \lambda \boldsymbol{e}_i \boldsymbol{e}_j^\top$. Easy to check:

$$\boldsymbol{E}\boldsymbol{E}^{-1} = (\boldsymbol{I} + \lambda \boldsymbol{e}_i \boldsymbol{e}_j^\top)(\boldsymbol{I} - \lambda \boldsymbol{e}_i \boldsymbol{e}_j^\top) = \boldsymbol{I}^2 - \lambda^2 \boldsymbol{e}_i \boldsymbol{e}_j^\top \boldsymbol{e}_i \boldsymbol{e}_j^\top = \boldsymbol{I} - 0 \cdot \lambda^2 \boldsymbol{e}_i \boldsymbol{e}_j^\top = \boldsymbol{I},$$

since $\boldsymbol{e}_j^\top \boldsymbol{e}_i = 0$ for $i \neq j$.

REPETITION: PERMUTATION MATRIX

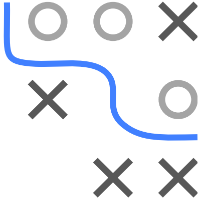
Permutation matrices contain exactly one 1 in each row and column, all other entries are 0.

Example:

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{e}_5 & \boldsymbol{e}_2 & \boldsymbol{e}_4 & \boldsymbol{e}_1 & \boldsymbol{e}_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Multiplying by a permutation matrix corresponds to one or more elementary transformations of type I.
Thus, an elementary matrix of type I is also a permutation matrix.

REMINDER: POSITIVE (SEMI-)DEFINITE



A symmetric matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ is **positive semi-definite** iff

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} \geq 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}.$$

Or equivalently: A matrix is positive semi-definite if all eigenvalues are non-negative.

A matrix is **positive-definite**, if the above equation can be rewritten with a "strict" greater than

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{0}.$$

or equivalently if all eigenvalues are positive.

REMINDER: ORTHOGONAL MATRICES

A matrix $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is **orthogonal** iff

$$\mathbf{Q}\mathbf{Q}^\top = \boldsymbol{I}.$$

The column vectors (or row vectors) of an orthogonal matrix are orthogonal to each other and normalized (to length 1). They form an orthonormal basis of $\mathbb{R}^n$.

The inverse of an orthogonal matrix is equal to its transpose, i.e.

$$\mathbf{Q}^{-1} = \mathbf{Q}^\top$$

Permutation matrices are orthogonal.

REMINDER: RANK OF A MATRIX

Definition: Rank of a matrix
A matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ has rank k if one of the following equivalent conditions is met:

- Maximum number of independent columns $= k$ (k-dimensional column space)
- Maximum number of independent rows $= k$ (k-dimensional row space)
- $\mathbf{A}$ can be factorized into matrices of rank k: $\mathbf{W} \in \mathbb{R}^{m \times k}$ and $\mathbf{H} \in \mathbb{R}^{k \times n}$

$$\mathbf{A} = \mathbf{W} \cdot \mathbf{H}$$