Vorlesung: MAT 460 Numerical Differential Equations

Version 21.10.14

Michael Karow, Andrea Dziubek

Lecture 3

Topics: Cholesky-factorization and Positive Definiteness

(At least) two very different algorithms for the Cholesky-factorization exist.

1. Cholesky-algorithm:

Input: symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$.

For
$$j = 1, \ldots, n$$

Compute
$$\ell_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} \ell_{jk}^2}$$

For
$$i = j + 1, ..., n$$

Compute
$$\ell_{ij} := \left(a_{ij} - \sum_{k=1}^{j-1} \ell_{ik}\ell_{jk}\right)/\ell_{jj}$$

Comment: This algorithm computes the elements of $L = [\ell_{ij}] \in \mathbb{R}^{n \times n}$

in the lower trianglular elements (including diagonal elements).

It works by equating coefficents.

Theorem of Cholesky-factorization

Let $A\in\mathbb{R}^{n\times n}$ be a **symmetric** and **positive definite** matrix. Then there exist exactly one lower triangular matrix $L\in\mathbb{R}^{n\times n}$ with positive diagonal elements, such that

$$A = L L^T$$
.

Example:

$$\underbrace{\begin{bmatrix} 25 & -5 & 15 \\ -5 & 10 & -15 \\ 15 & -15 & 29 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 5 & 0 & 0 \\ -1 & 3 & 0 \\ 3 & -4 & 2 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 5 & -1 & 3 \\ 0 & 3 & -4 \\ 0 & 0 & 2 \end{bmatrix}}_{L^{T}}.$$

Note: In contrast to the LU-factorization, the diagonal elements of the Cholesky-factor L are not necessarily 1.

Andre-Louis Cholesky (1875-1918) published Cholesky-factorization 1924.

Example for equating coefficients:

The ansatz

1

$$\begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} \ell_{11} & 0 & 0 \\ \ell_{21} & \ell_{22} & 0 \\ \ell_{31} & \ell_{32} & \ell_{33} \end{bmatrix} \begin{bmatrix} \ell_{11} & \ell_{21} & \ell_{31} \\ 0 & \ell_{22} & \ell_{32} \\ 0 & 0 & \ell_{33} \end{bmatrix}$$

reordering gives:

leads to the equations:

$$\begin{array}{lll} a_{11} &=& \ell_{11}^2 & & \ell_{11} &=& \sqrt{a_{11}} \\ a_{21} &=& \ell_{11} \ell_{21} & & \ell_{21} &=& a_{21}/\ell_{11} \\ a_{31} &=& \ell_{11} \ell_{31} & & \ell_{31} &=& a_{31}/\ell_{11} \\ a_{22} &=& \ell_{21}^2 + \ell_{22}^2 & & \ell_{22} &=& \ell_{21}^2 \\ a_{32} &=& \ell_{31} \ell_{21} + \ell_{32} \ell_{22} & & \ell_{32} &=& \ell_{31} \ell_{21})/\ell_{22} \\ a_{33} &=& \ell_{33}^2 + \ell_{31}^2 + \ell_{32}^2 & & \ell_{33} &=& \sqrt{a_{33} - \ell_{31}^2 - \ell_{32}^2} \end{array}$$

2. Cholesky-algorithm:

Input: symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$.

For $j = 1, \ldots, n$

Compute $a_{jj} := \sqrt{a_{jj}}$

For $i = j + 1, \dots, n$

Compute $a_{ij} := a_{ij}/a_{jj}$

For $k = j + 1, \dots, i$

Compute $a_{ik} := a_{ik} - a_{ij}a_{kj}$

Comment: This algorithm overwrites the entries in the lower triangular part of A (including the diagonal elements) with the entries of L.

In other words: After performing the algorithm we have

 $a_{ij} = \ell_{ij}$ for $i \ge j$.

It works by completing the square.

More at the end of this lecture.

Cholesky-algorithm, version 2:

Input: symmetric positive definite matrix $A \in \mathbb{R}^{n \times n}$.

For $j = 1, \ldots, n$

Compute $\ell_{ij} := \sqrt{a_{ij}}$

Für $i = j + 1, \ldots, n$

Compute $\ell_{ij} := a_{ij}/\ell_{jj}$

For $k = j + 1, \ldots, i$

Compute $a_{ik} := a_{ik} - \ell_{ij}\ell_{kj}$

Comment: The only difference between this and the last algorithm is: here the results are saved in the matrix $L = [\ell_{ij}]$ and not in A.

5

On the next slides we will answer the following questions:

- Why do only positive definite matrices have a Cholesky-factorization?
- What does positive definite actually mean?
- How can we determine definiteness?
- Why is positive definiteness important?
- Where do we find positive definite matrices in applications?

Quadratic forms und Definiteness

The scalar product as matrix product and the term x^TAx .

We know that the scalar product of two vectors $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \ y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ is defined as

$$x \cdot y = x_1 y_1 + \ldots + x_n y_n = \sum_{j=1}^n x_j y_j.$$

We can write the scalar product as matrix product: First, transpose the vector x, then multiply it with y according to the rules of matrix multiplication:

$$x^T y = [x_1 \dots x_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{j=1}^n x_j y_j = x \cdot y.$$

If we set y = Ax with any matrix $A = [a_{jk}] \in \mathbb{R}^{n \times n}$, we get

$$x^T A x = x^T y = \sum_{j=1}^n x_j y_j = \sum_{j=1}^n x_j \left(\sum_{k=1}^n a_{jk} x_k \right) = \sum_{j=1}^n \sum_{k=1}^n a_{jk} x_j x_k.$$

9

Example: quadratic forms in 4 variables

Let

$$x_{2}^{2} + 7x_{1}x_{2} - x_{4}^{2} = \begin{bmatrix} x_{1} \\ x_{2} \\ x_{3} \\ x_{4} \end{bmatrix}^{T} \begin{bmatrix} 0 & 3.5 & 0 & 0 \\ 3.5 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_{1} \\ x_{2} \\ x_{3} \\ x_{4} \end{bmatrix}$$

$$x_{1}^{2} + x_{2}^{2} + x_{3}^{2} + x_{4}^{2} = \begin{bmatrix} x_{1} \\ x_{2} \\ x_{3} \\ x_{4} \end{bmatrix}^{T} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1} \\ x_{2} \\ x_{3} \\ x_{4} \end{bmatrix}$$
$$= x^{T} I x \qquad I = \text{identity matrix}$$
$$= x^{T} x.$$

Quadratic Forms

Definition:

A quadratic form in the variables x_1, \ldots, x_n is an expression of the form

$$q_A(x) = x^T A x = \sum_{j=1}^n \sum_{k=1}^n a_{jk} x_j x_k,$$

where we can assume that the matrix $A \in \mathbb{R}^{n \times n}$ is symmetric.

Example: a quadratic form in two variables

Let
$$A = \begin{bmatrix} 2 & -2 \\ 8 & 4 \end{bmatrix}$$
 (not symmetric). Then

$$\begin{aligned} q_A(x) &= x^T A \, x \\ &= \left[x_1 \quad x_2 \right] \begin{bmatrix} 2 & -2 \\ 8 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= 2 \, x_1^2 - 2 \, x_1 x_2 + 8 \, x_2 x_1 + 4 \, x_2^2 \\ &= 2 \, x_1^2 + 3 \, x_1 x_2 + 3 \, x_2 x_1 + 4 \, x_2^2 \\ &= \left[x_1 \quad x_2 \right] \underbrace{ \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix} }_{=:\widetilde{A}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x^T \widetilde{A} \, x = q_{\widetilde{A}}(x). \end{aligned}$$

The matrix \widetilde{A} is symmetric.

Definitness of quadratic forms

A quadratic form (i. e. the corresponding matrix) is c d

positive definite if $x^TAx > 0$ for all $x \neq 0$, positive semidefinite if $x^TAx \geq 0$ for all x, negative definite if $x^TAx < 0$ for all $x \neq 0$, negative semidefinite if $x^TAx \leq 0$ for all $x \neq 0$, indefinite otherwise.

Note: A definite matrix is always semi-definite.

Example: Let A be a diagonal matrix:

$$A = \begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & a_{nn} \end{bmatrix}$$

Then

$$x^T A x = a_{11} x_1^2 + a_{22} x_2^2 + \ldots + a_{nn} x_n^2$$

A is positive definite, if and only if all diagonal elements are positive.

A is positive semi-definite, if and only if all diagonal elements are non negative.

A is indefinite, if and only if 2 diagonal elements have different signs.

10

are non ne

A necessary criteria for positive definiteness

If a symmetric matrix $A = [a_{jk}]$ is positive definite, than all its diagonal elements are positive.

Proof: Let

$$e_k = egin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \leftarrow & k ext{-te Zeile} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

It is easy to see, that $e_k^T A e_k = a_{kk}$, where a_{kk} is the k-th diagonal element of A. If A is positive definite, then

$$a_{kk} = e_k^T A e_k > 0$$

Practical example: The matrix

$$A = \begin{bmatrix} 2 & 3 & 0 & -6 \\ 3 & 7 & 4 & 8 \\ 0 & 4 & -2 & 7 \\ -6 & 8 & 7 & 1 \end{bmatrix}$$

is not positive definite, because $a_{33} = -2 \le 0$.

13

Motivation: Where do we find positive definite matrices in applications?

1. Mechanics

Linearised differential equation of a mechanical system with degrees of freedom $u(t) \in \mathbb{R}^n$:

$$M \ddot{u}(t) + D \dot{u}(t) + S u(t) = f(t).$$
 (*)

where

M = mass matrix D = damping matrix S = stiffness matrix

f = external forces (load)

From (*) we get the energy balance equation:

Definite quadratic forms:

Kinetic energy: $E_{kin} = \frac{1}{2} \dot{u}^T M \dot{u} > 0$

Elastic energy: $E_{elast} = \frac{1}{2}u^T Su \ge 0$

Energy dissipation: $\dot{u}^T D \dot{u} \geq 0$ (friction)

 $\frac{d}{dt}(E_{kin} + E_{elast}) = \dot{u}^T f - \dot{u}^T D \dot{u}.$

Warning:

We can not decide from only looking at the diagonal elements if a matrix is positive definite.

Example: The matrix $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$ has positive diagonal elements, but A is not positive.

Explanation: In general

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x^2 + 6x_1x_2 + 2x_2^2.$$

If we choose $x_1 = 1$, $x_2 = -1$, then

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}^T \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2 * 1^2 + 6 * 1 * (-1) + 2 * (-1)^2 = -2 < 0.$$

Practical example: A spring-mass system (here without damping)

Relaxed State

System in Motion m_1 m_2 m_3 m_3 m_4 Force

Energies: $(s_j = \text{spring constants})$

$$E_{elast} = \frac{s_1}{2}u_1^2 + \frac{s_2}{2}(u_2 - u_1)^2 + \frac{s_3}{2}(u_3 - u_2)^2 = \frac{1}{2} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}^T \begin{bmatrix} s_1 + s_2 & -s_2 & 0 \\ -s_2 & s_2 + s_3 & -s_3 \\ 0 & -s_3 & s_3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$$

$$E_{kin} = \frac{m_1}{2} \dot{u}_1^2 + \frac{m_2}{2} \dot{u}_2^2 + \frac{m_3}{2} \dot{u}_3^2 = \frac{1}{2} \begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \\ \dot{u}_3 \end{bmatrix}^T \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} \begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \\ \dot{u}_3 \end{bmatrix}$$

Equation of motion:

$$\underbrace{ \begin{bmatrix} m_1 & & \\ & m_2 & \\ & & m_3 \end{bmatrix}}_{M} \underbrace{ \begin{bmatrix} \ddot{u}_1(t) \\ \ddot{u}_2(t) \\ \ddot{u}_3(t) \end{bmatrix}}_{S} + \underbrace{ \begin{bmatrix} s_1 + s_2 & -s_2 & 0 \\ -s_2 & s_2 + s_3 & -s_3 \\ 0 & -s_3 & s_3 \end{bmatrix}}_{S} \underbrace{ \begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{bmatrix}}_{U} = \begin{bmatrix} 0 \\ 0 \\ f(t) \end{bmatrix}$$

2. Discretised heat equation (important for projects)

Stationary heat equation in domain Ω :

$$-\operatorname{div}(\lambda \, \nabla T) = f$$

Ansatz for temperature field:

$$T(x) = u_1 \phi_1(x) + u_2 \phi_2(x) + \ldots + u_n \phi_n(x)$$

with

 u_k = temperatures at the vertices,

 ϕ_k = hat functions.

Discretised heat equation:

$$Su = p$$

with $u = [u_1 \ u_2 \ \dots u_n]^T$ and with the <u>"stiffness matrix"</u>

$$S = [s_{jk}] \in \mathbb{R}^{n \times n}, \qquad s_{jk} = \int_{\Omega} \lambda (\nabla \phi_j)^T \nabla \phi_k$$

We have

$$u^{T}Su = \int_{\Omega} \lambda \|\nabla(u_{1}\phi_{1} + u_{2}\phi_{2} + \ldots + u_{n}\phi_{n})\|^{2} \geq 0.$$

 \Rightarrow S is semi-definite. (Even definite after including the boundary conditions.)

Eigenvalue criteria for definiteness

Remember: $\lambda \in \mathbb{C}$ is called eigenvalue of $A \in \mathbb{R}^{n \times n}$, if $0 \neq v \in \mathbb{R}^n$, such that $Av = \lambda v$

Theorem: Let $A \in \mathbb{R}^{n \times n}$ symmetric. Then:

- 1. All eigenvalues of A are real.
- 2. The eigenvectors can be choosen such that they form an orthonormal basis of \mathbb{R}^n .
- 3. The matrix *A* is positive (semi)definite if and only if all eigenvalues are positive (non negative).

Explanation of the eigenvalue criteria:

Let $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$ be an orthonormal basis of eigenvalues:

$$Av_j = \lambda_j v_j,$$
 $v_j^T v_k = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$

Every vector $x \in \mathbb{R}^n$ can be written as:

$$x = x_1 v_1 + x_2 v_2 + \ldots + x_n v_n, \quad x_i \in \mathbb{R}.$$

Straightforward computation gives: $x^T A x = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \ldots + \lambda_n x_n^2$

3. Analysis: Lokal extrema

Given: Function $f: \mathbb{R}^n \to \mathbb{R}$.

Taylor expansion about $x \in \mathbb{R}^n$ up to 2nd order:

$$f(x+h) = f(x) + \sum_{k=1}^{n} \frac{\partial f}{\partial x_{k}}(x) h_{k} + \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial^{2} f}{\partial x_{j} \partial x_{k}}(x) h_{j} h_{k} + o(\|h\|^{2})$$
$$= f(x) + \nabla f(x)^{T} h + \frac{1}{2} h^{T} H_{f}(x) h + o(\|h\|^{2}),$$

with the Hessian-matrix

$$H_f(x) = \left[\frac{\partial^2 f}{\partial x_j \partial x_k}(x)\right]_{jk}.$$

Necessary for local minimum at x:

$$\nabla f(x) = 0$$
, $H_f(x)$ positive semi-definite.

Sufficient for local minimum at x:

$$\nabla f(x) = 0$$
, $H_f(x)$ positive definite.

Check for definiteness by completing the square

A quadratic form in 2 variables can be transformed into a sum or difference of squares by completing the square.

(Method:
$$a^2 + 2ab = a^2 + 2ab + b^2 - b^2 = (a+b)^2 - b^2$$
.)

Example 1:

17

$$9 x_1^2 - 24 x_1 x_2 + 41 x_2^2 = (3x_1)^2 + 2(3x_1)(-4x_2) + 41 x_2^2$$

$$= (3x_1)^2 + 2(3x_1)(-4x_2) + (-4x_2)^2 - (-4x_2)^2 + 41 x_2^2$$

$$= (3x_1 - 4x_2)^2 - (-4x_2)^2 + 41 x_2^2$$

$$= (3x_1 - 4x_2)^2 + (5x_2)^2.$$

 \Rightarrow This form is positive definite.

Example 2:

$$36 x_1^2 + 4 x_1 x_2 = (6 x_1)^2 + 2 (6 x_1) (\frac{1}{3} x_2)$$

$$= (6 x_1)^2 + 2 (6 x_1) (\frac{1}{3} x_2) + (\frac{1}{3} x_2)^2 - (\frac{1}{3} x_2)^2$$

$$= (6 x_1 + \frac{1}{3} x_2)^2 - (\frac{1}{3} x_2)^2$$

⇒ This form is indefinite.

To see this, choose e. g. $x_2 = 1$ and x_1 such that $6x_1 + \frac{1}{3}x_2 = 0$.

2×2 matrices in general (for $a \neq 0$):

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^T \underbrace{\begin{bmatrix} a & b \\ b & c \end{bmatrix}}_{=A} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a x_1^2 + 2b x_1 x_2 + c x_2^2$$

$$= a \left(x_1^2 + 2 \frac{b}{a} x_1 x_2 \right) + c x_2^2$$

$$= a \left(x_1^2 + 2 \frac{b}{a} x_1 x_2 + \left(\frac{b}{a} \right)^2 x_2^2 \right)$$

$$- \frac{b^2}{a} x_2^2 + c x_2^2$$

$$= a \left(x_1 + \frac{b}{a} x_2 \right)^2 + \left(c - \frac{b^2}{a} \right) x_2^2$$

$$= a \left(x_1 + \frac{b}{a} x_2 \right)^2 + \frac{1}{a} \underbrace{\left(ac - b^2 \right)}_{=\det(A)} x_2^2$$

Conclusion: A is positive definite if and only if

$$a > 0$$
 and $det(A) > 0$.

21

Relation between completing the square and Cholesky-factorization

Example:

$$q_A(x) = x^T A x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}^T \begin{bmatrix} 25 & -5 & 15 \\ -5 & 10 & -15 \\ 15 & -15 & 29 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$
$$= 25x_1^2 - 10x_1x_2 + 30x_1x_3 + 10x_2^2 - 30x_2x_3 + 29x_3^2$$

Completing the square we get:

$$q_A(x) = (5x_1)^2 + 2(5x_1)(-x_2 + 3x_3) + (-\mathbf{x}_2 + 3\mathbf{x}_3)^2 - (-\mathbf{x}_2 + 3\mathbf{x}_3)^2 + 10x_2^2 - 30x_2x_3 + 29x_3^2$$

$$= (5x_1 - x_2 + 3x_3)^2 + \text{ quad. form in } x_2, x_3$$

$$\vdots$$

$$= (5x_1 - x_2 + 3x_3)^2 + (3x_2 - 4x_3)^2 + (2x_3)^2$$

$$= y_1^2 + y_2^2 + y_3^2 = y^T y$$

with

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 5x_1 - x_2 + 3x_3 \\ 3x_2 - 4x_3 \\ 2x_3 \end{bmatrix} = \underbrace{\begin{bmatrix} 5 & -1 & 3 \\ 0 & 3 & -4 \\ 0 & 0 & 2 \end{bmatrix}}_{=:T} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}$$

In summary: $x^T A x = y^T y = (L^T x)^T (L^T x) = x^T L L^T x.$ \Rightarrow $A = L L^T.$

The determinant criteria for positive definiteness

The criteria from the last slide for positive definiteness of a 2×2 matrix can be generalized.

Theorem: A symmetric matrix $A = [a_{jk}] \in \mathbb{R}^{n \times n}$ is positive definite if and only if the determinants of all upper left sub-matrices are positive, i. e.

$$\det[a_{11}] = a_{11} > 0, \quad \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} > 0, \quad \det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} > 0, \quad \dots \quad \det(A) > 0.$$

22

Why do only positive definite matrices have a Cholesky-factorization?

In general: Let $B \in \mathbb{R}^{n \times n}$ not singular (invertible). Then the matrix $A = B^T B$ is positive definite.

Explanation:

Let $x \in \mathbb{R}^n$, $x \neq 0$. Choose y = Bx. Then also $y \neq 0$, becaues B is not singular. And further:

$$x^{T}Ax = x^{T}B^{T}Bx = (Bx)^{T}Bx = y^{T}y = y_{1}^{2} + y_{2}^{2} + \dots + y_{n}^{2} > 0.$$

 \Rightarrow A is positive definite.

Conclusion: If $A = LL^T = (L^T)^T \underbrace{(L^T)}_B$ and $\det(L) \neq 0$, then A is positive definite.

Cholesky-factorization with MATLAB

U=chol(A) gives back a upper triangular matrix U, such that U'*U=A.

L=chol(A,'lower') gives back a lower triangular matrix L, such that L*L'=A.

If ${\tt A}$ is not symmetric and not positive definite, an error message is returned.