LINMA2380 - Matrix Computations

Homework 1

Final Version

Antoine Defosse (2775 2000)
Théo Hanon (3732 2000)
Thibault Lootvoet (1909 2000)
Laurian Paul (3243 2000)

16 October 2023

1 A1

We will make the proof by contradiction. We assume that the matrix

$$A = \begin{bmatrix} 0 & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \text{ with det(A)} \neq 0 \tag{1}$$

has a LU factorization.

Since in this homework we focus on the case without permutation, we can thus write $A = LU$ where L is a lower triangular matrix with diagonal elements equal to 1 and U is an upper triangular matrix. Its LU decomposition should look like :

$$\begin{bmatrix} 0 & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \tag{2}$$

From this system, we can extract the following equation :

$$0 = l_{11}u_{11} \tag{3}$$

Then at least one of l_{11} and u_{11} has to be zero, which implies that either L or U is singular. This is impossible if A is non singular. It contradicts our hypothesis that A was non singular ($det(A) \neq 0$). □

2 A2

Assumption: Let A be a nonsingular matrix with two distinct LU factorizations:

$$A = L_1U_1, \text{ where } L_1 \text{ is unit lower triangular and } U_1 \text{ is upper triangular,}$$
$$A = L_2U_2, \text{ where } L_2 \text{ is unit lower triangular and } U_2 \text{ is upper triangular.}$$

As we know that A is non-singular, L_1, L_2, U_1, U_2 are all invertible, we get the following:

$$L_2^{-1}L_1 = U_2U_1^{-1} \tag{4}$$

We know that L_1 and L_2 are unit lower triangular by definition of the LU factorization thus $L_2^{-1}L_1$ is also unit lower triangular. We also know that U_1 and U_2 are upper triangular by definition of the LU factorization thus $U_2U_1^{-1}$ is also upper triangular.

The only way to satisfy Equation 4 is that $L_2^{-1}L_1$ and $U_2U_1^{-1}$ are both identity matrices and thus that $L_1 = L_2$ and $U_1 = U_2$. □

3 A3

We will make the proof by induction on the order n of the matrix :

- **Base case** $n = 2$:
  Let

  $$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \tag{5}$$

  We can easily construct its LU decompostion

  $$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ c/a & 1 \end{bmatrix} \begin{bmatrix} a & b \\ 0 & d - \frac{bc}{a} \end{bmatrix}. \tag{6}$$

  Which is well defined as $m_1 = a \neq 0$. We can verify that

  $$\begin{cases} U_{11} & = m_1 & = a \\ U_{22} & = \frac{m_2}{m_1} & = \frac{ad-bc}{a} \end{cases}$$

- **Inductive step** $n = k + 1$:
  We assume that the assumption holds for $n = k$ and we verify it for $n = k + 1$. Let

$$A_{k+1} = \begin{bmatrix} A_k & z \\ w^T & a \end{bmatrix}. \tag{7}$$

Where $A_k \in \mathbb{R}^{k \times k},\ w^T \in \mathbb{R}^{1 \times k},\ z \in \mathbb{R}^{k \times 1},\ a \in \mathbb{R}$. As previously we can construct the LU decomposition of A_{k+1}

$$\begin{bmatrix} A_k & z \\ w^T & a \end{bmatrix} = \begin{bmatrix} L_k & 0 \\ u^T & 1 \end{bmatrix} \begin{bmatrix} U_k & v \\ 0 & x \end{bmatrix} \tag{8}$$

and we will choose $u^T \in \mathbb{R}^{1 \times k}, v \in \mathbb{R}^{k \times 1}$ and $x \in \mathbb{R}$ such that the following equality holds :

$$\begin{bmatrix} A_k & z \\ w^T & a \end{bmatrix} = \begin{bmatrix} L_k U_k & L_k v \\ u^T U_k & x + u^T v \end{bmatrix} \tag{9}$$

Since $A_k = L_k U_k$ by hypothesis, we have to solve these equations:

$$\begin{cases} L_k v = z \\ u^T U_k = w^T \\ a = x + u^T v \end{cases} \Leftrightarrow \begin{cases} v = L_k^{-1} z \\ u^T = w^T U_k^{-1} \\ x = a - u^T v \end{cases} \tag{10}$$

Our solutions v and u^T exist as $diag(L) = \{1, ..., 1\}$ and $diag(U) = \left\{ m_1, \frac{m_2}{m_1}, ..., \frac{m_k}{m_{k-1}} \right\}$ where by assumption $m_1, ..., m_k \neq 0$. Finally we have to verify that $x = \frac{m_{k+1}}{m_k}$. To do so, we develop our expression of x using (10)

$$\begin{aligned} x &= a - u^T v & (11) \\ &= a - w^T U_k^{-1} L_k^{-1} z & (12) \\ &= a - w^T A_k^{-1} z & (13) \end{aligned}$$

We can now apply the determinant's formula for block matrices on (7)which gives

$$det(A_{k+1}) = det(A_k)(a - w^T A_k^{-1} z) \tag{14}$$

Combining (13) and (14) we get

$$\begin{aligned} \frac{det(A_{k+1})}{det(A_k)} &= a - w^T A_k^{-1} z & (15) \\ \Leftrightarrow \frac{m_{k+1}}{m_k} &= x & (16) \end{aligned}$$

This claims our proof! □

4 A4

We can start from (A3) indeed the assumptions are respected here. Therefore, we know that we can write

$$A = LU \tag{17}$$

where

$$U = \begin{bmatrix} m_1 & \cdots & u_{1k} & \cdots & u_{1n} \\ & \ddots & \vdots & & \vdots \\ & & \frac{m_k}{m_{k-1}} & \cdots & u_{kn} \\ & & & \ddots & \vdots \\ & & & & \frac{m_n}{m_{n-1}} \end{bmatrix}. \tag{18}$$

From that the proof is straight forward, we can write U as the product of a diagonal matrix D and another upper triangular matrix $\bar{U}$ such that

$$U = \underbrace{\begin{bmatrix} m_1 & \cdots & 0 & \cdots & 0 \\ & \ddots & & & \vdots \\ & & \frac{m_k}{m_{k-1}} & & 0 \\ & & & \ddots & \vdots \\ & & & & \frac{m_n}{m_{n-1}} \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} 1 & \cdots & \frac{u_{1k}}{m_1} & \cdots & \frac{u_{1n}}{m_1} \\ & \ddots & & & \vdots \\ & & 1 & & \frac{u_{kn}m_{k-1}}{m_k} \\ & & & \ddots & \vdots \\ & & & & 1 \end{bmatrix}}_{\bar{U}} \tag{19}$$

Finally we can write,

$$A = LD\bar{U} \tag{20}$$

and the proof is over ! □

5 A5

We are going to make the proof in 2 steps by prooving the implication in both sense:

- $\{m_1, ..., m_n\} \neq 0 \Rightarrow \exists LU$
  We can see that the assumptions made in (A3) are respected. This means that the LU decomposition exists and it claims our proof.

- $\exists LU \Rightarrow \{m_1, ..., m_n\} \neq 0$
  We know from our hypothesis that

  $$A = LU \tag{21}$$

  and observe that

  $$A_k = L_k U_k \quad k = 1, ..., n \tag{22}$$

  where A_k, L_k and U_k are respectively the kth leading principal submatrix of A, L and U. Therefore, we have

  $$m_k = det(L_k)det(U_k) \quad k = 1, ..., n \tag{23}$$

  With (23) and the fact that $det(L_k) = 1$, we have

  $$det(A) = det(U) \neq 0 \;\Rightarrow\; U_{ii} \neq 0 \text{ for } i = 1, ..., n. \tag{24}$$

  Finally combining (23) and (24) we get

  $$m_k = det(U_k) \neq 0 \quad k = 1, ..., n \tag{25}$$

  This finishes our proof! □

6 A6

In our case, we want to prove that U is given by $U = L^T$ with the implication that the factorization can be written as $A = LDL^T$.

1. First we know that A can be factorized as $A = LDU$, where
   - A is a symmetric and nonsingular matrix
   - L is a unit lower-triangular matrix
   - U is a unit upper-triangular matrix
   - D is a diagonal matrix

2. Since A is symmetric, we have $A = A^T$: the transpose of A is equal to A.

3. Now, let's compute the transpose of the factorization of A :

$$\begin{aligned} & A^T = (LDU)^T \\ \Leftrightarrow \;& A^T = U^T D^T L^T \end{aligned}$$

4. Since D is a diagonal matrix, D^T is equal to D. Our equation becomes :

$$A^T = U^T D L^T$$

5. By indentification of the equation $A = A^T$, we can see that :

$$LDU = U^T D L^T$$

This leads to the following two conditions which must be met simultaneously

$$\begin{cases} L = U^T \\ U = L^T \end{cases}$$

6. In the end, we came across exactly what we wanted to prove :

$$U = L^T$$

We can now rewrite the decomposition of A in two equivalent ways :

$$A = U^T D U$$

$$\boxed{A = LDL^T}$$

This concludes our demonstration □

7 A7

For this question we are going to proof the first implication by contraposition and the other one directly:

- The negation of the second property leads to 2 cases :
  1. $\exists k \; s.t. \; m_k = 0$.
     Thus it means that $\exists x \neq 0 \; s.t.$
     $$Ax = 0 \tag{26}$$
     this implies that $x^T A x = 0$ which means that the first property cannot isn't respected and this proof the contraposition.
  2. $\exists k \; s.t. \; m_k < 0$
     We can combine (A6) and (A4) since the assumptions are respected. Therefore one can write:
     $$A = LDL^T \tag{27}$$
     $$\Longrightarrow x^T A x = x^T L D L^T x \tag{28}$$
     Let $z = L^T x$, it leads to
     $$x^T A x = z^T D z \tag{29}$$
     $$= z_1^2 d_1 + \cdots + z_n^2 d_n. \tag{30}$$
     We know that $d_k = \frac{m_k}{m_{k-1}}$, thus:
     (a) If $d_k < 0$, we can take $z = e_k$ (e.g. $e_1 = [1, 0, \; \dots \; , 0]$) and therefore $z^T D z < 0$.

(b) If $d_k > 0$, it means that $m_{k-1} < 0$. We can then look at d_{k-1}, if $d_{k-1} < 0$ we do the same that the previous point, otherwise we look at d_{k-2} and so on. We can continue this process until $d_2 > 0 \implies m_1 < 0$ and we can finally take $z = e_1$.

Once again we show the negation of the first property and we have covered all the cases.

- As all the leading minors are strictly positive we can write using (A4) and (A6)

$$A = LDL^T. \tag{31}$$

Let $z = L^T x$ which gives this quadratic form $z^T Dz$. D is a diagonal matrix thus all its eigenvalues are on the diagonal. Therefore we get that $\lambda_1, ..., \lambda_n > 0$. Combining this with the Inertia's Principle of Sylvester, we get

$$z^T Dz > 0 \tag{32}$$
$$\iff x^T Ax > 0. \tag{33}$$

We have shown the last implication and this concludes the entire proof. $\square$

8 A8

1. If A has a Cholesky factorization, then A is positive definite:

Suppose A admits a Cholesky factorization LL^T, to show that A is positive definite, we need to demonstrate that $x^T Ax > 0$ for any nonzero vector x.

Let x be a nonzero vector, and consider the quadratic form:

$$\begin{aligned} x^T Ax &= x^T (LL^T)x \\ &= (L^T x)^T (L^T x) \\ &= \|L^T x\|^2 \\ &\geq 0 \end{aligned}$$

This result allows us to say that A is semidefinite positive.

As we assume that L is a lower-triangular matrix with positive diagonal entries, L is invertible and
$Lx = 0 \iff x = 0$ thus A is positive definite.

2. If A is positive definite, then it has a Cholesky factorization :

From (A7) we can say that all leading principal minors of A are positive. This in turn implies that A admits a LDU factorization (A5). We can then say that A admits a LDL^T factorization (A6) as the matrix is symmetric and nonsingular. Let us define $\tilde{L} = LD^{1/2}$. We then have $\tilde{L}^T = (D^{1/2})^T L^T$ and the following :

$$\begin{aligned} \tilde{L}\tilde{L}^T &= LD^{1/2}(D^{1/2})^T L^T \\ &= LDL^T \\ &= A \end{aligned}$$

This proves that A has a Cholesky factorization. $\square$

9 A9

Here we want to illustrate by a counterexample that the condition of all leading principal minors of a symmetric matrix $A \in \mathbb{R}^{n\times n}$ being non negative , is not sufficient to establish that A is positive semidefinite.

For the case $n = 1$, we can trivially see that the condition is sufficient since the case $n = 1$ corresponds to the scalar case and that therefore the only minor corresponds to the element which is semi-positive according to the condition.

Let's now look at what happens in the case of $n = 2$. Without loss of generality, we can represent 2×2 symmetric matrices as :

$$A = \begin{pmatrix} \alpha & \gamma \\ \gamma & \beta \end{pmatrix}$$

Where, we can deduce the conditions for all minors being non negative as follows :

$$\begin{cases} m_1 = \alpha \geq 0 \\ m_2 = det\,(A) = \alpha\beta - \gamma^2 \geq 0 \end{cases}$$

Now all we need to do is find the α, β, γ parameters that meet our conditions above and make our defined A matrix semi-negative, negative or indefinite. This means restricting our eigenvalues respectively, such as :

$$\lambda_i \leq 0$$
$$\lambda_i < 0$$
$$sign(\lambda_1) \neq sign(\lambda_2)$$

Now, by taking $\alpha = 0, \beta = -1, \gamma = 0$ wich leads to the following matrix :

$$A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$$

We can see that our example is valid since $m_1 = 0 \geq 0$, $m_2 = 0 \geq 0$ and that the matrix is semi-negative because of his eigenvalues :

$$\lambda_1 = -1 \quad ; \quad \lambda_2 = 0$$
$$\Leftrightarrow \ \lambda_i \leq 0$$

Which concludes our illustration □

10 A10

We are going to make the proof in 2 steps by proving the implication in both sense:

1. $A \succeq 0$ (Proposition A) $\implies$ all principal minors of A are non-negative (Proposition B)

   Let $\mathcal{S} = \{i_1, i_2, ..., i_k\}$ where $1 \leq i_1 < i_2 < \ldots < i_k \leq n$, we know that $A[\mathcal{S}, \mathcal{S}]$ is a principal submatrix because it is obtained by keeping the k rows indexed by these values along with their corresponding columns indexed by the same values.
   Let's take the vector z such that

   $$z_i = \begin{cases} x_i & \text{if } i \in \mathcal{S} \\ 0 & \text{if } i \notin \mathcal{S} \end{cases}$$

   with $x_i \in \mathbb{R}$ and we denote $x \in \mathbb{R}^k$ the vector z without the zeros. We can remark that $z^\top A z = x^\top A[\mathcal{S}, \mathcal{S}] x \geq 0$. And this holds $\forall x \in \mathbb{R}^k$ as $A \succeq 0$. Therefore, for every principle submatrix $B = A[\mathcal{S}, \mathcal{S}] \in \mathbb{R}^{k \times k}$ we can write:

   $$x^T B x \geq 0 \quad , \forall x \in \mathbb{R}^k. \tag{34}$$

   We may now conclude that, as every submatrix of A is positive semidefinite their determinant cannot be negative.

2. (Proposition B) All principal minors of A are non-negative $\implies$ $A \succeq 0$ (Proposition A).

   We will prove that $\neg(A) \implies \neg(B)$. Before starting, we introduce the following **Lemma** 1, it will be crucial for the rest of the proof.

   Lemma 1. [1] *Let $A \in \mathbb{R}^{n \times n}$, a symmetric matrix and two eigenvalues λ_k, λ_{k+1} of A with $\lambda_1 \leq ... \leq \lambda_n$ and $k < n$. Any principal submatrix $B \in \mathbb{R}^{(n-1)\times(n-1)}$ of A has one eigenvalue γ s.t. $\lambda_k \leq \gamma \leq \lambda_{k+1}$.*

[1] The proof of this Lemma is an adaptation of the proof of the Cauchy's interlacing theorem given on `https://www.cmor-faculty.rice.edu/~caam440/chapter2.pdf`, page 49

Proof. Consider $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{(n-1)\times(n-1)}$ a principal submatrix of A. We represent the situation without loss of generality such as

$$A = \begin{bmatrix} B & u \\ u^T & d \end{bmatrix} \tag{35}$$

where $u \in \mathbb{R}^{n-1}$ and $d \in \mathbb{R}$. We introduce the following vector spaces:

$$V = span\,(q_k, ..., q_n) \qquad\qquad W = span\,(w_1, ..., w_k)$$

with $q_k, ..., q_n \in \mathbb{R}^n$ are eigenvectors of A and $w_1, ..., w_k \in \mathbb{R}^{n-1}$ eigenvectors of B with $\gamma_1 \le ... \le \gamma_k \le ... \le \gamma_n$. One can notice that we can find $v \in V$ and $\tilde{w} \in W$ such as

$$v = \begin{pmatrix} \tilde{w} \\ 0 \end{pmatrix}$$

and therefore we can obtain

$$\frac{v^T A v}{v^T v} = \frac{\tilde{w}^T B \tilde{w}}{\tilde{w}^T \tilde{w}}. \tag{36}$$

Finally we can bound our expression with the fact that for $v \in V$ and $\tilde{w} \in W$, the maximum and minimum values of the quadratic forms are in fact respectively γ_k and λ_k

$$\lambda_k = min_{v \in V} \quad \frac{v^T A v}{v^T v} \le \frac{v^T A v}{v^T v} = \frac{\tilde{w}^T B \tilde{w}}{\tilde{w}^T \tilde{w}} \le max_{\tilde{w} \in W} \quad \frac{\tilde{w}^T B \tilde{w}}{\tilde{w}^T \tilde{w}} = \gamma_k. \tag{37}$$

(38)

Similary we now take the following vector spaces:

$$V = span\,(q_1, ..., q_{k+1}) \qquad\qquad W = span\,(w_k, ..., w_n) \tag{39}$$

and once again we can find some $v \in V$ and $\tilde{w} \in W$

$$\frac{v^T A v}{v^T v} = \frac{\tilde{w}^T B \tilde{w}}{\tilde{w}^T \tilde{w}}. \tag{40}$$

But this time we need to take care because the sets V and W are different. Indeed we can lower bound the right with γ_k and we can upper bound the left with λ_{k+1}

$$\lambda_{k+1} = max_{v \in V} \quad \frac{v^T A v}{v^T v} \ge \frac{v^T A v}{v^T v} = \frac{\tilde{w}^T B \tilde{w}}{\tilde{w}^T \tilde{w}} \ge min_{\tilde{w} \in W} \quad \frac{\tilde{w}^T B \tilde{w}}{\tilde{w}^T \tilde{w}} = \gamma_k. \tag{41}$$

(42)

Finally we can write,

$$\lambda_k \le \gamma_k \le \lambda_{k+1} \tag{43}$$

□

Lets start the proof, by contraposition we assume that $\exists x$ s.t. $x^\top A x < 0$ meaning that $\exists \lambda(A) < 0$. Let $r \in \mathbb{N}$ be the rank of the matrix, we first reduce A by keeping $n - r = k$ independent row/column and we denote this new matrix $\bar{A}$. We can observe that by construction $det(\bar{A}) \neq 0$.

Using **Lemma** 1 we can conclude that $\bar{A}$ must have at least one negative eigenvalue. Indeed each time we remove one row/column of A, at least one new eigenvalue γ of the reduced matrix should be s.t. $0 > \gamma \ge \lambda_-$ with λ_- a negative eigenvalue of A. The left inequality is strict otherwise it means that we keep nonindependent row/columns. Therefore we have 2 cases:

(a) $det(\bar{A}) < 0 \implies$ we found a negative principal minor

(b) $det(\bar{A}) > 0$ means that we have an even number of negative eigenvalues. We will find the negative principle minor by induction on the order i of the matrix $\bar{A}$:

- $i = 2$
  By hypothesis, $det(\bar{A}) = \lambda_1\lambda_2 = ad - bc > 0$ and $tr(\bar{A}) = \lambda_1 + \lambda_2 = a + d < 0$
  We see that either a or b should be negative and therefore there exists at least one negative principal minor.

- $i = k + 1$
  Assume $B \in \mathbb{R}^k$ to be a principal submatrix of $\bar{A}$. As $\bar{A}$ is full rank we can find B s.t. $det(B) \neq 0$ by keeping k independent rows/columns. Since $\bar{A}$ has at least two negative eigenvalues (by hypothesis), we have with our **Lemma** 1 that B must have at least one negative eigenvalue. Thus, we can conclude :
  - If $det(B) < 0 \implies$ We found a negative principle minor of $\bar{A}$
  - If $det(B) > 0 \implies$ We have an even number of negative eigenvalues. Then, as $B \in \mathbb{R}^{k\times k}$ by induction hypothesis we can find a negative principal minor inside B.

Every principle submatrix of $\bar{A}$ is also a principle submatrix of A. Therefore we just prove $\neg$ (B) since we can always find a negative principle minor in A. $\square$