TMA 4215 Numerical Mathematics: Lecture 02

André Massing (`andre.massing@ntnu.no`)

Aug 22, 2019

Notice.

Before we start: We (as in you :)) need/want to set-up a **reference group** of 2-3 students (favorably with one exchange student) to ensure a (semi-continuous) quality assessment of the course based on the students' feedback. Reference group and I will meet at least 3 times during the course, and the reference group will be asked to hand in a final report.

The is an important duty, so please come forward with some volunteers as soon as possible!

1 Linear System

Notice.

We expect that you are familiar with fundamental concepts from linear algebra. For a quick recall, skim through Ch.1 in **YEB**, in particular Section 1.1-1.5, 1.7,1.8, 1.10.

Notice.

Material in Lecture 1 and 2 are built upon

- Sections 3.3.1 in **YEB**
- Section 2.1 - 2.3 in **BLUB**

Also, the first two pages (in the Lecture 02 PDF document) are from Lecture 01 to quickly recall where we stop last time.

This chapter is concerned with solving linear systems of the form

$$\left.\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{1n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n & = & b_n \end{array}\right\} \tag{1}$$

As usual, we denote by $\mathbb{R}^{m\times n}$ the set of $m \times n$ matrices with entries in $\mathbb{R}$. The the system of linear equations can be written in matrix form

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{n1} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}}_{=:A} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{:=\boldsymbol{x}} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}}_{:=\boldsymbol{b}} \tag{2}$$

or more compact, $A\boldsymbol{x} = \boldsymbol{b}$ where $A \in \mathbb{R}^{n\times n}$ is now a *square matrix*, $\boldsymbol{x}$ is the vector of unknows, and $\boldsymbol{b}$ is the right-hand side vector given by the data.

1.1 Cramer's Rule: A Way to Solve Linear Systems Numerically?

You have probably learned about one way to solve the linear system (2) a basic course on linear algebra, namely **Cramer's rule**,

$$x_i = \frac{D_i}{D} \text{ for } i = 1, \dots, n, \tag{3}$$

where $D = \det(A)$ and D_i is the determinant of the matrix obtained by replacing the i-th column in A by $\boldsymbol{b}$.

Exercise 1: Computing the inverse A^{-1} using Cramer's rule

Show that inverse matrix A^{-1} of A is given by

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} A_{11} & A_{21} & \dots & A_{1n} \\ A_{12} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix} \tag{4}$$

where $A_{ij} := (-1)^{i+j}\mathrm{Cof}(a_{ij})$, and the **cofactor** $\mathrm{Cof}(a_{ij})$ denotes the **cofactor** of a_{ij} denotes the determinant of the $n-1 \times n-1$ matrix which is obtained by deleting row i and column j from the matrix A.

Hint 1: Recall the possibility to compute the determinant of a matrix by expanding it into subdeterminants,

$$\det(A) = a_{i1}A_{i1} + a_{i2}A_{i2} + \cdots + a_{in}A_{in} \tag{5}$$

for any row i (with a similar expansion for any column j)

Hint 2: Realize that to compute the inverse A^{-1} you have to solve n linear systems of the form $A\boldsymbol{x} = \boldsymbol{b}$ with a particular rhs vector $\boldsymbol{b}$ (one for each column of A^{-1})

Question.

Is Cramer's rule a feasible method to solve the linear system (2)? To answer this question, we have to take a closer look at the computational complexity when computing determinants.

Its time for some Blackboard writing....

1.2 The Gaussian Elimination Method

Solving linear systems involving triangular matrices.

Definition 1: Square matrices $R, L \in \mathbb{R}^{n,n}$ of the form

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & \cdots & u_{2n} \\ \vdots & & \ddots & & \vdots \\ 0 & & & \ddots & \vdots \\ 0 & & & & u_{nn} \end{pmatrix}, \quad L = \begin{pmatrix} l_{11} & 0 & & & 0 \\ l_{21} & l_{22} & & & \vdots \\ \vdots & \vdots & \ddots & & 0 \\ \vdots & \vdots & & \ddots & 0 \\ l_{n1} & l_{n2} & \cdots & \cdots & l_{nn} \end{pmatrix} \tag{6}$$

are called **upper triangular** and **lower triagonal**, respectively. If in addition $l_{ii} = 1$ for $i = 1, \ldots, n$, the L is called a **unit lower triangular** matrix.

Exercise 2: A solution algorithm for $U\boldsymbol{x} = \boldsymbol{b}$

The nice with upper or lower triangular matrices is that the corresponding matrix equation is particularly easy to solve. For instance, looking at

$$U\boldsymbol{x} = \boldsymbol{b}$$

you can write up a simple algorithm to successively solve for $x_n, x_{n-1}, \ldots x_1$, starting from x_n.

Solution. Successively compute

$$\begin{aligned}
x_n &= b_1/u_{nn}, \quad \text{if } u_{nn} \neq 0,\\
x_{n-1} &= \left(b_{n-1} - u_{n-1,n}x_n\right)/u_{n-1,n-1}, \quad \text{if } u_{n-1,n-1} \neq 0,\\
&\vdots\\
x_1 &= \left(b_1 - \sum_{i=2}^{n-1} u_{1,i}x_i\right)/u_{11}, \quad \text{if } u_{11} \neq 0.
\end{aligned}$$

Incidentally, we notice that the system can be solved $\Leftrightarrow u_{11}, \dots u_{nn} \neq 0 \Leftrightarrow \det(U) = u_{11} \cdots u_{nn} \neq 0 \Leftrightarrow U$ is invertible.

The Gaussian Elimination Method (GUM) is an algorithm which transforms a given matrix equation $A\boldsymbol{x} = \boldsymbol{b}$ to an equivalent system (that is, having the same solution, if any) of the form $L\boldsymbol{x} = \widetilde{\boldsymbol{b}}$. Lets start by relabeling A and $\boldsymbol{b}$ to $A^{(1)}$ and $\boldsymbol{b}^{(1)}$, respectively.

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{n1}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}^{(1)} & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(1)} \\ \vdots \\ b_n^{(1)} \end{pmatrix} \tag{7}$$

Assuming the $a_{11}^{(1)}$ is not 0, unknown x_1 can be eliminated from equation row 2 to n:

$$l_{i1} := \frac{a_{i1}^{(1)}}{a_{11}^{(1)}} \qquad \text{for } i = 2, 3, \dots n, \tag{8}$$

$$a_{ij}^{(2)} := a_{ij}^{(1)} - l_{i1}a_{1j}^{(1)} \qquad \text{for } i, j = 2, \dots n, \tag{9}$$

$$b_i^{(2)} := b_i^{(1)} - l_{i1}b_1^{(1)} \qquad \text{for } i = 2, \dots n. \tag{10}$$

This will lead to an equivalent linear system of the form $A^{(2)}\boldsymbol{x} = \boldsymbol{b}^{(2)}$ of the form

$$\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{n1}^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1^{(2)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(2)} \end{pmatrix} \tag{11}$$

We can repeat this process to eliminate the unknown x_2 from row 3 to n. Repeating this process will lead to a sequence of equivalent linear systems

$$(A^{(1)}, \boldsymbol{b}^{(1)}) \mapsto (A^{(2)}, \boldsymbol{b}^{(2)}) \mapsto \cdots \mapsto (A^{(n)}, \boldsymbol{b}^{(n)}) = (U, \boldsymbol{b}^{(n)})$$

So for $k = 1, \dots, n-1$, we pass from $(A^{(k)}, \boldsymbol{b}^{(k)})$ to $(A^{(k+1)}, \boldsymbol{b}^{(k+1)})$ by calculating

$$l_{ik} := \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} \quad \text{for } i = k+1, \dots n, \tag{12}$$

$$a_{ij}^{(k+1)} := a_{ij}^{(k)} - l_{ik} a_{kj}^{(k)} \quad \text{for } i, j = k+1, \dots n, \tag{13}$$

$$b_i^{(k+1)} := b_i^{(k)} - l_{ik} b_k^{(k)} \quad \text{for } i = k+1, \dots n. \tag{14}$$

Question.

What did I miss in the definition of these reduction steps?

Yes, $a_{kk}^{(k)}$ needs to be non-zero!

1.3 LU Factorization

Next we want to reinterpret what we just have done. We define the matrix E_{ik} to be the matrix which only non-zero element is $e_{ik} = 1$. Also, we set $L_{ik} = Id - l_{ik} E_{ik}$.

Question: What is the effect of the matrix multiplication $L_{ik} A$?

A. Rescales column i with l_{ik} and subtracts it from column k.
B. Rescales column k with l_{ik} and subtracts it from column i.
C. Rescales row k with l_{ik} and subtracts it from row i.
D. Rescales row i with l_{ik} and subtracts it from row k.
Solution: **A**: Wrong. **B**: Wrong. **C**: Right. **D**: Wrong.

Exercise 3: Elimination as Matrix Multiplication

Show that you can perform the transformation step

$$(A^{(k)}, \boldsymbol{b}^{(k)}) \mapsto (A^{(k+1)}, \boldsymbol{b}^{(k+1)})$$

via a simple matrix multiplication

$$L_k A^{(k)} = A^{(k+1)}, \quad L_k b^{(k)} = b^{(k+1)}$$

where L_k is unit lower triangular and can be written in terms of L_{ik}. How does the final matrix L_k look like?

Solution. Starting from

$$A^{(k)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & \cdots & a_{n1}^{(1)} \\ 0 & a_{22}^{(2)} & & & & a_{2n}^{(2)} \\ \vdots & & \ddots & & & \vdots \\ 0 & \dots & 0 & a_{kk}^{(k)} & \dots & a_{kn}^{(k)} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{nk}^{(k)} & \dots & a_{nn}^{(k)} \end{pmatrix} \tag{15}$$

Now we apply $L_{k+1,k}$ to obtain

$$L_{k+1,k}A^{(k)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & \cdots & \cdots & a_{n1}^{(1)} \\ 0 & a_{22}^{(2)} & & & & & a_{2n}^{(2)} \\ \vdots & & \ddots & & & & \vdots \\ 0 & \dots & 0 & a_{kk}^{(k)} & \dots & \dots & a_{kn}^{(k)} \\ 0 & \dots & 0 & 0 & a_{k+1,k+1}^{(k+1)} & \dots & a_{k+1,n}^{(k+1)} \\ 0 & \dots & 0 & a_{k+2,k}^{(k)} & \dots & \dots & a_{k+2,n}^{(k)} \\ \vdots & & \vdots & \vdots & & & \vdots \\ 0 & \dots & 0 & a_{nk}^{(k)} & \dots & \dots & a_{nn}^{(k)} \end{pmatrix} \tag{16}$$

Continuing like this and multiplying from the left with $L_{k+2}, \dots, L_{nk}$ we obtain

$$\underbrace{L_{nk}L_{n-1,k}\cdots L_{k+1,k}}_{:=L_k}A^{(k)} = \underbrace{\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & \cdots & \cdots & a_{n1}^{(1)} \\ 0 & a_{22}^{(2)} & & & & a_{2n}^{(2)} \\ \vdots & & \ddots & & & \vdots \\ 0 & \dots & 0 & a_{k+1k+1}^{(k+1)} & \dots & a_{k+1n}^{(k+1)} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a_{nk+1}^{(k+1)} & \dots & a_{nn}^{(k+1)} \end{pmatrix}}_{A^{(k+1)}} \tag{17}$$

The final matrix L_k goes often under the name **Frobenius matrix** and can be computed to

$$L_k = \begin{pmatrix} 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & & 1 & 0 & & 0 \\ 0 & \dots & -l_{k+1,k} & 1 & \dots & 0 \\ & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & -l_{n,k} & 0 & \dots & 1 \end{pmatrix} \tag{18}$$

To summarize, we see that

$$L_{n-1} \cdots L_2 L_1 A = A^{(n)} = U \tag{19}$$

Theorem 1.

1. The product of two (unit) lower triangular matrices of order n is againa a (unit) lower triangular matrix of order n;
2. A lower triangular matrix is invertible if and only if all diagonal entries are not zero;
3. a lower triangular matrix is invertible if and only if all diagonal entries are not zero;
4. the inverse of a nonsingular lower triangular matrix of order n is lower triangular of order n

Problem 4: Prove the last statements

The first 3 statements should not be to hard to prove, the challenging one is the forth. Guidance will be given as part of the first computer lab.

Using these properties of (unit) lower triangular matrices of order n, we can now conclude that rewriting the GUM in in terms of successively applying unit lower triangular matrices of the form L_k yields us to the **LU factorization** of A:

$$A = (L_{k-1} \cdots L_2 L_1)^{-1} U = \underbrace{L_1^{-1} L_2^{-1} \cdot L_{k-1}^{-1}}_{L} U \tag{20}$$

Question.

We have already seen, that the GUM can break down if at some point the diagonal element $a_{kk}^k = 0$ for any $k \in 1, \ldots, n-1$. So under what conditions does a square matrix possess a LU factorization?

Definition 2: For a square matrix $A \in \mathbb{R}^{n,n}$, the **leading principal submatrix** of order k with $1 \leqslant k \leqslant n$ is defined as the matrix $A^{(k)} \in \mathbb{R}^{k,k}$ whose elements $a_{ij}^{(k)}$ satisfies $a_{ij}^{(k)} = a_{ij}$ for $1 \leqslant i, j \leqslant k$. (Roughly speaking we "erase" any rows and columns from A with row or column index larger than k)

Theorem 2: Let $A \in \mathbb{R}^{n,n}$ with $n \geqslant 2$ and assume that every leading principal submatrix of order k with $1 \leqslant k \leqslant n-1$ is invertible. Then A admits a LU factorization, where L is unit lower triangular of order n, and R is upper triangular of order n.

Proof. The following steps will guide you through the proof and is built on a classical proof by induction.

1. (**Base case**) Prove the theorem for $n = 2$. We could have started with $n = 1$ as one as well, but the case $n = 2$ already illustrates the main idea, when going from n to $n+1$. Assume that A fullfills the assumption of the theorem and write it as

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \tag{21}$$

What can you say about a? Now we wish to establish the existence of

$$L = \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix}, \quad U = \begin{pmatrix} u & v \\ 0 & \eta \end{pmatrix}, \tag{22}$$

such that $LU = A$. Multiply out LU and equal every element in the product matrix to the corresponding elements of A. You get then 4 equations which you have to consider.

2. (**Inductive step**) Assume the theorem is proven for all matrices of order k, $2 \leqslant k \leqslant n$, and suppose that $A \in \mathbb{R}^{n+1,n+1}$, satisfying the assumptions of the theorem.

a) Inspired by the case $n = 2$, rewrite A as a block matrix

$$A = \begin{pmatrix} A^{(n)} & \boldsymbol{b} \\ \boldsymbol{c}^T & d \end{pmatrix}, \tag{23}$$

where $A^{(n)} \in \mathbb{R}^{n,n}$ is a not-singular matrix. Here $\boldsymbol{b}$ and $\boldsymbol{c}$ are column vectors, and η is a scalar again.

b) Assuming that A satisfy the assumptions of the theorem and that the theorem is already proved for n. Show that $A^{(n)}$ admits a LU factorization and make the ansatz

$$L = \begin{pmatrix} L^{(n)} & \boldsymbol{0} \\ \boldsymbol{m}^T & 1 \end{pmatrix}, \quad U = \begin{pmatrix} U^{(n)} & \boldsymbol{v} \\ 0 & \eta \end{pmatrix}, \tag{24}$$

Now proceed similar as in step and compute the (block) matrix product of LU. then the resuliting blocks entries need to equal the corresponding ones in block representation of A. Again carefully consider the resulting 4 "block equations". Why can you solve it? At some point you might also need to recall the Binet-Cauchy theorem stating that $\det(XY) = \det(X)\det(Y)$ for square matrices X and Y.

Congratulations!! You have successfully proved that for a particular type of invertible matrices, you can always find a LU factorization. What about general invertible matrices which do not satisfy the specific assumptions made in Theorem 2? That will be the content of the next lecture.