# Numerical Mathematics Lab session 3

Stefan Evanghelides (s2895323) and Frank te Nijenhuis (S2462575)

May 15, 2018

## 1 Preparation

### 1.1 LU factorisation and partial pivoting

2. Having $Ax = b$, where

$$A = \begin{pmatrix} \epsilon & 1 \\ 1 & 1 \end{pmatrix}$$

- For $\epsilon \neq 1$ and $b = (1, 2)^T$,
  (a) $det(a) = \epsilon - 1 \neq 0$ for $\epsilon \neq 1$, so A is nonsingular (invertible).
  (b) For $Ax = 0$, where $x = (x_1, x_2)^T$, we obtain the following system:

$$\begin{cases} \epsilon x_1 + x_2 = 0 \\ x_1 + x_2 = 0 \end{cases} \overset{R_1 - R_-2 \to R_1}{=} \begin{cases} (\epsilon - 1)x_1 \quad = 0 \\ x_1 + x_2 = 0 \end{cases} \to \begin{cases} x_1 = 0 \\ x_2 = 0 \end{cases}$$

  So, the null space of A has only one element, the vector $x = (0, 0)^T$
  The bases of the range of A has 2 vectors $v_1 = (\epsilon, 1)^T$ and $v_2 = (1, 1)^T$
  Since the above 2 vectors form the basis of the range space, it also holds that any linear combination of them are part of the range. Having $\epsilon = 0.5$ will result in the fact that $b = 2 \cdot v1$, so $b = (1, 2)^T$ is in the range of A only for $\epsilon = 0.5$.
  (c) Solving the following system will yield:

$$\begin{cases} \epsilon x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases} \overset{R_1 - R_-2 \to R_1}{=} \begin{cases} (\epsilon - 1)x_1 \quad = -1 \\ x_1 + x_2 = 2 \end{cases} \to \begin{cases} x_1 = \frac{1}{1-\epsilon} \\ x_2 = 2 - x_1 = \frac{1-2\epsilon}{1-\epsilon} \end{cases}$$

  So, there is only 1 solution, $x = (\frac{1}{1-\epsilon}, \frac{1-2\epsilon}{1-\epsilon})^T$.
  (d) There is no general form for the solution, as there is a unique solution presented above.

- For $\epsilon = 1$ and $b = (1, 2)^T$,
  (a) $det(a) = \epsilon - 1 = 0$ for $\epsilon = 1$, so A is singular.
  (b) For $Ax = 0$, where $x = (x_1, x_2)^T$ and $\epsilon = 1$

$$\begin{cases} x_1 + x_2 = 0 \\ x_1 + x_2 = 0 \end{cases} \to \begin{cases} x_1 = -x_2 \end{cases}$$

  So, the null space of A has an infinite amount of elements, the vectors of the form $x = (x_1, -x_1)^T$.
  The basis of the range of A has 1 vector $v = (1, -1)^T$ and $b \notin Range(A)$.
  (c) The system has an infinite amount of solutions
  (d) Having $x_1 = \alpha$, then the general form of the solution is $x = (\alpha, -\alpha)^T$

- For $\epsilon = 1$ and $b = (1,1)^T$,
  (a) $det(a) = \epsilon - 1 = 0$ for $\epsilon = 1$, so A is singular.
  (b) Same answer as in the previous case.
  (c) Solving the following system for $\epsilon = 1$ will yield:

$$\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_2 = 1 \end{cases} \quad \to \begin{cases} x_1 = 1 - x_2 \end{cases}$$

  So, there are an infinite amount of solutions.
  (d) Having $x_1 = \alpha$, the general form of the solutions is $x = (\alpha,\ 1-\alpha)^T$

3. Show that $K_*(A) = ||A||_*||A^{-1}||_*$ satisfies $K_*(AB) \leq K_*(A)K_*(B)$,
for all nonsingular matrices $A, B \in R^{nxn}$.

$$K_*(AB) = ||AB||_*||(AB)^{-1}||_* = ||AB||_*||B^{-1}A^{-1}||_* \ \leq \ ||A||_*||B||_*||B^{-1}||_*||A^{-1}||_*$$

Since the norms are real numbers, they commute, so we can rearrange the right-hand side of the inequality in the following way:

$$\begin{aligned} K_*(AB) \ &\leq \ ||A||_* \ ||A^{-1}||_* \ ||B||_* \ ||B^{-1}||_* \iff \\ K_*(AB) \ &\leq \ K_*(A) \ K_*(B) \end{aligned}$$

4. An upper bound for the relative error after perturbing the whole system is given by the inequality

$$\frac{||\mathbf{a} - \hat{\mathbf{a}}||}{||\mathbf{a}||} \leq \frac{K_2(M)}{1 - K_2(A)||\delta M||_2/||M||_2} \left( \frac{||\delta M||_2}{||M||_2} + \frac{||\delta \mathbf{f}||}{||\mathbf{f}||} \right)$$

If we only perturb $\mathbf{f}$ we are left with the much simpler bound

$$\frac{||\mathbf{a} - \hat{\mathbf{a}}||}{||\mathbf{a}||} \leq K(M) \frac{||\delta \mathbf{f}||}{||\mathbf{f}||}$$

5. We know that $r = f - M\hat{a} = Ma - M\hat{a} = M(a - \hat{a})$
We also know that $\delta b = M(\hat{a} - a) = -r$
So, the scaled residual norm can be expressed as follows: $\frac{||\delta f||_2}{||f||_2} = \frac{||r||_2}{||f||_2}$
Therefore, the inequality presented above can rewritten as follows:

$$\frac{||\mathbf{a} - \hat{\mathbf{a}}||_2}{||\mathbf{a}||_2} \leq K_2(M) \frac{||\mathbf{r}||_2}{||\mathbf{f}||_2}$$

If $K_2(A)$ is relatively small, then the residual is small so the above inequality is a good measure of the relative error. Otherwise, not.

## 1.2 Conditioning of the least squares problem

2. For $A \in \mathbb{R}^{m \times n}$, $\bar{c} \in \mathbb{R}^n$ and $\bar{y} \in \mathbb{R}^m$ with $m \geq n$ we want to minimize

$$||A\mathbf{c} - \mathbf{y}||_2^2$$

This means we need to find minimal

$$\Phi(a_0, a_1, \ldots, a_m) = \sum_{i=0}^{n} [(a_0 + a_1 c_i + \ldots + a_m c_i^m) - y_i]^2$$

Setting the partial derivatives to 0,

$$\frac{\partial \Phi}{\partial a_i} = 0$$

we obtain a system of equations

$$\sum_{i=0}^{n}[(a_0 + a_1 c_i + \ldots + a_m c_i^m) - y_i] = 0$$

which can be written in matrix form as

$$A\mathbf{c} - \mathbf{y} = 0$$

or

$$A\mathbf{c} = \mathbf{y}$$

or, by multiplying both sides by $A^t$, such that $A^t A$ is a square invertible matrix, as the normal equations

$$A^t A\mathbf{c} = A^t\mathbf{y}$$

3. A unique solution to the normal equations exists if $A$ has full rank, $rank(A) = \min(m, n) = m$.

4. We have $A^t A\bar{x}^* = A^t\bar{b}$ with $A = QR$ and, subsequently $A^t = (QR)^t = R^t Q^t$.

$$A^t A\bar{x}^* = A^t\bar{b}$$

Substituting the definitions

$$R^t Q^t Q R\bar{x}^* = R^t Q^t\bar{b}$$

Using the orthogonality of $Q$

$$R^t R\bar{x}^* = R^t Q^t\bar{b}$$

Multiplying by $(R^t)^{-1}$

$$R\bar{x}^* = Q^t\bar{b}$$

Finally, multiplying by $R^{-1}$

$$\bar{x}^* = R^{-1} Q^t\bar{b}$$

## 2 Lab Experiments

### 2.1 LU factorisation and partial pivoting

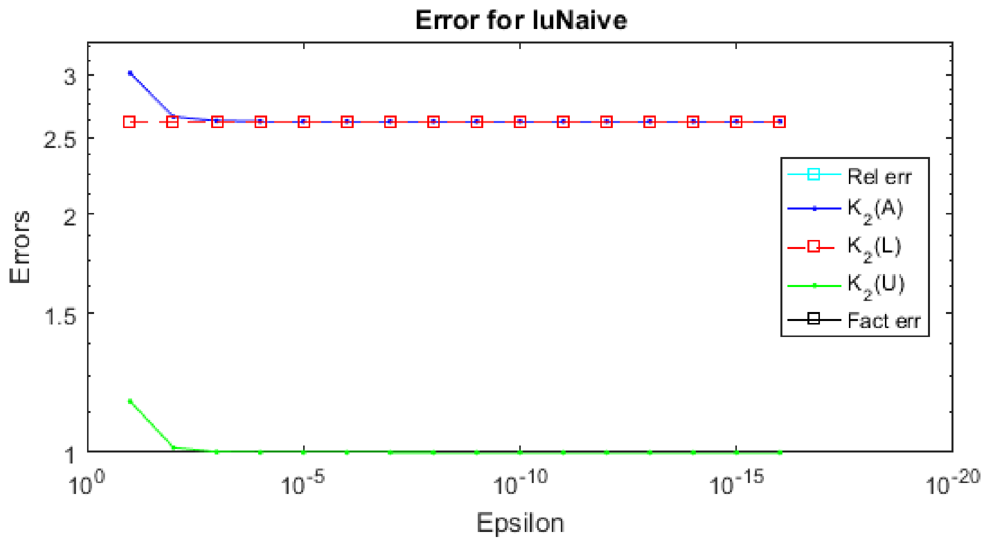



Figure 1: Quantities for naive LU factorisation

The relative error and factorisation errors from figure 1 above are 0 or nearly 0 so they don't appear in the plot.

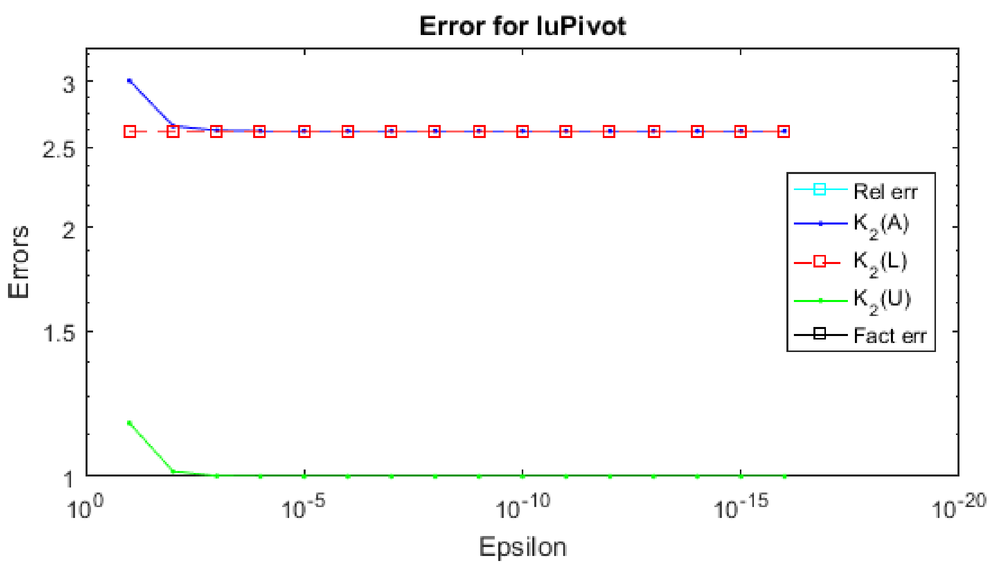
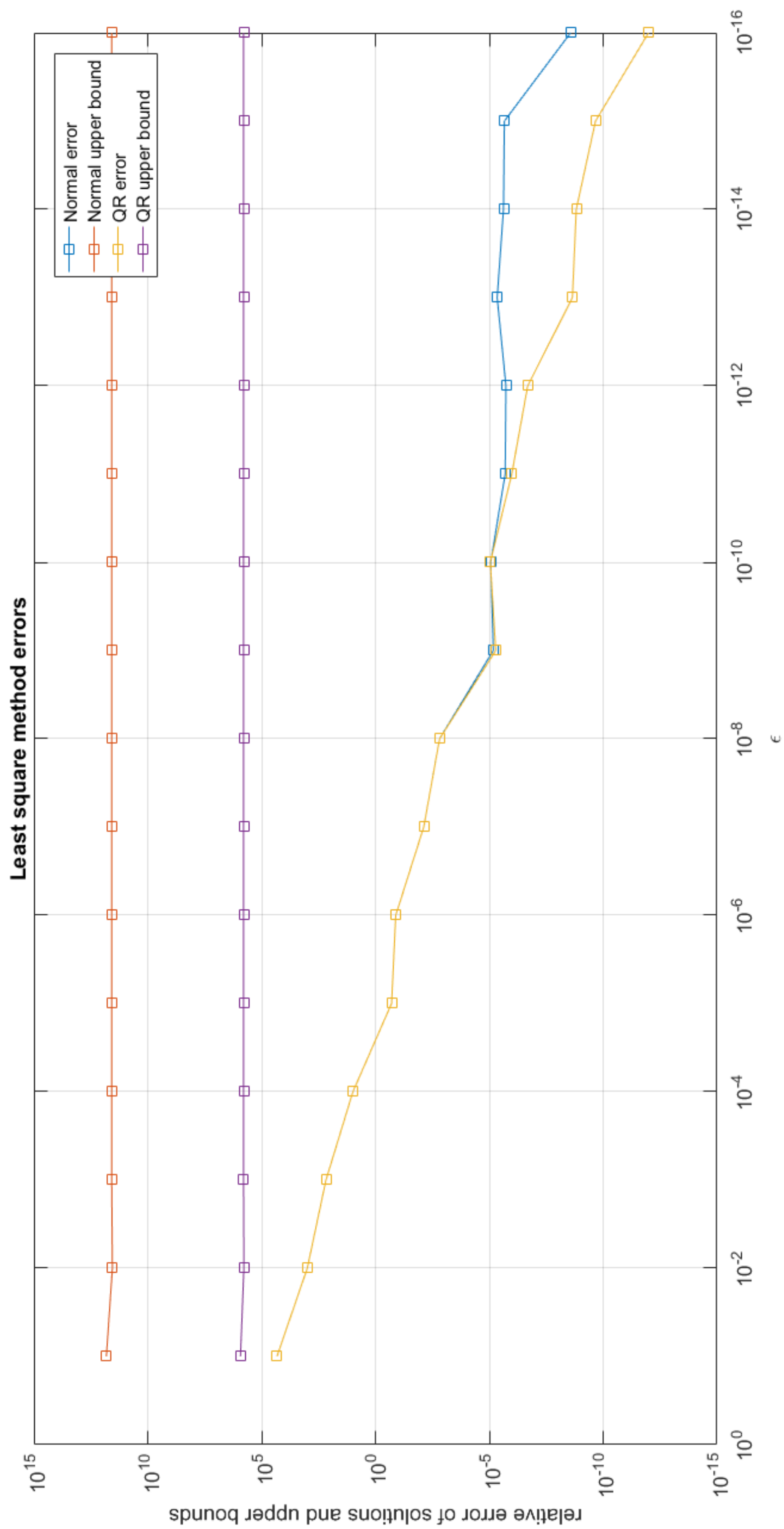



Figure 2: Quantities for partial pivoting LU factorisation

The relative error and factorisation errors from figure 2 above are 0 or nearly 0 so they don't appear in the plot.

### 2.2 Conditioning of the least squares problem


Least square method errors
Normal error
Normal upper bound
QR error
QR upper bound
relative error of solutions and upper bounds
ϵ
10^15
10^10
10^5
10^0
10^-5
10^-10
10^-15
10^0
10^-2
10^-4
10^-6
10^-8
10^-10
10^-12
10^-14
10^-16

# 3 Discussion

## 3.1 LU factorisation and partial pivoting

1. The issue with the normal factorisation is that it is dependent on the main diagonal's values to be different than 0. Partial pivoting is used to switch rows such that the main diagonal will not contain any zeros.

2. The overall LU factorization takes $\frac{2}{3}n^3$ operations, where $n$ is the dimension of the square matrix. If the matrix is banded with $b$ bands, then the number of operations are $\frac{2}{3}b^2n$. This can be justified by the fact that instead of going through $n$ elements for both rows and columns, we go through only $b$ elements.

3.

(a) $Ax = b \to LUx = b$. We can thus note $Ux = y$ so the remaining system to solve is $Ly = b$.

(b) $K(A)$ is almost as large as $K(L)$ while $K(U)$ is close to 1.

(c) $K(A) = K(LU) \leq K(L) \cdot K(U)$, which means that the condition number of $A$ can never be larger than the product of $L$ and $U$.

(d) Yes, looking at the graphs we see that $K(L)$ is almost equal to $K(A)$, while $K(U)$ is slightly greater than 1.

4.

(a) Let A = $\begin{bmatrix} \epsilon & 1 \\ 0 & 1 - \frac{1}{\epsilon} \end{bmatrix}$

and b = $\begin{bmatrix} 1 \\ 2 - \frac{1}{\epsilon} \end{bmatrix}$

For $Ax = b \to x = A^{-1}b$

Where $A^{-1} = \begin{bmatrix} \frac{1}{\epsilon} & \frac{1}{1-\epsilon} \\ 0 & \frac{\epsilon}{\epsilon-1} \end{bmatrix}$ such that $x = \begin{bmatrix} \frac{1}{1-\epsilon} \\ \frac{2\epsilon}{\epsilon-1} \end{bmatrix}$ The values of $\epsilon$ do not cause any problems in the denominators as $\epsilon \to 0$ so the problem is well-defined.

(b) Let $\frac{||\delta \mathbf{b}||}{||\mathbf{x}||} = \epsilon_m$, where we perturb the system by the smallest possible amount on a machine. Then using the fact that $A$ is positive semidefinite, we can use the bound

$$\frac{||\mathbf{x} - \hat{\mathbf{x}}||}{||\mathbf{x}||} \leq \frac{1}{\lambda_{min}} \frac{\delta \mathbf{x}}{\mathbf{x}} = \frac{\epsilon_m}{\epsilon}$$

where we use the fact that $\lambda_{min} = \epsilon$

(c) This equation implies that the error should grow once $\epsilon$ is smaller than the machine precision. In our experiments, we do not reach this point, so our result are in agreement with the theory as much as possible.

## 3.2 Conditioning of the least squares problem

1. We obtain the $K_2$ numbers of the matrices using the MATLAB commands `cond(A.'*A,2)` and `cond(R,2)`, respectively. For the first matrix, we obtain a value of $3.8117 \cdot 10^{11}$, for the second one we get $6.1739 \cdot 10^5$. We can conclude that the QR factorization is better conditioned, since it has a lower condition number. This agrees with our results, which indicate a lower error when using the QR factorization as opposed to the normal equations.
2. As seen on the plots the theoretical bounds are much higher than the actual errors found by the system. Therefore the bounds found in question 4 are in agreement with the practical results.