CS724: Topics in Algorithms
Solving Linear Systems
Slide Set 6

Prof. Dan A. Simovici
1. Linear Systems and Matrices

2. The Row Echelon Form of Matrices

3. Solving Linear Systems in MATLAB

4. Condition Numbers for Matrices
Consider the following set of linear equalities

\[
\begin{align*}
    a_{11}x_1 + \ldots + a_{1n}x_n &= b_1, \\
    a_{21}x_1 + \ldots + a_{2n}x_n &= b_2, \\
    \vdots \quad &\vdots \\
    a_{m1}x_1 + \ldots + a_{mn}x_n &= b_m,
\end{align*}
\]

where \( a_{ij} \) and \( b_i \) belong to a field \( F \). This set constitutes a system of linear equations. Solving this system means finding \( x_1, \ldots, x_n \) that satisfy all equalities.
The system can be written succinctly in a matrix form as $Ax = b$, where

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix},$$

and

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$  

If the set of solutions of a system $Ax = b$ is not empty we say that the system is *consistent*. Note that $Ax = b$ is consistent if and only if $b \in \text{range}(A)$. 
Let $Ax = b$ be a linear system in matrix form, where $A \in \mathbb{C}^{m \times n}$. The matrix $[A \ b] \in \mathbb{C}^{m \times (n+1)}$ is the augmented matrix of the system $Ax = b$. 
Theorem

Let $A \in \mathbb{C}^{m \times n}$ be a matrix and let $b \in \mathbb{C}^{n \times 1}$. The linear system $Ax = b$ is consistent if and only if $\text{rank}(Ab) = \text{rank}(A)$.

Proof.

If $Ax = b$ is consistent and $x' = (x_1, \ldots, x_n)$ is a solution of this system, then $b = x_1c_1 + \cdots + x_nc_n$, where $c_1, \ldots, c_n$ are the columns of $A$. This implies $\text{rank}([A \ b]) = \text{rank}(A)$.

Conversely, if $\text{rank}(Ab) = \text{rank}(A)$, the vector $b$ is a linear combination of the columns of $A$, which means that $Ax = b$ is a consistent system.
**Definition**

An **homogeneous linear system** is a linear system of the form \( Ax = 0_m \), where \( A \in \mathbb{C}^{m \times n} \), \( x \in \mathbb{C}^{n,1} \) and \( 0 \in \mathbb{C}^{m \times 1} \).

Clearly, any homogeneous system \( Ax = 0_m \) has the solution \( x = 0_n \). This solution is referred to as the **trivial solution**. The set of solutions of such a system is null(\( A \)), the null space of the matrix \( A \).
Let \( u \) and \( v \) be two solutions of the system \( Ax = b \). Then \( A(u - v) = 0_m \), so \( z = u - v \) is a solution of the homogeneous system \(Ax = 0_m\), or \( z \in \text{null}(A)\). Thus, the set of solutions of \( Ax = b \) can be obtained as a “translation” of the null space of \( A \) by any particular solution of \( Ax = b \). In other words the set of solution of \( Ax = b \) is \( \{x + z \mid z \in \text{null}(A)\} \).

Thus, for \( A \in \mathbb{C}^{m \times n} \), the system \( Ax = b \) has a unique solution if and only if \( \text{null}(A) = \{0_n\} \), that is, if \( \text{rank}(A) = n \).
Theorem

Let $A \in \mathbb{C}^{n \times n}$. Then, $A$ is invertible (which is to say that $\text{rank}(A) = n$) if and only if the system $Ax = b$ has a unique solution for every $b \in \mathbb{C}^{n}$.

Proof.

If $A$ is invertible, then $x = A^{-1}b$, so the system $Ax = b$ has a unique solution.

Conversely, if the system $Ax = b$ has a unique solution for every $b \in \mathbb{C}^{n}$, let $c_1, \ldots, c_n$ be the solution of the systems $Ax = e_1, \ldots, Ax = e_n$, respectively. Then, we have

$$A(c_1| \cdots |c_n) = I_n,$$

which shows that $A$ is invertible and $A^{-1} = (c_1| \cdots |c_n)$. 

\[\Box\]
Corollary

An homogeneous linear system $Ax = 0$, where $A \in \mathbb{C}^{n \times n}$ has a non-trivial solution if and only if $A$ is a singular matrix.
**Definition**

A matrix $A \in \mathbb{C}^{n \times n}$ is **diagonally dominant** if

$$|a_{ii}| > \sum\{|a_{ik}| \mid 1 \leq k \leq n \text{ and } k \neq i\}.$$ 

**Theorem**

A diagonally dominant matrix is non-singular.
Proof

Suppose that $A \in \mathbb{C}^{n \times n}$ is a diagonally dominant matrix that is singular. The homogeneous system $Ax = 0$ has a non-trivial solution $x \neq 0$. Let $x_k$ be a component of $x$ that has the largest absolute value. Since $x \neq 0$, we have $|x_k| > 0$. We can write

$$a_{kk}x_k = - \sum \{ a_{kj}x_j \mid 1 \leq j \leq n \text{ and } j \neq k \},$$

which implies

$$|a_{kk}| |x_k| = \left| \sum \{ a_{kj}x_j \mid 1 \leq j \leq n \text{ and } j \neq k \} \right| \leq \sum \{ |a_{kj}| |x_j| \mid 1 \leq j \leq n \text{ and } j \neq k \} \leq |x_k| \sum \{ |a_{kj}| \mid 1 \leq j \leq n \text{ and } j \neq k \}.$$

Thus, we obtain

$$|a_{kk}| \leq \sum \{ |a_{kj}| \mid 1 \leq j \leq n \text{ and } j \neq k \},$$

which contradicts the fact that $A$ is diagonally dominant.
We begin with a class of linear systems that can easily be solved.

**Definition**

A matrix $C \in \mathbb{C}^{m \times n}$ is in row echelon form if the following conditions are satisfied:

- rows that contain a non-zero elements precede zero rows (that is, rows that contain only zeros);
- if $c_{ij}$ is the first non-zero element of the row $i$, all elements in the $j^{th}$ column located below $c_{ij}$, that is, entries of the form $c_{kj}$ with $k > j$ are zero;
- if $i < \ell$, $c_{ij_i}$ is the first non-zero element of the row $i$, and $c_{\ell j_\ell}$ is the first non-zero element of the row $\ell$, then $j_i < j_\ell$.

The first non-zero element of a row $i$ (if it exists) is called the **pivot of the row $i$**.
Example

Let $C \in \mathbb{R}^{4 \times 5}$ be the matrix

$$C = \begin{pmatrix}
1 & 2 & 0 & 2 & 0 \\
0 & 2 & 3 & 0 & 1 \\
0 & 0 & 0 & -1 & 2 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

It is clear that $C$ is in row echelon form; the pivots of the first, second and third rows are $c_{11} = 1$, $c_{22} = 2$, and $c_{34} = -1$. 
**Theorem**

Let $C \in \mathbb{C}^{m \times n}$ be a matrix in row echelon form such that the rows that contain non-zero elements are the first $r$ rows. Then, $\text{rank}(C) = r$.

**Proof:** Let $c_1, \ldots, c_r$ be the non-zero rows of $C$. Suppose that the row $c_i$ has the first non-zero element in the column $j_i$ for $1 \leq i \leq r$. By the definition of the echelon form we have $j_1 < j_2 < \cdots < j_r$. 
Suppose that \( a_1 c_1 + \cdots + a_r c_r = 0 \). This equality can be written as:

\[
\begin{align*}
    a_1 c_{1j_1} &= 0, \\
    a_1 c_{1j_2} + a_2 c_{2j_2} &= 0, \\
    &\vdots \\
    a_1 c_{1n} + a_2 c_{2n} + \cdots + a_r c_{rn} &= 0.
\end{align*}
\]

Since \( c_{1j_1} \neq 0 \), we have \( a_1 = 0 \). Substituting \( a_1 \) by 0 in the second equality implies \( a_2 = 0 \) because \( c_{2j_2} \neq 0 \), etc. Thus, we obtain \( a_1 = a_2 = \ldots = a_r = 0 \), which proves that the rows \( c_1, \ldots, c_r \) are linearly independent. Since this is a maximal set of rows of \( C \) that is linearly independent, it follows that \( \text{rank}(C) = r \).
Linear systems whose augmented matrices are in row echelon form can be easily solved using a process called *back substitution*. Consider the augmented matrix in row echelon form of a system with $m$ equations and $n$ unknowns:

$$
\begin{pmatrix}
0 & \cdots & 0 & a_{1j_1} & \cdots & \cdots & a_{1n} & b_1 \\
0 & \cdots & 0 & 0 & a_{2j_2} & \cdots & a_{2n} & b_2 \\
\vdots & \cdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & 0 & 0 & \cdots & a_{rj_r} & \cdots & b_r \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & b_{r+1} \\
\vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & b_m \\
\end{pmatrix}
$$

The system of equations has the form:

\[
\begin{align*}
a_{1j_1}x_{j_1} + \cdots + a_{1n}x_n &= b_1 \\
a_{2j_2}x_{j_2} + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{rj_r}x_{j_r} + \cdots + a_{rn}x_n &= b_r
\end{align*}
\]
The variables $x_{j_1}, x_{j_2}, \ldots, x_{j_r}$ that correspond to the columns where the pivot element occur are referred to as the *basic variables*. The remaining variables are *non-basic*.

Note that we have $r \leq \min\{m, n\}$. If $r < m$ and there exists $b_\ell \neq 0$ for $r < \ell \leq m$, then the system is inconsistent and no solutions exist. If $r = m$ or $b_\ell = 0$ for $r < \ell \leq m \leq n$, one can choose the variables that do not correspond to the pivot elements, $\{x_i \mid i \notin \{j_1, j_2, \ldots, j_r\}\}$ as parameters and express the basic variables as functions of these parameters.
The process starts with the last basic variable, $x_{jr}$ (because every other variable in the equation $a_{rj}x_{jr} + \cdots + a_{rn}x_n = b_r$ is a parameter), and then substitutes this variable in the previous equality. This allows us to express $x_{jr-1}$ as a function of parameters, etc. This explains the term \textit{back substitution} previously introduced. If $r = n$, then no parameters exist.

To conclude, if $r < m$ the system has a solution if and only if $b_j = 0$ for $j > r$. If $r = m$, the system has a solution. This solution is unique if $r = n$. 
Example

Consider the system

\[
\begin{align*}
\ x_1 & \ + \ 2x_2 & \ + & \ 0 & \ = \ b_1 \\
2x_2 & \ + \ 3x_3 & \ + & \ x_4 & \ = \ b_2 \\
1 & \ = \ b_3 & \ = \ b_4
\end{align*}
\]

The augmented matrix of this system is

\[
\begin{pmatrix}
1 & 2 & 0 & 2 & 0 & b_1 \\
0 & 2 & 3 & 0 & 1 & b_2 \\
0 & 0 & 0 & -1 & 2 & b_3 \\
0 & 0 & 0 & 0 & 0 & b_4
\end{pmatrix}
\]

The basic variables are \( x_1, x_2 \) and \( x_4 \). If \( b_4 = 0 \) the system is consistent. Under this assumption we can choose \( x_3 \) and \( x_5 \) as parameters. Let \( x_3 = p \) and \( x_5 = q \).
Example

The third equation yields $x_4 = q - b_3$. Similarly, the second equation implies $x_2 = 0.5(b_2 - 3p - q)$. Substituting these values in the first equation allows us to write $x_1 = b_1 - b_2 + 2b_3 - 3p - q$.

Further transformations of this system allow us to construct an equivalent linear system whose matrix contain the columns of the matrix $I_3$.

Subtracting the second row from the first yields:

$$
\begin{pmatrix}
1 & 0 & -3 & 2 & -1 & b_1 - b_2 \\
0 & 2 & 3 & 0 & 1 & b_2 \\
0 & 0 & 0 & -1 & 2 & b_3 \\
0 & 0 & 0 & 0 & 0 & b_4 \\
\end{pmatrix}
$$
Example

Then, dividing the second row by 2 will produce:

\[
\begin{pmatrix}
1 & 0 & -3 & 2 & -1 & b_1 - b_2 \\
0 & 1 & 3/2 & 0 & 1/2 & b_{2/2} \\
0 & 0 & 0 & -1 & 2 & b_3 \\
0 & 0 & 0 & 0 & 0 & b_4
\end{pmatrix}
\]

which creates the second column of \( I_3 \). Next, multiply the third row by \(-1\):

\[
\begin{pmatrix}
1 & 0 & -3 & 2 & -1 & b_1 - b_2 \\
0 & 1 & 3/2 & 0 & 1/2 & b_{2/2} \\
0 & 0 & 0 & 1 & -2 & -b_3 \\
0 & 0 & 0 & 0 & 0 & b_4
\end{pmatrix}
\]

Finally, multiply the third by \(-2\) and add it to the first row:

\[
\begin{pmatrix}
1 & 0 & -3 & 0 & 1 & b_1 - b_2 + 2b_3 \\
0 & 1 & 3/2 & 0 & 1/2 & b_{2/2} \\
0 & 0 & 0 & 1 & -2 & -b_3 \\
0 & 0 & 0 & 0 & 0 & b_4
\end{pmatrix}
\]
Example

Thus, if we choose for the basic variables

\[ x_1 = b_1 - b_2 + 2b_3, \quad x_2 = \frac{b_2}{2}, \quad \text{and} \quad x_4 = -b_3 \]

and for the non-basic variables \( x_3 = x_5 = 0 \) we obtain a solution of the system.
We shall see that the extended echelon form of a system can be achieved by applying certain transformations on the rows of the augmented matrix of the system (which amount to transformations involving the equations of the system). In preparation, a few special invertible matrices are introduced in the next examples.
Example

Consider the matrix

\[ T(i \leftrightarrow j) = \begin{pmatrix}
1 & \cdots & \cdots & \cdots \\
\vdots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & \cdots & 1 \\
\vdots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & \ddots \\
1 & 0 & \cdots & \cdots \\
0 & \cdots & \cdots & 1 \\
\end{pmatrix}, \]

where line \( i \) contains exactly one 1 in position \( j \) and line \( j \) contains exactly one 1 in position \( i \). If \( T(i \leftrightarrow j) \in \mathbb{C}^{p \times p} \) and \( A \in \mathbb{C}^{p \times q} \), it is easy to see that the matrix \( T(i \leftrightarrow j)A \) is obtained from the matrix \( A \) by permuting the lines \( i \) and \( j \).
Example

For instance, consider the matrix $T^{(2)\leftrightarrow(4)} \in \mathbb{C}^{4 \times 4}$ defined by:

$$T^{(2)\leftrightarrow(4)} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix}$$

and the matrix $A \in K^{4 \times 5}$. We have:

$$T^{(2)\leftrightarrow(4)} A = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0
\end{pmatrix} \begin{pmatrix}
a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\
a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\
a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\
a_{41} & a_{42} & a_{43} & a_{44} & a_{45}
\end{pmatrix} = \begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
a_{41} & a_{42} & a_{43} \\
a_{31} & a_{32} & a_{33} \\
a_{21} & a_{22} & a_{23}
\end{pmatrix}$$

The inverse of $T^{(i)\leftrightarrow(j)}$ is $T^{(i)\leftrightarrow(j)}$ itself.
Example

Let \( T^{a(i)} \in \mathbb{C}^{p \times p} \) be the matrix

\[
T^{a(i)} = \begin{pmatrix}
1 & 0 & \ldots & 0 & \ldots & 0 \\
0 & 1 & \ldots & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \ddots & 1 & \ldots & 0 \\
0 & 0 & \ldots & 0 & \ldots & 1
\end{pmatrix}
\]

that has \( a \in F - \{0\} \) on the \( i^{th} \) diagonal element, 1 on the remaining diagonal elements and 0 everywhere else. The product \( T^{a(i)}A \) is obtained from \( A \) by multiplying the \( i^{th} \) row by \( a \). The inverse of this matrix is \( T^{\frac{1}{a(i)}} \).
Example

Consider the matrix $T^{3(2)} \in \mathbb{C}^{4 \times 4}$ given by

$$
T^{3(2)} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 3 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}.
$$

If $A \in \mathbb{C}^{4 \times 5}$ the matrix $T^{3(2)}A$ is obtained from $A$ by multiplying its second line by 3. We have

$$
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 3 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\
a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\
a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\
a_{41} & a_{42} & a_{43} & a_{44} & a_{45}
\end{pmatrix}
= 
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\
3a_{21} & 3a_{22} & 3a_{23} & 3a_{24} & 3a_{25} \\
a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\
a_{41} & a_{42} & a_{43} & a_{44} & a_{45}
\end{pmatrix}.
$$
Example

Let $T^{(i)+a(j)} \in \mathbb{C}^{p \times p}$ be the matrix whose entries are identical to the matrix $I_p$ with the exception of the element located in row $i$ and column $j$ that equals $a$:

$$T^{(i)+a(j)} = \begin{pmatrix}
1 & 0 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \\
0 & 1 & \cdots & \cdots & \cdots & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & 0 \\
0 & 0 & \cdots & a & \cdots & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \ddots & 1
\end{pmatrix}.$$

The result of the multiplication $T^{(i)+a(j)}A$ is a matrix that can be obtained from $A$ by adding the $j^{\text{th}}$ line of $A$ multiplied by $a$ to the $i^{\text{th}}$ line of $A$. The inverse of the matrix $T^{(i)+a(j)}A$ is $T^{(i)-a(j)}A$. 
Example

We have

\[ T^{(4)} + 2(2) A = \]

\[
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 2 & 0 & 1
\end{pmatrix}
\]

\[
\begin{pmatrix}
a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\
a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\
a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\
a_{41} & a_{42} & a_{43} & a_{44} & a_{45}
\end{pmatrix}
\]

\[
= \begin{pmatrix}
a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\
a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\
a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\
a_{41} + 2a_{21} & a_{42} + 2a_{22} & a_{43} + 2a_{23} & a_{44} + 2a_{24} & a_{45} + 2a_{25}
\end{pmatrix}.
\]
It is easy to see that if one multiplies a matrix $A$ at the right by $T^{(i)\leftrightarrow(j)}$, $T^{a(i)}$, and $T^{(i)+a(j)}$ the effect on $A$ consists of exchanging the columns $i$ and $j$, multiplying the $i^{th}$ column by $a$, and adding the $j^{th}$ column multiplied by $a$ to the $i^{th}$ column, respectively.
Definition

Let $\mathbb{F}$ be a field, $A, C \in \mathbb{F}^{m \times n}$ and let $b, d \in \mathbb{F}^{m \times 1}$. Two systems of linear equations $Ax = b$ and $Cx = d$ are **equivalent** if they have the same set of solutions.

If $Ax = b$ is a system of linear equations in matrix form, where $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^{m \times 1}$, and $T \in \mathbb{C}^{m \times m}$ is a matrix that has an inverse, then the systems $Ax = b$ and $(TA)x = (Tb)$ are equivalent. Indeed, any solution of $Ax = b$ satisfies the system $(TA)x = (Tb)$. Conversely, if $(TA)x = (Tb)$, by multiplying this equality by $T^{-1}$ to the left, we get $(T^{-1}T)Ax = (T^{-1}T)b$, that is, $Ax = b$. 
The matrices $T^{(i)\leftrightarrow(j)}$, $T^{a(i)}$, and $T^{(i)+a(j)}$ play a special role in an algorithm that transforms a linear system $Ax = b$ into an equivalent system in row echelon form. These transformations are known as *elementary transformation matrices*. 
Algorithm for the Row Echelon Form of a Matrix

**Data:** An matrix $A \in \mathbb{F}^{p \times q}$.

**Result:** A row echelon form of $A$.

$r = 1; c = 1;$

**while** ($r \leq p$ and $c \leq q$) **do** {

**while** ($A(\ast, c) = 0$) **{** c=c+1 }

$j = r;$

**while** ($A(j, c) = 0$) **{** j = j+1 }

**if** ($j \neq r$) **{** exchange line $r$ with line $j$ }

multiply line $r$ by $\frac{1}{A(r,c)}$

**ForEach** ($k = r + 1$ to $p$)

**{** add line $r$ multiplied by $-A(k, c)$ to line $k$ }

$r = r + 1; c = c + 1;$

**}
Example

Consider the linear system

\[
\begin{align*}
    x_1 + 2x_2 + 3x_3 &= 4 \\
    x_1 + 2x_2 + x_3 &= 3 \\
    x_1 + 3x_2 + x_3 &= 1.
\end{align*}
\]

The augmented matrix of this system is

\[
[A|b] = \begin{pmatrix}
    1 & 2 & 3 & 4 \\
    1 & 2 & 1 & 3 \\
    1 & 3 & 1 & 1
\end{pmatrix}.
\]

By subtracting the first row from the second and the third we obtain the matrix

\[
T^{(3)-1(1)} T^{(2)-1(1)} [A|b] = \begin{pmatrix}
    1 & 2 & 3 & 4 \\
    0 & 0 & -2 & -1 \\
    0 & 1 & -2 & -3
\end{pmatrix}.
\]
Example

Next, the second and third row are exchanged yielding the matrix

\[
T^{(2)\leftrightarrow (3)} T^{(3)-1(1)} T^{(2)-1(1)} [A|b] = \begin{pmatrix}
1 & 2 & 3 & 4 \\
0 & 1 & -2 & -3 \\
0 & 0 & -2 & -1
\end{pmatrix}.
\]

To obtain an 1 in the pivot of the third row we multiply the third row by \(-\frac{1}{2}\):

\[
T^{-0.5(3)} T^{(2)\leftrightarrow (3)} T^{(3)-1(1)} T^{(2)-1(1)} [A|b] = \begin{pmatrix}
1 & 2 & 3 & 4 \\
0 & 1 & -2 & -3 \\
0 & 0 & 1 & 0.5
\end{pmatrix},
\]

which is the row echelon form of the matrix \([A|b]\).
Example

To achieve the row echelon form we needed to multiply the matrix $[A|b]$ by the matrix

$$T = T_{-0.5(3)} T_{(2)\leftrightarrow(3)} T_{(3)-1(1)} T_{(2)-1(1)}.$$

The solutions of the system can now be obtained by back substitution from the linear system

$$x_1 + 2x_2 + 3x_3 = 4,$$
$$x_2 - 2x_3 = -3,$$
$$x_3 = 0.5.$$

The last equation yields $x_3 = 0.5$. Substituting $x_3$ in the second equation implies $x_2 = -2$; finally, from the first equality we have $x_1 = 6.5$. 
Theorem

Let $T^{a(i)}$, $T^{(p)\leftrightarrow(q)}$, and $T^{(i)+a(j)}$ be the matrices in $\mathbb{R}^{m\times m}$ that correspond to the row transformations applied to matrices in $\mathbb{R}^{m\times n}$, where $i \neq j$ and $p \neq q$. We have:

$$T^{a(i)} T^{(p)\leftrightarrow(q)} = \begin{cases} T^{(p)\leftrightarrow(q)} T^{a(i)} & \text{if } i \not\in \{p, q\}, \\ T^{(p)\leftrightarrow(q)} T^{a(q)} & \text{if } i = p, \\ T^{(p)\leftrightarrow(q)} T^{a(p)} & \text{if } i = q, \end{cases}$$

$$T^{(i)+a(j)} T^{(p)\leftrightarrow(q)} = \begin{cases} T^{(p)\leftrightarrow(q)} T^{(i)+a(j)} & \text{if } \{i, j\} \cap \{p, q\} = \emptyset, \\ T^{(q)+a(j)} T^{(p)\leftrightarrow(q)} & \text{if } i = p \text{ and } j \neq q, \\ T^{(i)+a(p)} T^{(p)\leftrightarrow(q)} & \text{if } i \neq p \text{ and } j = q, \\ T^{(q)+a(p)} T^{(p)\leftrightarrow(q)} & \text{if } i = p \text{ and } j = q. \end{cases}$$
The matrices that describe elementary transformations are of two types: lower triangular matrices of the form $T^a(i)$ or $T(i) + a(j)$ or permutations matrices of the form $T(p) \leftrightarrow(q)$.

If all pivots encountered in the construction of the row echelon form of the matrix $A$ are not-zero then there is no need to have any permutation matrix $T(p) \leftrightarrow(q)$ among the matrices that multiply $A$ at the left. Thus, there is a lower matrix $T$ and an upper triangular matrix $U$ such that $TA = U$.

The matrix $T$ is a product of invertible matrices and therefore it is invertible. Since the inverse $L = T^{-1}$ of a lower triangular matrix is lower triangular, it follows that $A = LU$; in other words, $A$ can be decomposed into a product of a lower triangular and an upper triangular matrix. This decomposition is known as an $LU$-decomposition of $A$. 

Prof. Dan A. Simovici
CS724: Topics in Algorithms Solving Linear Systems Slide Set 6
Example

Let $A \in \mathbb{R}^{3 \times 3}$ be the matrix

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 1 & -1 & 2 \end{pmatrix}$$

Initially, we add the first row multiplied by $-2$ to the second row, and the same first row, multiplied by $-1$ to the third row. This amounts to

$$T^{(3),-1} T^{(2),-2} A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$
Example

Next, we add the second row to the third to produce the matrix

\[
T^{(3)+(2)} T^{(3)}, -(1) T^{(2)}, -2(1) A = \begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & -1 \\
0 & 0 & 0
\end{pmatrix},
\]

which is an upper triangular matrix. We can conclude that \( \text{rank}(A) = 2 \)
and we can write:

\[
A = (T^{(2), -2(1)})^{-1} (T^{(3), -(1)})^{-1} (T^{(3)+(2)})^{-1} \begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & -1 \\
0 & 0 & 0
\end{pmatrix}
\]
Example

Thus, the lower triangular matrix we are seeking is

\[ L = \left( T^{(2)} - 2(1) \right)^{-1} \left( T^{(3)} + (1) \right)^{-1} \left( T^{(3)} + (2) \right)^{-1} = \left( T^{(2)} + 2(1) \right) \left( T^{(3)} + (1) \right) \left( T^{(3)} - (2) \right) \]

\[
= \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix},
\]

which shows that \( A \) can be written as:

\[
A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix},
\]

where the first matrix is lower triangular and the second is upper triangular.
Suppose that during the construction of the matrix $U$ some of the elementary transformation matrices are permutations matrices of the form $T(p)\leftrightarrow(q)$.

Matrices of the form $T(p)\leftrightarrow(q)$ can be shifted to the right. Therefore, instead on the previous factorization of the matrix $A$ we have a lower triangular matrix $T$ and a permutation matrix (which results as a product of all permutation matrices of the form $T(p)\leftrightarrow(q)$ used in the algorithm such that $TPA = U$. In this case we obtain an $LU$-factorization of $PA$ instead of $A$. 
Theorem

Let $A \in \mathbb{R}^{m \times n}$ be a matrix. If $\text{rank}(A) = k$, then a largest non-singular square submatrix $B$ of $A$ is a $k \times k$-matrix.
The function `inv` computes the inverse of an invertible square matrix.

Example

Let $A$ be the invertible matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$

Its inverse is given by;

$$\gg \text{inv}(A)$$

$$\text{ans} = \begin{pmatrix} -3.0000 & 2.0000 \\ 2.0000 & -1.0000 \end{pmatrix}$$
Example

On other hand, if inv is applied to a singular matrix

\[ A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \]

an error message is posted:

Warning: Matrix is singular to working precision.
If $A$ is non-singular, the function $\text{inv}$ can be used to solve the system $A\mathbf{x} = \mathbf{b}$ by writing $\mathbf{x} = \text{inv}(A)\mathbf{b}$, although a better method is described below.

**Example**

For

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 13 \\ 23 \end{pmatrix}$$

the solution of the system $A\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} = \text{inv}(A)\mathbf{b}$$

$\mathbf{x} =$

7.0000  
3.0000
This is not the best way for solving a system of linear equations. In certain circumstances, this method produces errors and has a poor time performance.

A better approach is for solving a linear system $Ax = b$ is to use the backslash operator $x = A \backslash b$ or $x = \text{mldivide}(A,b)$. The term \text{mldivide} is related to the position of the matrix $A$ at the left of $x$. 
Example

Define \( A \) and \( b \) as

\[
\begin{align*}
\text{>> } A &= \begin{bmatrix} 5 & 11 & 2 \\ 10 & 6 & -4 \\ -2 & 9 & 7 \end{bmatrix} \\
A &= \\
&= \\
&= \\
\text{>> } b &= \begin{bmatrix} 53 \\ 26 \\ 48 \end{bmatrix} \\
b &= \\
&= \\
&=
\end{align*}
\]

Then either \( x=A\backslash b \) or \( x=mldivide(A,b) \) produces

\[
\begin{align*}
x &= \\
&= \\
&= 1.0000 \\
&= 4.0000 \\
&= 2.0000
\end{align*}
\]
The system $xA = c$, where $A \in \mathbb{R}^{n \times m}$ and $c' \in \mathbb{R}^m$ can be solved using either $x = A / b$ or $x = \text{mrdivide}(A,b)$.

It is easy to see that these operations are related by:

$$A\backslash b = (A'/b')'.$$
The function \texttt{rref} produces the reduced row echelon form of a matrix \( A \), when called as \( R = \texttt{rref}(A) \).

A variant of this function, \( [R,r] = \texttt{rref}(A) \) also yields a vector \( r \) so that \( r \) indicates the non-zero pivots, \( \text{length}(r) \) is the rank of \( A \), and \( A(:,r) \) is a basis for the range of \( A \). Roundoff errors may cause this algorithm to produce a rank for \( A \) that is different from the actual rank.

A pivot tolerance \( \texttt{tol} \) used by the algorithm to determine negligible columns can be specified using \( \texttt{rref}(A,\texttt{tol}) \).
Example

Starting from the matrix

\[
A =
\begin{bmatrix}
1 & 2 & 3 & 4 & 5 & 6 \\
7 & 8 & 9 & 10 & 11 & 12 \\
1 & 3 & 5 & 7 & 9 & 11 \\
\end{bmatrix}
\]

the function call \([R,r]=rref(A)\) returns

\[
R =
\begin{bmatrix}
1 & 0 & -1 & -2 & -3 & -4 \\
0 & 1 & 2 & 3 & 4 & 5 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

\[
r =
\begin{bmatrix}
1 & 2 \\
\end{bmatrix}
\]

showing that the rank of \(A\) is 2.
Let $Au = b$ be a linear system, where $A \in \mathbb{C}^{n \times n}$ is a non-singular matrix and $b \in \mathbb{R}^n$.

We examine the sensitivity of the solution of this system to small variations of $b$. So, together with the original system, we work with a system of the form $Av = b + h$, where $h \in \mathbb{R}^n$ is the perturbation of $b$.

Note that $A(v - u) = h$, so $v - u = A^{-1}b$. Using a vector norm $\| \cdot \|$ and its corresponding matrix norm $\| \cdot \|$ we have

$$\| v - u \| = \| A^{-1}h \| \leq \| A^{-1} \| \| h \|.$$
Since $\| b \| = \| A u \| \leq \| A \| \| u \|$ it follows that

$$\frac{\| v - u \|}{\| u \|} \leq \frac{\| A^{-1} \| \| h \|}{\| b \| \| A \|} = \frac{\| A \| \| A^{-1} \| \| h \|}{\| b \|}.$$  (1)

Thus, the relative variation of the solution, $\frac{\| v - u \|}{\| u \|}$ is upper bounded by the number $\frac{\| A \| \| A^{-1} \| \| h \|}{\| b \|}$.
Definition

Let $A \in \mathbb{C}^{n \times n}$ be a non-singular matrix. The condition number of $A$ relative to the matrix norm $\|\cdot\|$ is the number $\text{cond}(A) = \|A\|\|A^{-1}\|$. 
The Equality

\[
\frac{\|v - u\|}{\|u\|} = \frac{\|A\| \|A^{-1}\| \|h\|}{\|b\|}.
\]

implies that if the condition number is large, then small variations in \(b\) may generate large variations in the solution of the system \(Au = b\), especially when \(b\) is close to \(0\). When this is the case, we say that the system \(Au = b\) is \textit{ill-conditioned}. Otherwise, the system \(Au = b\) is \textit{well-conditioned}.
Theorem

Let $A \in \mathbb{C}^{n \times n}$ be a non-singular matrix. The following statements hold for every matrix norm induced by a vector norm:

- $\text{cond}(A) = \text{cond}(A^{-1})$;
- $\text{cond}(cA) = |c| \text{cond}(A)$;
- $\text{cond}(A) \geq 1$. 

Proof.

We prove here only Part (iii). Since $AA^{-1} = I$, by the properties of a matrix norm induced by a vector norm we have $\text{cond}(A) = \|A\| \|A^{-1}\| \geq \|AA^{-1}\| = \|I_n\| = 1$. 

\[\square\]
Let $A, B$ be two non-singular matrices in $\mathbb{C}^{n \times n}$ such that $B = aA$, where $a \in \mathbb{C}$. We have $B^{-1} = aA^{-1}$, $\|B\| = |a|\|B\|$ and $\|B^{-1}\| = |a|\|A^{-1}\|$ so $\text{cond}(B) = |a|^2\text{cond}(A)$. On another hand, $\det(B) = a^n \det(A)$. Thus, if $n$ is large enough and $a < 1$, then $\det(B)$ can be quite close to 0, while the condition number of $B$ may be quite large. This shows that the determinant and the condition number are relatively independent.
Example

Let $A \in \mathbb{C}^{2 \times 2}$ be the matrix

$$A = \begin{pmatrix} a & a + \alpha \\ a + \alpha & a + 2\alpha \end{pmatrix},$$

where $a > 0$ and $\alpha < 0$. We have

$$A^{-1} = \begin{pmatrix} -\frac{a+2\alpha}{\alpha^2} & \frac{a+\alpha}{\alpha^2} \\ \frac{a+\alpha}{\alpha^2} & -\frac{a}{\alpha^2} \end{pmatrix},$$

so $\|A\|_1 = a$ and $\|A^{-1}\| = \frac{a}{\alpha^2}$. Thus, $\text{cond}(A) = \left(\frac{a}{\alpha}\right)^2$ and, if $|\alpha|$ is small a system of the for $Au = b$ may be ill-conditioned.
Ill-conditioned linear systems $Au = b$ may occur when large differences in scale exists among the columns of $A$, or among the rows of $A$.

**Theorem**

Let $A = (a_1 \cdots a_n)$ be an invertible matrix in $\mathbb{C}^{n \times n}$, where $a_1, \ldots, a_n$ are the columns of $A$. Then,

$$cond(A) \geq \max \left\{ \frac{\|a_i\|}{\|a_j\|} \mid 1 \leq i, j \leq n \right\}.$$
Proof

Since \( \text{cond}(A) = \|A\| \|A^{-1}\| \), we have

\[
\text{cond}(A) = \frac{\max\{\|Ax\| | \|x\| = 1\}}{\min\{\|Ax\| | \|x\| = 1\}}
\]

Note that \( Ae_k = a_k \), where \( a_k \) is the \( k \)th column of \( A \) and that \( \| e_k \| = 1 \). Therefore,

\[
\max\{\|Ax\| | \|x\| = 1\} \geq \| a_i \|, \quad \min\{\|Ax\| | \|x\| = 1\} \leq \| a_j \|,
\]

which implies

\[
\text{cond}(A) \geq \frac{\| a_i \|}{\| a_j \|}
\]

for all \( 1 \leq i, j \leq n \). This yields the inequality of the theorem.
Example

Let

\[ A = \begin{pmatrix} 1 & 0 \\ 1 & \alpha \end{pmatrix}, \]

where \( \alpha \in \mathbb{R} \) and \( \alpha > 0 \). The matrix \( A \) is invertible and

\[ A^{-1} = \begin{pmatrix} 1 & 0 \\ -\frac{1}{\alpha} & \frac{1}{\alpha} \end{pmatrix}. \]

It is easy to see that the condition number of \( A \) relative to the Frobenius norm is

\[ \text{cond}(A) = \frac{2 + \alpha^2}{\alpha}. \]

Thus, if \( \alpha \) is sufficiently close to 0, the condition number can reach arbitrarily large values.
In general, we use as matrix norms, norms of the form $\| \cdot \|_p$. The corresponding condition number of a matrix $A$ is denoted by $\text{cond}_p(A)$.

**Example**

Let $A = \text{diag}(a_1, \ldots, a_n)$ be a diagonal matrix. Then,  
$$\| A \|_2 = \max_{1 \leq i \leq n} |a_i|.$$  
Since $A^{-1} = \text{diag}\left(\frac{1}{a_1}, \ldots, \frac{1}{a_n}\right)$, it follows that  
$$\| A^{-1} \|_2 = \frac{1}{\min_{1 \leq i \leq n} |a_i|},$$  
so $\text{cond}_2(A) = \frac{\max_{1 \leq i \leq n} |a_i|}{\min_{1 \leq i \leq n} |a_i|}$. 
The condition number of a matrix $A$ is computed using the function $\text{cond}(A,p)$ which returns the $p$-norm condition of matrix $A$. When used with a single parameter, as in $\text{cond}(A)$, the 2-norm condition number of $A$ is returned.
Example

Let $A$ be the matrix

$$A = \begin{bmatrix} 10.1 & 6.2 \\ 5.1 & 3.1 \end{bmatrix}$$

The condition number $\text{cond}(A)$ is $567.966$, which is quite large indicating significant sensitivity to inverse calculations. The inverse of $A$ is

$$\text{inv}(A) = \begin{bmatrix} -10.0000 & 20.0000 \\ 16.4516 & -32.5806 \end{bmatrix}$$
Example

If we make a small change in $A$ yielding the matrix

```matlab
>> B=[10.2 6.3;5.1 3.1]
B =
    10.2000   6.3000
    5.1000   3.1000
```

the inverse of $B$ changes completely:

```matlab
>> inv(B)
ans =
   -6.0784  12.3529
    10.0000 -20.0000
```
Values of the condition number close to 1 indicate a well-conditioned matrix, and the opposite is true for large values of the condition number.

**Example**

Consider the linear systems:

\[
\begin{align*}
10.1x_1 + 6.2x_2 &= 12 \\
5.1x_1 + 3.1x_2 &= 6
\end{align*}
\]

and

\[
\begin{align*}
10.2x_1 + 6.3x_2 &= 12 \\
5.1x_1 + 3.1x_2 &= 6
\end{align*}
\]

that correspond to \(Ax = b\) and \(Bx = b\), where \(b = \begin{pmatrix} 12 \\ 6 \end{pmatrix}\). In view of the resemblance of \(A\) and \(B\) one would expect their solutions to be close.
Example

However, this is not the case. The solution of $Ax = b$ is

```matlab
>> x=inv(A)*b
x =
    0
   1.9355
```

while the solution of $Bx = b$ is

```matlab
>> x=inv(B)*b
x =
   1.1765
   1.1765
    0
```