Solving Systems of Linear Equations

Given a set of equations,
\[
\begin{align*}
  x_1 - x_2 &= -2 \\
  2x_1 - 5x_2 &= -7
\end{align*}
\]
First check the determinant: \((1)(-5) - (-1)(2) = -3\)

The following is what you did in high school

\[
\begin{align*}
  x_1 - x_2 &= -2 \\
  2x_1 - 5x_2 &= -7
\end{align*}
\]
\[
\begin{align*}
  x_1 &= -1 \\
  x_2 &= 1
\end{align*}
\]

Matrix Form:

\[
\begin{pmatrix}
  1 & -1 \\
  2 & -5
\end{pmatrix}
\begin{pmatrix}
  x_1 \\
  x_2
\end{pmatrix}
= \begin{pmatrix}
  -2 \\
  -7
\end{pmatrix}
\]

\[
\begin{pmatrix}
  1 & -1 \\
  0 & -3
\end{pmatrix}
\begin{pmatrix}
  x_1 \\
  x_2
\end{pmatrix}
= \begin{pmatrix}
  -2 \\
  -3
\end{pmatrix}
\]

\[
\begin{pmatrix}
  1 & 0 \\
  0 & 1
\end{pmatrix}
\begin{pmatrix}
  x_1 \\
  x_2
\end{pmatrix}
= \begin{pmatrix}
  -1 \\
  1
\end{pmatrix}
\]

\[\mathbf{U}\] is an upper triangular matrix.

Definition: \(\mathbf{x} = \mathbf{A} \backslash \mathbf{b}\) denotes the solution of \(\mathbf{x}\) obtained using Gaussian Elimination

Importance: Mathematically \(\mathbf{A} \backslash \mathbf{b} = \mathbf{A}^{-1} \mathbf{b}\), but computationally we do not need \(\mathbf{A}^{-1}\) when calculating \(\mathbf{A} \backslash \mathbf{b}\). So \(\mathbf{A} \backslash \mathbf{b}\) is more efficient.

In Matlab, \(\mathbf{A} \backslash \mathbf{b}\) is executed by \(\mathbf{A} \backslash \mathbf{b}\) and \(\mathbf{A}^{-1} \mathbf{b}\) is executed by \(\text{inv}(\mathbf{A}) \ast \mathbf{b}\)
LU decomposition

Purpose: Given a square $n \times n$ matrix $A$, decompose $A$ into a product of two square matrices: $A = LU$, where $L$ is lower triangular and $U$ is upper triangular.

$U$ is obtained using Gaussian Elimination. But how to get $L$? $L$ is to record how we make a zero in $A$. First write

$$L = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Consider this intermediate step:

$$\begin{pmatrix} 1 & -1 \\ 2 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 \\ -7 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -1 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 \\ -3 \end{pmatrix}$$

We have multiplied 2 to the first row to use it to subtract the second row, resulting a zero in the $(2,1)$-th element. So put 2 into the same location where $A$ is made zero:

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

Check:

$$LU = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & -3 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 2 & -5 \end{pmatrix} = A$$

Then

$$Ax = b \Rightarrow (LU)x = b \Rightarrow Ux = Lb \Rightarrow x = U \backslash (L \backslash b)$$

Here

$A = LU$ manifests Gaussian Elimination $\sim O\left(\frac{2}{3}n^3\right)$ float-point operations (flops)

$L \backslash b$ manifests forward substitution $\sim O(n^2)$ flops

$U \backslash b$ manifests back substitution $\sim O(n^2)$ flops
Pivoting and Permutation

There is situation where \( A = LU \) cannot be done. Consider

\[
\begin{align*}
x_2 &= 1 \\
2x_1 - 5x_2 &= -7
\end{align*}
\]

or equivalently,

\[
\begin{pmatrix} 0 & 1 \\ 2 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -7 \end{pmatrix}
\]

This system cannot be made upper triangular by Gaussian Elimination. But one can swap the order of the rows of \( A \) by applying a permutation matrix

\[
P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\]

so that

\[
A' = PA = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 2 & -5 \end{pmatrix} = \begin{pmatrix} 2 & -5 \\ 0 & 1 \end{pmatrix}
\]

And \( A' \) is LU-decomposable (although not necessary in this example). In Matlab, when \( A \backslash b \) is executed, the rows of \( A \) are first rearranged by a permutation matrix \( P \) to avoid zero pivots. Then \( P \cdot A \) is decomposed into \( L \) and \( U \). Finally, \( x \) is obtained by \( U \backslash (L \cdot P \cdot b) \).

\[
A x = b \Rightarrow (P \cdot A) x = P b \Rightarrow (L \cdot U) x = P b \Rightarrow U x = L \backslash (P b) \Rightarrow x = U \backslash (L \backslash (P b))
\]

Matlab commands

\[
[L, U, P] = \text{lu}(A) \text{ such that } P \cdot A = L \cdot U
\]

\[A \backslash b \text{ or } \text{inv}(A) \ast b?\]

To obtain \( A \backslash b \), \( O \left(2.67n^3\right) \) flops are taken. To obtain \( \text{inv}(A) \ast b \), \( O \left(5.67n^3\right) \) flops are taken. So always use \( A \backslash b \).

Given any matrices, check

\[\text{det}(A): \text{If } |\text{det}(A)| >> 10^{-16}, \text{ unique solution of } x \text{ exists.}\]

But does not mean accurate. Check

\[\text{cond}(A): \text{Fractional error } \frac{\delta x}{x} = \text{cond}(A) \times \epsilon, \text{ where } \epsilon = 10^{-16} \text{ is the machine precision.}\]