Chapter 8

8.1 Gaussian Elimination

The process of Gaussian elimination is a fundamental tool in solving linear systems of equations.

Example 1: Consider the linear system:

$$x_1 + x_2 + x_3 = 3 \tag{8.1}$$
$$x_1 + 2x_2 + 4x_3 = 7 \tag{8.2}$$
$$x_1 + 3x_2 + 9x_3 = 13 \tag{8.3}$$

The traditional way of solving this system is to subtract the first equation from the second and the third to obtain

$$x_1 + x_2 + x_3 = 3 \tag{8.4}$$
$$x_2 + 3x_3 = 4 \tag{8.5}$$
$$2x_2 + 8x_3 = 12 \tag{8.6}$$

Now subtract 2 times the second equation from the third to obtain

$$x_1 + x_2 + x_3 = 3 \tag{8.7}$$
$$x_2 + 3x_3 = 4 \tag{8.8}$$
$$2x_3 = 2 \tag{8.9}$$

Now we can perform *back substitution*:

$$x_3 = 1 \tag{8.10}$$

$$x_2 = 3 - 3x_3 = 4 - 3 \cdot 1 = 1 \tag{8.11}$$
$$x_1 = 3 - x_2 - x_3 = 3 - 1 - 1 = 1. \tag{8.12}$$

Instead of performing the same process for *every* right hand side, it is more advantageous to use matrix factorizations instead. Write the System 8.1-8.3 as

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \\ 13 \end{bmatrix} \tag{8.13}$$

In general linear systems are written as

$$Ax = b \tag{8.14}$$

or

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{8.15}$$

where we assume that leading principal minors $A(1:k, 1:k)$, $k = 1, 2, \dots, n$, of the n-by-n matrix $A = [a_{ij}]_{i,j=1}^n$ are nonzero.

Some linear systems are easy to solve. For example if A is *triangular* or *diagonal*.

If A is (lower or upper) triangular nonsingular matrix then $Ax = b$ can be solved via *back substitution*. The system

$$\begin{bmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & \vdots & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{8.16}$$

(the zero entries of the upper triangular part have been omitted) is equivalent to

$$a_{11}x_1 = b_1 \tag{8.17}$$
$$a_{21}x_1 + a_{22}x_2 = b_2 \tag{8.18}$$
$$\dots$$
$$a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n = b_n \tag{8.19}$$

and is solved by computing x_1 from the first equation, substituting into the second and so on:

$$x_1 = \frac{b_1}{a_{11}} \tag{8.20}$$

$$x_2 = \frac{1}{a_{21}}(b_2 - a_{21}x_1) \tag{8.21}$$

$$\dots$$

$$x_n = \frac{1}{a_{nn}}(b_n - a_{n1}x_1 - a_{n2}x_2 - \dots - a_{n,n-1}x_{n-1}) \tag{8.22}$$

The solution to a diagonal linear system is trivial:

$$\begin{bmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{8.23}$$

implies $x_i = \frac{b_i}{d_i}$, $i = 1, 2, \dots, n$.

Definition: A matrix A is called *unit lower triangular* if $a_{ij} = 0$, $1 \leq i < j \leq n$ and $a_{ii} = 1$, $1 \leq i \leq n$.

An *unit upper triangular* matrix is defined analogously.

Example: The following 4-by-4 matrix is unit lower triangular

$$\begin{bmatrix} 1 & & & \\ 2 & 1 & & \\ 3 & 5 & 1 & \\ 4 & 6 & 7 & 1 \end{bmatrix} \tag{8.24}$$

Exercise: Prove that if A and B are unit lower triangular matrices, then so are A^{-1} and AB.

Definition: Let A be a nonsingular matrix, then a decomposition of A as a product of a unit lower triangular matrix L, a diagonal matrix D and a unit upper triangular matrix U:

$$A = LDU \tag{8.25}$$

is called *an LDU decomposition of A*.

The main idea in what follows is to use Gaussian elimination to compute the LDU decomposition of A.

Once we have the LDU decomposition of A, the equation $Ax = b$ becomes $LDUx = b$, which is easy to solve. First compute the solution y to the lower triangular system $Ly = b$, then the solution z to the diagonal system $Dz = y$, and finally the solution x to the upper triangular system $Ux = z$. Finally,

$$Ax = LDUx = LD(Ux) = L(Dz) = Ly = b \tag{8.26}$$

as desired.

So how does one compute the LDU decomposition of a nonsingular matrix A?

First we represent a subtraction of a multiple of one row from another in matrix form. Consider the matrix:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 3 & 9 & 27 \end{bmatrix} \tag{8.27}$$

In order to introduce a zero in position $(3, 1)$ we need to subtract 3 times the first row from the third. This is equivalent to multiplication by the matrix

$$\begin{bmatrix} 1 & & \\ 0 & 1 & \\ -3 & 0 & 1 \end{bmatrix} \tag{8.28}$$

namely

$$\begin{bmatrix} 1 & & \\ 0 & 1 & \\ -3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 3 & 9 & 27 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 0 & 6 & 24 \end{bmatrix} \tag{8.29}$$

Since

$$\begin{bmatrix} 1 & & \\ 0 & 1 & \\ -3 & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 3 & 0 & 1 \end{bmatrix} \tag{8.30}$$

the equality 8.29 implies

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 3 & 9 & 27 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 0 & 6 & 24 \end{bmatrix} \tag{8.31}$$

Next, subtract the first row from the second to analogously obtain

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 0 & 6 & 24 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 6 & 24 \end{bmatrix} \tag{8.32}$$

Now observe that the matrices used for elimination combine very nicely:

$$\begin{bmatrix} 1 & & \\ 0 & 1 & \\ 3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 3 & 0 & 1 \end{bmatrix} \tag{8.33}$$

therefore

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 3 & 9 & 27 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 6 & 24 \end{bmatrix} \tag{8.34}$$

Then continue by induction–subtract 6 times the second row from the third, obtaining the decomposition

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 6 & 24 \end{bmatrix} = \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & 6 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 6 \end{bmatrix} \tag{8.35}$$

Therefore

$$\begin{aligned} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 3 & 9 & 27 \end{bmatrix} &= \underbrace{\begin{bmatrix} 1 & & \\ 1 & 1 & \\ 3 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & & \\ 0 & 1 & \\ 0 & 6 & 1 \end{bmatrix}} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 6 \end{bmatrix} & (8.36) \\ &= \begin{bmatrix} 1 & & \\ 1 & 1 & \\ 3 & 6 & 1 \end{bmatrix} \cdot \overbrace{\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 6 \end{bmatrix}} & (8.37) \\ &= \underbrace{\begin{bmatrix} 1 & & \\ 1 & 1 & \\ 3 & 6 & 1 \end{bmatrix}}_{L} \cdot \underbrace{\begin{bmatrix} 1 & & \\ & 1 & \\ & & 6 \end{bmatrix}}_{D} \cdot \underbrace{\begin{bmatrix} 1 & 1 & 1 \\ & 1 & 3 \\ & & 1 \end{bmatrix}}_{U} & (8.38) \end{aligned}$$

Algorithm: [Gaussian Elimination] The following algorithm computes the LDU decomposition of a matrix A whose leading principal minors are nonzero.

$U = A$, $L = I$, $D = I$
for $i = 1 : n-1$
  for $j = i+1 : n$
    $l_{ji} = u_{ji}/u_{ii}$
    $u_{j,i:n} = u_{j,i:n} - l_{ji}u_{i,i:n}$
  endfor
  $d_{ii} = u_{ii}$
  $u_{i,i:n} = u_{i,i:n}/d_{ii}$
endfor

8.2 Pivoting, Partial and Complete

$$P_1 A = L_1 A_1 \tag{8.39}$$

$$
\begin{aligned}
P_2A_2 &= L_2A_2 && (8.40)\\
\dots & \\
A &= P_1^T L_1 A_1 \\
&= P_1^T L_1 P_2^T L_2 A_2 \\
&= P_1^T L_1 P_2^T L_2 P_3^T L_3 U \\
&= P_1^T P_2^T P_3^T ((P_2^T P_3^T)^{-1} L_1 P_2^T P_3^T)(P_3 L_2 P_3^T) L_3 U \\
&= P_1^T P_2^T P_3^T L_1 L_2 L_3 U \\
&= P^T LU && (8.41)
\end{aligned}
$$

Complete pivoting: two sided

$$A = P^T LUQ^T \quad (8.42)$$

Operation count: $\frac{2}{3}n^3$ for GENP and GEPP, for complete pivoting

$$
\begin{aligned}
n^2 + (n-1)^2 + \cdots &= \frac{n(n+1)(2n+1)}{6} \\
&= \frac{2}{3}n^3 \textit{ additional} && (8.43)
\end{aligned}
$$

8.3 Stability of GE

Example:

$$
\begin{aligned}
A &= LU && (8.44)\\
\begin{bmatrix} 10^{-16} & 1 \\ 1 & 1 \end{bmatrix} &= \begin{bmatrix} 1 & 0 \\ 10^{16} & 1 \end{bmatrix} \begin{bmatrix} 10^{-16} & 1 \\ 0 & 1-10^{16} \end{bmatrix} && (8.45)\\
\|A\| &= O(1) && (8.46)\\
\|L\|, \|U\| &= O(10^{16}) && (8.47)\\
Ax &= b && (8.48)\\
LUx &= x && (8.49)\\
Ly &= b && (8.50)\\
Ux = y & && (8.51)\\
(L+\delta L)\hat{y} &= b && (8.52)\\
\frac{\|\delta L\|}{\|L\|} &= O(\epsilon) && (8.53)\\
\|\delta L\| &= O(1) && (8.54)\\
(U+\delta U)\hat{x} &= y && (8.55)
\end{aligned}
$$

$$\begin{aligned}
\frac{\|\delta U\|}{\|U\|} &= O(\epsilon) && (8.56)\\
\|\delta U\| &= O(1) && (8.57)\\
(L+\delta L)(U+\delta U)\hat{x} &= b && (8.58)\\
A + \underbrace{\delta LU + U\delta L + \delta L\delta U}_{\delta A} &= O(1\cdot 10^{16} + 1\cdot 10^{16} + 1\cdot 1)\\
&= O(10^{16}) && (8.59)\\
\frac{\|\delta A\|}{\|A\|} &= O(10^{16}), && (8.60)
\end{aligned}$$

while we expected ϵ.

Theorem: A: nonsingular. Let $A = LU$ be computed by GENP in floating point arithmetic. If A has an LU factorization, then for sufficiently small $\epsilon_{\text{machine}}$, the factorization completes successfully and

$$\tilde{L}\tilde{U} = A + \delta A \tag{8.61}$$

where,

$$\frac{\|\delta A\|}{\|L\|\cdot\|U\|} = O(\epsilon_{\text{machine}}) \tag{8.62}$$

for some $\delta A \in \mathcal{C}^{n\times n}$.

Backward Stability?

Need

$$\frac{\|\delta A\|}{\|A\|} = O(\epsilon_{\text{machine}}) \tag{8.63}$$

then

$$\textit{backward stability} \Leftrightarrow \|L\|\cdot\|U\| = O(\|A\|) \quad ? \tag{8.64}$$

Need to measure $\frac{\|L\|\cdot\|U\|}{\|A\|}$ and make sure it is $O(\epsilon_{\text{machine}})$.

No pivoting: unstable.

Partial pivoting:

$$\|L\|_\infty \leq n \tag{8.65}$$

since

$$|l_{ij}| < 1 \tag{8.66}$$

so the question moves down to the size of

$$\frac{\|U\|}{\|A\|} \simeq \frac{\max|u_{ij}|}{\max|a_{ij}|} \equiv \rho \tag{8.67}$$

where ρ is called growth factor.
Therefore,

$$\tilde{L}\tilde{U} = A + \delta A \tag{8.68}$$

$$\frac{\|\delta A\|}{\|A\|} = O(\rho \epsilon_{\text{machine}}) \tag{8.69}$$

stable if $\rho = O(1)$.
Partial Pivoting: $\rho \leq 2^n$, attainable, but never happens. Usually $\leq n^{\frac{1}{2}}$.
Complete Pivoting: $\rho = O(1)$, but cost $\frac{4}{3}n^3$, same as QR.