Special Linear Systems
Linear systems where Gaussian elimination does not require pivoting.

Symmetric, Positive Definite Systems
Used in
1. Differential equations
2. Statistics — covariance matrices

Properties
1. $A = A^T$ symmetry
2. $x^T A x > 0$, $x \neq 0$. positive definiteness

All such matrices have the form

$$A = B^T B$$

for some nonsingular $B$. We will show how to construct such a $B$ (it is not unique).

You can prove that

$$\max_{1 \leq i \leq n} a_{ii} = \max_{(i,j)} |a_{ij}|.$$ 

The largest element is always on the diagonal (though it is not necessarily the pivot element).

In fact, Von Neumann showed that during Gaussian elimination

$$\max_{1 \leq i \leq n} a_{ii} = \max_{(i,j,k)} |a_{ij}^{(k)}|$$

where $a_{ij}^{(k)}$ is the value of $a_{ij}$ after $k$ elimination steps. There is no growth in the elements!

It is better than that, symmetric positive definite matrices have their own form of LU factorization! It is called Cholesky factorization (it was invented by a French artillery captain) and is about half the operations of ordinary Gaussian elimination.
It produces an upper triangular matrix $R$ such that
\[
A = R^T R.
\]

In MATLAB, it is the command
\[
\gg R = \text{chol}(A)
\]
We will develop the algorithm inductively based upon the dimension of $A$. Take $n = 1$. Then
\[
A = (a_{11}), \quad a_{11} > 0.
\]
We have that
\[
R = (r_{11}), \quad r_{11}^2 = a_{11}.
\]
By convention, we choose
\[
r_{11} = \sqrt{a_{11}}.
\]
Assume that we can do Cholesky for $n - 1$, do for $n$.
\[
A = \begin{pmatrix} n & 1 \\ 1 & n-1 \end{pmatrix} \begin{pmatrix} A_{n-1} & z \\ z^T & a_n \end{pmatrix},
\]
For instance, if $A$ is $3 \times 3$,
\[
A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}
\]
and
\[
A_{n-1} = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}, \quad z = \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}.
\]
Suppose
\[
R = \begin{pmatrix} n & 1 \\ 1 & n-1 \end{pmatrix} \begin{pmatrix} R_{n-1} & w \\ w^T & r_{nn} \end{pmatrix},
\]
Then we have that
\[
R^T R = \begin{pmatrix} R_{n-1}^T & 0 \\ w^T & r_{nn} \end{pmatrix} \begin{pmatrix} R_{n-1} & w \\ 0 & r_{nn} \end{pmatrix}
\]
\[
= \begin{pmatrix} R_{n-1}^T R_{n-1} & R_{n-1}^T w \\ w^T R_{n-1} & r_{nn}^2 + w^T w \end{pmatrix}.
\]
Matching blocks with $A$ yields

$$A_{n-1} = R_{n-1}^T R_{n-1}$$  \hspace{1cm} (1)

$$R_{n-1}^T w = z$$  \hspace{1cm} (2)

$$r_{nn}^2 + w^T w = a_{nn}$$  \hspace{1cm} (3)

The Cholesky factorization of the $(n - 1) \times (n - 1)$ principal submatrix yields $R_{n-1}$. The forward substitution (2) yields $w$, and $r_{nn}$ is recovered from

$$r_{nn} = \sqrt{a_{nn} - w^T w}.$$

This assumes that $a_{nn} - w^T w > 0$ which positive definiteness assures.

The following is a quick proof (that I did not do in class). Let

$$x = \begin{pmatrix} -R_{n-1}^{-1} w \\ 1 \end{pmatrix} = \begin{pmatrix} -A^{-1}z \\ 1 \end{pmatrix}.$$  

Then, from the definition of $w$, we have

$$Ax = \begin{pmatrix} A_{n-1} & z \\ z^T & a_{nn} \end{pmatrix} \begin{pmatrix} -A^{-1}z \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} -A_{n-1}^{-1} z^2 + z \\ -z^T A_{n-1}^{-1} z + a_{nn} \end{pmatrix} = \begin{pmatrix} n - 1 & 0 \\ a_{nn} - z^T A_{n-1}^{-1} z \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ a_{nn} - z^T R_{n-1}^{-1} w \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ a_{nn} - w^T w \end{pmatrix}$$

Thus

$$x^T A x = \begin{pmatrix} -A_{n-1}^{-1} z \\ 1 \end{pmatrix}^T \begin{pmatrix} 0 \\ a_{nn} - w^T w \end{pmatrix}$$

$$= a_{nn} - w^T w > 0.$$  

Thus if $A$ is positive definite, this value is ALWAYS positive. Now for an algorithm
function R = cholesky(A)
    n = size(A);
    R = zeros(n, n);
    R(1, 1) = sqrt(A(1, 1));
    for k = 2:n
        R(1:k-1, k) = R(1:k-1, 1:k-1)' \ A(1:k-1, k);
        % Do the forward solve R_{n-1}^T w = z;
        R(k, k) = sqrt(A(k, k) - R(1:k-1, k) * R(1:k-1, k));
        % r_{kk} = \sqrt{a_{kk} - w^T w};
    end

The Cholesky algorithm is also a test for positive definiteness. A symmetric matrix $A$ is positive definite if and only the value $A(k, k) - R(1:k-1, k)' \ R(1:k-1, k)$ is positive for every value of $k$. In floating point arithmetic, if $A$ is symmetric, positive definite and nonsingular to machine precision, the Cholesky algorithm will always “succeed ” in this sense.

Here is a $2 \times 2$ example.

**Example 1**

$$A = \begin{pmatrix} 9 & 12 \\ 12 & 41 \end{pmatrix}$$

The algorithm yields

$$
\begin{align*}
    r_{11} &= \sqrt{a_{11}} = 3 \\
    r_{12} &= \frac{a_{12}}{r_{11}} \\
    r_{22} &= \sqrt{a_{22} - r_{12}^2} = \sqrt{41 - 4^2} = 5
\end{align*}
$$

So

$$R = \begin{pmatrix} 3 & 4 \\ 0 & 5 \end{pmatrix}.$$ 

When solving

$$Ax = b$$

we solve

$$R^T y = b$$

$$Rx = y$$

Two well know examples of symmetric positive definite matrices are:
Example 2  The Hilbert matrices from the command `hilb(n)`. The 4×4 one is

\[
A = \begin{pmatrix}
1 & 1/2 & 1/3 & 1/4 \\
1/2 & 1/3 & 1/4 & 1/5 \\
1/3 & 1/4 & 1/5 & 1/6 \\
1/4 & 1/5 & 1/6 & 1/7 \\
\end{pmatrix}.
\]

Its Cholesky factor is

\[
R = \begin{pmatrix}
1.00000 & 0.50000 & 0.33333 & 0.25000 \\
0 & 0.28868 & 0.28868 & 0.25981 \\
0 & 0 & 0.07454 & 0.11180 \\
0 & 0 & 0 & 0.01890 \\
\end{pmatrix}
\]

This can be produced from formulas. For your pleasure, my private stash routine to produce it has been placed www.cse.psu.edu/~barlow/cse455/hilbert.m. It has no documentation.

A second class of example is the following matrix from an approximation to the second derivative. The 6 × 6 version is given.

Example 3

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

Its Cholesky factor (from `octave`, the MATLAB clone) is

\[
R = \begin{pmatrix}
1.414 & -0.707 & 0 & 0 & 0 & 0 \\
0 & 1.225 & -0.816 & 0 & 0 & 0 \\
0 & 0 & 1.155 & -0.866 & 0 & 0 \\
0 & 0 & 0 & 1.118 & -0.894 & 0 \\
0 & 0 & 0 & 0 & 1.095 & -0.913 \\
0 & 0 & 0 & 0 & 0 & 1.080 \\
\end{pmatrix}.
\]

Another important class of matrices that do not require pivoting in Gaussian elimination is the set of diagonally dominant matrices.
The are two common types of diagonal dominance, *column* diagonal dominance and *row* diagonal dominance.

Column diagonal dominance is

\[ |a_{jj}| \geq \sum_{i \neq j} |a_{ij}| \]  

(4)

The absolute sum of off diagonal elements in a *column* is less than or equal to the diagonal element.

Row diagonal dominance is

\[ |a_{ii}| \geq \sum_{j \neq i} |a_{ij}| \]

The absolute sum of off diagonal elements in a *row* is less than or equal to the diagonal element.

For Gaussian elimination, both are equally good. Neither requires pivoting for stability. Diagonally dominant matrices arise in differential equations and in Markov chains.

Example 3 is both row and column diagonally dominant.

There is a significant overlap between diagonally dominant matrices and symmetric positive definite ones. (*I did not state this in class and you don’t have to know it.*)

**Theorem 1** If \( A \) is symmetric, diagonally dominant (row or column), has positive diagonal elements, and the inequality (4) is strict for at least one \( i \), then it is positive definite.