Gaussian Transformation and LU Factorization

To obtain a matrix factorization description of Gaussian elimination, it is easier to use a matrix description of the zeroing process.

For a given vector $v \in \mathbb{R}^n$ with $v_k \neq 0$ for some $1 \leq k \leq n$, let

$$\ell_{ik} = \frac{v_i}{v_k}, \quad i = k+1, \ldots, n, \qquad l_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \ell_{k+1,k} \\ \vdots \\ \ell_{n,k} \end{bmatrix},$$

and

$$M_k = I - l_k e_k^T = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & -\ell_{k+1,k} & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & -\ell_{n,k} & 0 & \cdots & 1 \end{bmatrix}$$

Then one can verify that

$$M_k v = \begin{bmatrix} v_1 \\ \vdots \\ v_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

That is, M_k zeros $v_{k+1}, \ldots, v_n$ and keeps $v_1, \ldots, v_k$ unchanged.

matrix M_k is called a Gaussian transformation, the vector l_k a Gauss vecto

The entries $l_{k+1,k}, \ldots, l_{n,k}$ are multipliers

Furthermore, one can verify that

$$M_k^{-1} = (I - l_k e_k^T)^{-1} = I + l_k e_k^T = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & \ell_{k+1,k} & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \ell_{n,k} & 0 & \cdots & 1 \end{bmatrix}$$

Gaussian Elimination and LU Factorization

Given a nonsingular matrix $A \in \mathbb{R}^{n\times n}$, Gauss transformations $M_1, M_2, \ldots, M_{n-1}$ can usually be found such that $M_{n-1}\cdots M_2 M_1 A \equiv U$ is upper triangular. To see this, we denote $A^{(1)} = [a_{ij}^{(1)}] = A$. If $a_{11}^{(1)} \neq 0$, then the Gauss transformation

$$M_1 = I - l_1 e_1^T, \quad \text{where} \quad l_1 = \begin{bmatrix} 0 \\ \ell_{21} \\ \vdots \\ \ell_{n1} \end{bmatrix}, \qquad \ell_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}, \quad i = 2, \ldots, n,$$

can be formed such that

$$A^{(2)} = M_1 A^{(1)} = \begin{bmatrix} a_{11}^{(2)} & a_{12}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{bmatrix},$$

where

$$a_{ij}^{(2)} = \begin{cases} a_{ij}^{(1)}, & \text{for } i = 1 \text{ and } j = 1, \ldots, n; \\ a_{ij}^{(1)} - \ell_{i1} * a_{1j}^{(1)}, & \text{for } i = 2, \ldots, n \text{ and } j = 2, \ldots, n. \end{cases}$$

The procedure can proceed to zero the entries in the second column below the diagonal, and so on. In general, at the k-th step, we are confronted with a matrix

$$\begin{aligned} A^{(k)} &= M_{k-1}\cdots M_2 M_1 A^{(1)} \\ &= \left[\begin{array}{cccc|ccc} a_{11}^{(k)} & a_{12}^{(k)} & \cdots & a_{1,k-1}^{(k)} & a_{1k}^{(k)} & \cdots & a_{1n}^{(k)} \\ 0 & a_{22}^{(k)} & \cdots & a_{2,k-1}^{(k)} & a_{2k}^{(k)} & \cdots & a_{2n}^{(k)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{k-1,k-1}^{(k)} & a_{k-1,k}^{(k)} & \cdots & a_{k-1,n}^{(k)} \\ \hline 0 & 0 & \cdots & 0 & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{kn}^{(k)} & \cdots & a_{nn}^{(k)} \end{array}\right] \\ &= \left[\begin{array}{cccc|ccc} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1,k-1}^{(1)} & a_{1k}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2,k-1}^{(2)} & a_{2k}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{k-1,k-1}^{(k)} & a_{k-1,k}^{(k)} & \cdots & a_{k-1,n}^{(k)} \\ \hline 0 & 0 & \cdots & 0 & a_{kk}^{(k)} & \cdots & a_{kn}^{(k)} \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{kn}^{(k)} & \cdots & a_{nn}^{(k)} \end{array}\right] \end{aligned}$$

If the pivot $a_{kk}^{(k)} \neq 0$, then the multipliers

$$\ell_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, \quad i = k+1, \ldots, n,$$

can be computed and the Gaussian transformation

$$M_k = I - l_k e_k^T, \quad \text{where} \quad l_k = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \ell_{k+1,k} \\ \vdots \\ \ell_{nk} \end{bmatrix},$$

can be applied to the left of $A^{(k)}$ to obtain

$$A^{(k+1)} = M_k A^{(k)} = \left[\begin{array}{cccc|cccc} a_{11}^{(k+1)} & a_{12}^{(k+1)} & \cdots & a_{1,k-1}^{(k+1)} & a_{1k}^{(k+1)} & \cdots & \cdots & a_{1n}^{(k+1)} \\ 0 & a_{22}^{(k+1)} & \cdots & a_{2,k-1}^{(k+1)} & a_{2k}^{(k+1)} & \cdots & \cdots & a_{2n}^{(k+1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \cdots & \vdots \\ 0 & 0 & \cdots & a_{k-1,k-1}^{(k+1)} & a_{k-1,k}^{(k+1)} & \cdots & \cdots & a_{k-1,n}^{(k+1)} \\ \hline 0 & 0 & \cdots & 0 & a_{kk}^{(k+1)} & \cdots & \cdots & a_{kn}^{(k+1)} \\ \vdots & \vdots & & \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \cdots & a_{nn}^{(k+1)} \end{array}\right],$$

$$A^{(k+1)} = M_k A^{(k)} = \left[\begin{array}{cccc|cccc} a_{11}^{(k+1)} & a_{12}^{(k+1)} & \cdots & a_{1,k-1}^{(k+1)} & a_{1k}^{(k+1)} & \cdots & \cdots & a_{1n}^{(k+1)} \\ 0 & a_{22}^{(k+1)} & \cdots & a_{2,k-1}^{(k+1)} & a_{2k}^{(k+1)} & \cdots & \cdots & a_{2n}^{(k+1)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \cdots & \vdots \\ 0 & 0 & \cdots & a_{k-1,k-1}^{(k+1)} & a_{k-1,k}^{(k+1)} & \cdots & \cdots & a_{k-1,n}^{(k+1)} \\ \hline 0 & 0 & \cdots & 0 & a_{kk}^{(k+1)} & \cdots & \cdots & a_{kn}^{(k+1)} \\ \vdots & \vdots & & \vdots & 0 & \ddots & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \cdots & a_{nn}^{(k+1)} \end{array}\right],$$

in which

$$a_{ij}^{(k+1)} = \begin{cases} a_{ij}^{(k)}, & \text{for } i = 1, \ldots, k, \; j = 1, \ldots, n; \\ 0, & \text{for } i = k+1, \ldots, n, \; j = k; \\ a_{ij}^{(k)} - \ell_{ik} a_{kj}^{(k)}, & \text{for } i = k+1, \ldots, n, \; j = k+1, \ldots, n. \end{cases}$$

Upon the completion,

$$U \equiv A^{(n)} = M_{n-1} \cdots M_2 M_1 A$$

is upper triangular.

Hence

$$A = M_1^{-1} M_2^{-1} \cdots M_{n-1}^{-1} U \equiv LU,$$

where

$$\begin{aligned} L \equiv M_1^{-1} M_2^{-1} \cdots M_{n-1}^{-1} &= (I - l_1 e_1^T)^{-1} (I - l_2 e_2^T)^{-1} \cdots (I - l_{n-1} e_{n-1}^T)^{-1} \\ &= (I + l_1 e_1^T)(I + l_2 e_2^T) \cdots (I + l_{n-1} e_{n-1}^T) \\ &= I + l_1 e_1^T + l_2 e_2^T + \cdots + l_{n-1} e_{n-1}^T \\ &= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \ell_{21} & 1 & 0 & \cdots & 0 \\ \ell_{31} & \ell_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdots & 1 \end{bmatrix} \end{aligned}$$

is unit lower triangular.

This matrix factorization is called the LU-factorization of A.

Alternatively, the algorithm can be derived directly by componentwise comparison in $A = LU$. Suppose that the factors L and U are written as

$$L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \ell_{21} & 1 & 0 & \cdots & 0 \\ \ell_{31} & \ell_{32} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n1} & \ell_{n2} & \ell_{n3} & \cdots & 1 \end{bmatrix}, \quad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix}.$$

By comparing the components in $A = LU$ as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ \ell_{21}u_{11} & \ell_{21}u_{12} + u_{22} & \ell_{21}u_{13} + \ell_{22}u_{23} & \cdots & l_{21}u_{1n} + u_{2n} \\ \ell_{31}u_{11} & \ell_{31}u_{12} + \ell_{32}u_{22} & \ell_{31}u_{13} + \ell_{32}u_{23} + u_{33} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n1}u_{11} & \ell_{n1}u_{12} + \ell_{n2}u_{2n} & \ell_{n1}u_{13} + \ell_{n2}u_{23} + \ell_{n3}u_{33} & \cdots & \ell_{n1}u_{1n} + \cdots + u_{nn} \end{bmatrix},$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}$$

$$= \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ \ell_{21}u_{11} & \ell_{21}u_{12}+u_{22} & \ell_{21}u_{13}+\ell_{22}u_{23} & \cdots & l_{21}u_{1n}+u_{2n} \\ \ell_{31}u_{11} & \ell_{31}u_{12}+\ell_{32}u_{22} & \ell_{31}u_{13}+\ell_{32}u_{23}+u_{33} & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \ell_{n1}u_{11} & \ell_{n1}u_{12}+\ell_{n2}u_{2n} & \ell_{n1}u_{13}+\ell_{n2}u_{23}+\ell_{n3}u_{33} & \cdots & \ell_{n1}u_{1n}+\cdots+u_{nn} \end{bmatrix},$$

we see that $u_{11} = a_{11}, u_{12} = a_{12}, \ldots, u_{1n} = a_{1n}$. Once u_{11} is determined, $\ell_{21} = a_{21}/u_{11}, \ell_{31} = a_{31}/u_{11}, \ldots, \ell_{n1} = a_{n1}/u_{11}$ can be computed. Next, $a_{22} = \ell_{21}u_{12} + u_{22}$. Since ℓ_{21} and u_{12} are known at this stage, $u_{22} = a_{22} - \ell_{21}u_{12}$ is computed. Once u_{22} is determined, $u_{22}, \ldots, u_{2n}$, and $\ell_{22}, \ldots, \ell_{n2}$ can be obtained. Then we proceed to compute the third row of U and the third column of L, and so on.

The principal of the procedure follows from the identity

$$a_{ij} = \sum_{p=1}^{\min(i,j)} \ell_{ip}u_{pj}.$$

Suppose that after $k-1$ steps we have computed the first $k-1$ columns of L and the first $k-1$ rows of U. At the k-th step, we can first determine u_{kk} since

$$a_{kk} = \sum_{p=1}^{k} \ell_{kp}u_{pk} = \sum_{p=1}^{k-1} \ell_{kp}u_{pk} + u_{kk},$$

and ℓ_{kp}, u_{pk}, $p = 1, 2, \ldots, k-1$, are known.

The computation is done by performing

$$u_{kk} = a_{kk} - \sum_{p=1}^{k-1} \ell_{kp}u_{pk}.$$

The computation is done by performing

$$u_{kk} = a_{kk} - \sum_{p=1}^{k-1} \ell_{kp} u_{pk}.$$

Once u_{kk} is determined and $u_{kk} \neq 0$, we can compute the k-th column of L and the k-th row of U. By using the identities

$$a_{ik} = \sum_{p=1}^{k-1} l_{ip} u_{pk} + l_{ik} u_{kk}, \qquad i = k+1, \ldots, n,$$

and

$$a_{kj} = \sum_{p=1}^{k-1} l_{kp} u_{pj} + u_{kj}, \qquad j = k+1, \ldots, n,$$

we can compute

Doolittle LU-factorization

$$\ell_{ik} = \left(a_{ik} - \sum_{p=1}^{k-1} \ell_{ip} u_{pk} \right) \Big/ u_{kk}, \qquad i = k+1, \ldots, n,$$

and

$$u_{kj} = a_{kj} - \sum_{p=1}^{k-1} \ell_{kp} u_{pj}, \qquad j = k+1, \ldots, n.$$

Algorithm **(Doolittle LU-Factorization)** *Given a nonsingular square matrix* $A \in \mathbb{R}^{n\times n}$*, this algorithm computes a unit lower triangular matrix* L *and an upper triangular matrix* U *such that* $A = LU$.

```
Initialize L = 0, U = 0
for k = 1,...,n do
   L(k,k) = 1
   t = 0
   for p = 1,...,k-1 do
      t = t + L(k,p) * U(p.k)
   end for
   U(k,k) = A(k,k) - t
   for i = k+1,...,n do
      t = 0
      for p = 1,...,k-1 do
         t = t + L(i,p) * U(p,k)
      end for
      L(i,k) = (A(i,k) - t)/U(k,k)
   end for
   for j = k+1,...,n do
      t = 0
      for p = 1,...,k-1 do
         t = t + L(k,p) * U(p,j)
      end for
      U(k,j) = A(k,j) - t
   end for
end for
```

Find an LU factorization of $A = \begin{bmatrix} 1 & 2 & 0 & 2 \\ 1 & 3 & 2 & 1 \\ 2 & 3 & 4 & 0 \end{bmatrix}$.

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 6 & -5 \end{bmatrix}$$

Solution

One way to find the LU factorization is to simply look for it directly. You need

$$\begin{bmatrix} 1 & 2 & 0 & 2 \\ 1 & 3 & 2 & 1 \\ 2 & 3 & 4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ x & 1 & 0 \\ y & z & 1 \end{bmatrix} \begin{bmatrix} a & d & h & j \\ 0 & b & e & i \\ 0 & 0 & c & f \end{bmatrix}.$$

Then multiplying these you get

$$\begin{bmatrix} a & d & h & j \\ xa & xd+b & xh+e & xj+i \\ ya & yd+zb & yh+ze+c & yj+iz+f \end{bmatrix}$$

and so you can now tell what the various quantities equal. From the first column, you need $a = 1, x = 1, y = 2$. Now go to the second column. You need $d = 2, xd + b = 3$ so $b = 1, yd + zb = 3$ so $z = -1$. From the third column, $h = 0, e = 2, c = 6$. Now from the fourth column, $j = 2, i = -1, f = -5$. Therefore, an LU factorization is

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 & 2 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 6 & -5 \end{bmatrix}.$$

You can check whether you got it right by simply multiplying these two.

For practical implementation, the storage of A can be overwritten by the entries of L and U. The upper triangular components of U overwrite upper triangular part of A, and the multipliers (the strictly lower triangular part of L) can be stored in the strictly lower triangualr part of A. The diagonal part of L needs not to be stored since they are all 1. An outer-product version of the LU-factorization is as follows.

Algorithm 3.7 (LU Factorization) *Given a nonsingular square matrix $A \in \mathbb{R}^{n\times n}$, this algorithm computes a unit lower triangular matrix L and an upper triangular matrix U such that $A = LU$. The matrix A is overwritten by L and U.*

```
for k = 1, ..., n − 1 do
  for i = k + 1, ..., n do
    A(i, k) = A(i, k)/A(k, k)
    for j = k + 1, ..., n do
      A(i, j) = A(i, j) − A(i, k) ⋆ A(k, j)
    end for
  end for
end for
```

This algorithm requires

$$\sum_{k=1}^{n-1}\sum_{i=k+1}^{n} 2(n-k) = \frac{2}{3}n^3 - \frac{1}{2}n^2 + \frac{1}{3}n$$

flops.

Existence and Uniqueness of LU Factorization

Definition 3.1 (Leading principal minor) *Let A be an $n \times n$ matrix. The upper left $k \times k$ submatrix, denoted as*

$$A_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix},$$

is called the leading $k \times k$ principal submtrix, and the determinant of A_k, $\det(A_k)$, is called the leading principal minor.

Theorem 3.1 *If all leading principal minor of $A \in \mathbb{R}^{n\times n}$ are nonzero, that is, all leading principal submatrices are nonsingular, then A has an LU-factorization.*

Theorem 3.1 *If all leading principal minor of $A \in \mathbb{R}^{n\times n}$ are nonzero, that is, all leading principal submatrices are nonsingular, then A has an LU-factorization.*

Proof: Proof by mathematical induction.

When $n = 1$, $A_1 = [a_{11}]$ is nonsingular, then $a_{11} \neq 0$.

Let $L_1 = [1]$ and $U_1 = [a_{11}]$. Then $A_1 = L_1U_1$. The theorem holds.

Assume that the leading principal submatrices $A_1, \ldots, A_k$ are nonsingular and A_k has an LU-factorization $A_k = L_kU_k$, where L_k is unit lower triangular and U_k is upper triangular. We shall show that there exist an unit lower triangular matrix L_{k+1} and an upper triangular matrix U_{k+1} such that $A_{k+1} = L_{k+1}U_{k+1}$.

Write

$$A_{k+1} = \begin{bmatrix} A_k & v_k \\ w_k^T & a_{k+1,k+1} \end{bmatrix}, \quad \text{where} \quad v_k = \begin{bmatrix} a_{1,k+1} \\ a_{2,k+1} \\ \vdots \\ a_{k,k+1} \end{bmatrix} \quad \text{and} \quad w_k = \begin{bmatrix} a_{k+1,1} \\ a_{k+1,2} \\ \vdots \\ a_{k+1,k} \end{bmatrix}.$$

Since A_k is nonsingular, both L_k and U_k are nonsingular..

Therefore the system $L_ky_k = v_k$ has a unique solution $y_k \in \mathbb{R}^k$,

and $z^tU_k = w_k^T$ has a unique solution $z_k \in \mathbb{R}^k$.

Let

$$L_{k+1} = \begin{bmatrix} L_k & 0 \\ z_k^T & 1 \end{bmatrix} \quad \text{and} \quad U_{k+1} = \begin{bmatrix} U_k & y_k \\ 0 & a_{k+1,k+1} - z_k^T y_k \end{bmatrix}.$$

Then L_{k+1} is unit lower triangular, U_{k+1} is upper triangular, and

$$L_{k+1}U_{k+1} = \begin{bmatrix} L_kU_k & L_ky_k \\ z_k^TU_k & z_k^Ty_k + a_{k+1,k+1} - z_k^Ty_k \end{bmatrix} = \begin{bmatrix} A_k & v_k \\ w_k^T & a_{k+1,k+1} \end{bmatrix} = A_{k+1}.$$

This proves the theorem.

Theorem 3.2 *If A is nonsingular and the LU factorization exists, then the LU factorization is unique and* $\det(A) = u_{11} \cdots u_{nn}$.

Proof: Suppose both $A = L_1U_1$ and $A = L_2U_2$ are LU factorizations. Since A is nonsingular, L_1, U_1, L_2, U_2 are all nonsingular, and

$$A = L_1U_1 = L_2U_2 \Longrightarrow L_2^{-1}L_1 = U_2U_1^{-1}.$$

Since L_1 and L_2 are unit lower triangular which implies $L_2^{-1}L_1$ is unit lower triangular, and U_1 and U_2 are upper triangular which implies $U_2U_1^{-1}$ is upper triangular, it follows that $L_2^{-1}L_1 = I = U_2U_1^{-1}$. Therefore $L_1 = L_2$ and $U_1 = U_2$. ■

Symmetric Positive Definite System and Cholesky Factorization

An $n \times n$ matrix A is positive definite if $x^TAx > 0$, for all $x \in \mathbb{R}^n$, $x \neq 0$.

If A is both

symmetric and positive definite (*spd*), then we can derive a stable LU factorization called the Choleseky factorization.

Lemma 3.1 *If $A \in \mathbb{R}^{n\times n}$ is positive definite, then A is nonsingular and $a_{ii} > 0$ for $i = 1, \ldots, n$.*

Proof: Suppose A is singular. Then there exists $x \in \mathbb{R}^n$ and $x \neq 0$ such that $Ax = 0$. This implies $x^TAx = 0$, which contradicts the fact that A is positive definite. Therefore A is nonsingular. It is also easy to see that

$$a_{ii} = e_i^TAe_i > 0,$$

where e_i is the i-th column of the $n \times n$ identify matrix. ■

Lemma 3.2 *If $A \in \mathbb{R}^{n\times n}$ is positive definite, then all leading principal submatrices of A are nonsingular.*

Proof: For $1 \le k \le n$, let $z_k = [x_1, \ldots, x_k]^T \in \mathbb{R}^k$ and $x = [x_1, \ldots, x_k, 0, \ldots, 0]^T \in \mathbb{R}^n$, where $x_1, \ldots, x_k \in \mathbb{R}$ are not all zero. Since A is positive definite,

$$z_k^T A_k z_k = x^T A x > 0,$$

where A_k is the $k \times k$ leading principal submatrix of A. This shows that A_k are also positive definite, hence A_k are nonsingular. ■

Theorem 3.3 *If $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite, then there exists a unique lower triangular matrix $G \in \mathbb{R}^{n\times n}$ with positive diagonal entries such that A has the factorization* $A = GG^T$.

Proof: From Lemma 3.2, all leading principal submatrices of a positive definite matrix are nonsingular, hence A has the LU factorization $A = LU$, where L is unit lower triangular and U is upper triangular. Since A is symmetric,

$$LU = A = A^T = U^T L^T \quad \Longrightarrow \quad U(L^T)^{-1} = L^{-1}U^T.$$

Note that $U(L^T)^{-1}$ is upper triangular and $L^{-1}U^T$ is lower triangular, this forces $U(L^T)^{-1}$ to be a diagonal matrix, say, $U(L^T)^{-1} = D$. Then $U = DL^T$. Hence

$$A = LDL^T.$$

Since A is positive definite,

$$x^T A x > 0 \quad \Longrightarrow \quad x^T LDL^T x = (L^T x)^T D (L^T x) > 0.$$

This means D is also positive definite, and hence $d_{ii} > 0$. Thus $D^{1/2}$ is well-defined and we have

$$A = LDL^T = LD^{1/2}D^{1/2}L^T \equiv GG^T,$$

where $G \equiv LD^{1/2}$. Since the LU factorization is unique, G is unique. ■

To derive an algorithm for computing the Cholesky factorization we assume that the first $k-1$ columns of G have been determined after $k-1$ steps. By componentwise comparison with equation one has

$$a_{kk} = \sum_{j=1}^{k} g_{kj}^2,$$

which gives

$$g_{kk}^2 = a_{kk} - \sum_{j=1}^{k-1} g_{kj}^2.$$

Moreover,

$$a_{ik} = \sum_{j=1}^{k} g_{ij} * g_{kj}, \qquad i = k+1, \ldots, n,$$

hence the k-th column of G can be computed by

$$g_{ik} = \left(a_{ik} - \sum_{j=1}^{k-1} g_{ij} * g_{kj} \right) \Big/ g_{kk}, \qquad i = k+1, \ldots, n.$$

Algorithm ° ¹° **(Cholesy Factorization)** *Given an $n \times n$ symmetric positive definite matrix A, this algorithm computes the Cholesky factorization $A = GG^T$.*

Initialize $G = 0$
for $k = 1, \ldots, n$ **do**
$\quad G(k,k) = \sqrt{A(k,k) - \sum_{j=1}^{k-1} G(k,j) * G(k,j)}$
$\quad$**for** $i = k+1, \ldots, n$ **do**
$\quad\quad G(i,k) = \left(A(i,k) - \sum_{j=1}^{k-1} G(i,j) * G(k,j) \right) \Big/ G(k,k)$
$\quad$**end for**
end for

In addition to n square root operations, there are approximately

$$\sum_{k=1}^{n} [2k - 1 + 2k(n-k)] = \frac{1}{3}n^3 + n^2 - \frac{1}{3}n$$

Diagonally Dominant Systems

A matrix $A \in \mathbb{R}^{n\times n}$ is said to be diagonally dominant if

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ij}|.$$

We shall show that Guassian elimination without pivoting can be safely used for this class of matrices.

Lemma 3.3 *If $A \in \mathbb{R}^{n\times n}$ is diagonally dominant, then A is nonsingular.*

Proof: Suppose A is singular. Then there exists $x \in \mathbb{R}^n$, $x \neq 0$ such that $Ax = 0$. Let k be the integer index such that

$$|x_k| = \max_{1\le i\le n} |x_i| \quad \Longrightarrow \quad \frac{|x_i|}{|x_k|} < 1, \quad \forall\, i \neq k.$$

Since $Ax = 0$, for the fixed k, we have

$$\sum_{j=1}^{n} a_{kj}x_j = 0 \quad \Longrightarrow \quad a_{kk}x_k = -\sum_{j=1, j\neq k}^{n} a_{kj}x_j \quad \Longrightarrow \quad |a_{kk}||x_k| \le \sum_{j=1, j\neq k}^{n} |a_{kj}||x_j|,$$

which implies

$$|a_{kk}| \le \sum_{j=1, j\neq k}^{n} |a_{kj}|\frac{|x_j|}{|x_k|} < \sum_{j=1, j\neq k}^{n} |a_{kj}|.$$

But this contradicts the assumption that A is diagonally dominant. Therefore A must be nonsingular. ∎