A short note on LU Decomposition

Suppose we have the system of equations of the form

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \quad \rightarrow \mathbf{A} = \mathbf{L} \cdot \mathbf{U} \tag{1}$$

where we are able to write the matrix $\mathbf{A}$ as a product of two matrices $\mathbf{L}$, the *lower triangular* having elements only on the diagonals and below, and $\mathbf{U}$, the *upper triangular* having elements only on the diagonal and above. For example, a 4×4 matrix $\mathbf{A}$ will look like

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \begin{pmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{pmatrix} \cdot \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix} \tag{2}$$

No matter how formidable (2) look like at first sight, this decomposition is extremely useful and not nearly as hard to deal with. If you do the matrix multiplication in RHS of (2), we get the following,

$$\begin{pmatrix} l_{11}u_{11} & l_{11}u_{12} & l_{11}u_{13} & l_{11}u_{14} \\ l_{21}u_{11} & l_{21}u_{12} + l_{22}u_{22} & l_{21}u_{13} + l_{22}u_{23} & l_{21}u_{14} + l_{22}u_{24} \\ l_{31}u_{11} & l_{31}u_{12} + l_{32}u_{22} & l_{31}u_{13} + l_{32}u_{23} + l_{33}u_{33} & l_{31}u_{14} + l_{32}u_{24} + l_{33}u_{34} \\ l_{41}u_{11} & l_{41}u_{12} + l_{42}u_{22} & l_{41}u_{13} + l_{42}u_{23} + l_{43}u_{33} & l_{41}u_{14} + l_{42}u_{24} + l_{43}u_{34} + l_{44}u_{44} \end{pmatrix} \tag{3}$$

To determine l_{ij} and u_{ij} we can equate each elements of the $\mathbf{LU}$ matrix with the corresponding elements of $\mathbf{A}$ and solve the equations, a couple of such equations are shown below in the equations (4 – 7).

$$\begin{aligned} l_{11}u_{11} &= a_{11} & l_{11}u_{12} &= a_{12} \quad \ldots & (4) \\ l_{21}u_{11} &= a_{21} & l_{21}u_{12} + l_{22}u_{22} &= a_{22} \quad \ldots & (5) \\ l_{31}u_{12} + l_{32}u_{22} &= a_{32} & l_{31}u_{13} + l_{32}u_{23} + l_{33}u_{33} &= a_{33} \quad \ldots & (6) \\ l_{41}u_{13} + l_{42}u_{23} + l_{43}u_{33} &= a_{43} & l_{41}u_{14} + l_{42}u_{24} + l_{43}u_{34} + l_{44}u_{44} &= a_{44} \quad \ldots & (7) \end{aligned}$$

Now, here is a catch! We have $4 \times 4 = 16$ equations but $4 \times (4 + 1) = 20$ variables, hence we cannot have a unique solutions for l_{ij} and u_{ij}.

The trick here is to put either all four $l_{ii} = 1$, known as **Crout decomposition**, or all four $u_{ij} = 1$, known as **Doolittle decomposition**.

Any of the above LU decompositions *i.e.* determination of the $\mathbf{L}$ and $\mathbf{U}$ can proceed iteratively. For Crout decomposition,

- Set $l_{ii} = 1$ for all $i = 1, \ldots, N$ and $u_{1j} = a_{1j}$

- For each $j = 1, 2, \ldots N$ do both the calculations below in the order they appear

$$u_{ij} = a_{ij} - \sum_{k=1}^{i-1} l_{ik} u_{kj} \qquad \text{for } \; i = 2, \ldots j \tag{8}$$

$$l_{ij} = \frac{1}{u_{jj}} \left(a_{ij} - \sum_{k=1}^{j-1} l_{ik} u_{kj} \right) \qquad \text{for } \; i = j+1, j+2, \ldots, N \tag{9}$$

An interesting point to note that every a_{ij} in (8 and 9) is used only once and never again. This means that the u_{ij} and l_{ij} can be stored in the location that corresponding a_{ij} is to occupy. Don't store the l_{ii} at all, just modify the evolution equations accordingly. For Crout decomposition this means,

$$\mathbf{LU} = \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ l_{21} & u_{22} & u_{23} & u_{24} \\ l_{31} & l_{32} & u_{33} & u_{34} \\ l_{41} & l_{42} & l_{43} & u_{44} \end{pmatrix} \quad \rightarrow \quad \mathbf{A} \tag{10}$$

But one point worth remembering – $\mathbf{A}$ can have LU decomposition only if all its *leading submatrices* have non-zero determinant. For instance, taking a 3×3 matrix,

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 2 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \;\Rightarrow\; \mathbf{A_1} = 1, \; \mathbf{A_2} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \text{ and } \mathbf{A_3} = \begin{pmatrix} 1 & 0 & 2 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \tag{11}$$

Since $\det\mathbf{A_1} = 1$, $\det\mathbf{A_2} = 1$ and $\det\mathbf{A_3} = -3$ are all non-zero, $\mathbf{A}$ has a guaranteed LU decomposition. But the matrix $\mathbf{B}$ below in (12) cannot be LU decomposed,

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 0 & 2 \\ -1 & 1 & 2 \end{pmatrix} \;\Rightarrow\; \mathbf{B_1} = 1, \; \mathbf{B_2} = \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix} \text{ and } \mathbf{B_3} = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 0 & 2 \\ -1 & 1 & 2 \end{pmatrix} \tag{12}$$

because although $\det\mathbf{B_1} = 1$ and $\det\mathbf{B_3} = 4$ are non-zero, the $\det\mathbf{B_2} = 0$. However, since the matrix $\mathbf{B}$ itself is invertible (as its determinant is non-zero), we can always perform partial pivoting to avoid the problem of zero determinant of one of the submatrices. But otherwise you are doomed!

Backward substitution

Before we get any further in LU decomposition technique, let us reflect on the $\mathbf{U}$ matrix for a few second. This matrix is the same one we got by Gaussian elimination previously of a matrix whose determinant we sought. This $\mathbf{U}$ matrix also offers us solutions of simultaneous equations by the method of *backward substitution.* Again, specializing to 4×4 matrix,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} \Rightarrow$$

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \\ \bar{b}_4 \end{pmatrix} \tag{13}$$

The solutions x_i are obvious if we start from the last row *i.e.* move *backwards*,

$$u_{44}x_4 = \bar{b}_4 \qquad x_4 = \frac{\bar{b}_4}{u_{44}} \tag{14}$$

$$u_{33}x_3 + u_{34}x_4 = \bar{b}_3 \qquad x_3 = \frac{\bar{b}_3 - u_{34}x_4}{u_{33}} = \frac{1}{u_{33}}\left(\bar{b}_3 - \frac{u_{34}}{u_{44}}\bar{b}_4\right) \tag{15}$$

$$u_{22}x_2 + u_{23}x_3 + u_{24}x_4 = \bar{b}_2 \qquad x_2 = \frac{\bar{b}_2 - u_{23}x_3 - u_{24}x_4}{u_{22}} \tag{16}$$

$$u_{11}x_1 + u_{12}x_2 + u_{13}x_3 + u_{14}x_4 = \bar{b}_1 \qquad x_1 = \frac{\bar{b}_1 - u_{12}x_2 - u_{13}x_3 - u14x_4}{u_{11}} \tag{17}$$

So generically speaking, for $N \times N$ matrix, backward solution is

$$x_i = \frac{1}{u_{ii}}\left(\bar{b}_i - \sum_{j=i+1}^{N} u_{ij}x_j\right), \quad \text{where } x_N = \frac{\bar{b}_N}{u_{NN}} \text{ and } i = N-1, N-2, \ldots, 1 \tag{18}$$

LU decomposition : forward-backward substitution

Let us now turn our attention to LU decomposition and use it to solve systems of equations. Importantly, it will also help us to obtain the determinant and inverse. However, irrespective of problem LU is applied to it is always advisable *to perform partial pivoting to make the method stable.* The procedure can be performed as follows,

1. Consider the equation (1),

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \quad \Rightarrow \quad \mathbf{L} \cdot \mathbf{U} \cdot \mathbf{x} = \mathbf{b} \tag{19}$$

2. Split up the above equation (19) in two equations

$$\mathbf{U} \cdot \mathbf{x} = \mathbf{y} \quad \Rightarrow \mathbf{L} \cdot \mathbf{y} = \mathbf{b} \tag{20}$$

3. In the first step solve for $\mathbf{y}$ from $\mathbf{L}\cdot\mathbf{y}=\mathbf{b}$ using *forward substitution* and then use it to solve for $\mathbf{x}$.

The *forward* solution proceed exactly the *backward* way except that here we use $\mathbf{L}$. Rewriting the $\mathbf{L}$ and $\mathbf{U}$ matrices again for ready reference,

$$\mathbf{L}=\begin{pmatrix} 1 & 0 & 0 & 0 \\ l_{21} & 1 & 0 & 0 \\ l_{31} & l_{32} & 1 & 0 \\ l_{41} & l_{42} & l_{43} & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{U}=\begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix} \tag{21}$$

The solutions for y_i and consequently x_i are thus

$$y_i=\bar{b}_i-\sum_{j=1}^{i-1} l_{ij}y_j, \qquad \text{where } y_1=\bar{b}_1 \text{ and } i=2,3,\ldots,N \tag{22}$$

$$x_i=\frac{1}{u_{ii}}\left(y_i-\sum_{j=1}^{N} u_{ij}x_j\right), \qquad \text{where } x_N=\frac{y_N}{u_{NN}} \text{ and } i=N-1,N-2,\ldots,1 \tag{23}$$

We get the determinant of $\mathbf{A}$ for free : $\det\mathbf{A}=\det\mathbf{LU}=\det\mathbf{L}\times\det\mathbf{U}=(-1)^n\prod_i u_{ii}$. For the inverse, use the combination of (22) and (23) to iterate through each column of the identity matrix $\mathbf{B}=\mathbb{1}$.

Cholesky decomposition

In linear algebra, whenever a matrix is written as a product of matrices it is called *decomposition* or *factorization*. So far we have studied the LU decomposition and how it can be used to solve linear algebra problems. Another useful factorization scheme is called **Cholesky decomposition**. It is a decomposition of a Hermitian, positive-definite matrix into a product of lower triangular $\mathbf{L}$ matrix and its conjugate transpose. Cholesky is about twice as efficient as the LU decomposition for solving system of linear equations. The decomposition is of the form,

$$\mathbf{A}=\mathbf{L}\,\mathbf{L}^{\dagger} \xrightarrow{\text{real matrix}} \mathbf{L}\,\mathbf{L}^{T} \Rightarrow \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22}a_{23} & \\ a_{31} & a_{32} & a_{33} \end{pmatrix}=\begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix}\begin{pmatrix} l_{11} & l_{12} & l_{13} \\ 0 & l_{22} & l_{23} \\ 0 & 0 & l_{33} \end{pmatrix} \tag{24}$$

When $\mathbf{A}$ is real and positive definite, then

$$l_{ii} = \pm\sqrt{a_{ii}-\sum_{j=1}^{i-1} l_{ij}^2} \tag{25}$$

$$l_{ij} = \frac{1}{l_{ii}}\left(a_{ij}-\sum_{k=1}^{i-1} l_{ik}l_{kj}\right) \quad \text{for } i<j \tag{26}$$

The expression under the square root in (25) is always positive and the sign before the square root is inconsequential. Figure out yourself the corresponding expressions for complex Hermitian matrix. An examples of Cholesky decomposition of a real, symmetric matrix is (taken from Wikipedia),

$$\begin{pmatrix} 4 & 12 & -16 \\ 12 & 37 & -43 \\ -16 & -43 & 98 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 6 & 1 & 0 \\ -8 & 5 & 3 \end{pmatrix} \begin{pmatrix} 2 & 6 & -8 \\ 0 & 1 & 5 \\ 0 & 0 & 3 \end{pmatrix} \tag{27}$$

Apart from being used for numerical solution of linear equations (1), Cholesky decomposition is also used in non-linear optimization for multiple variable, monte carlo simulation for decomposing covariance matrix, inversion of Hermitian matrices etc.