Solving Linear Systems with Triangular Matrices: Forward/Backward Substitution

Let $U = (u_{ij})_{ij} \in \mathbb{R}^{n \times n}$ be an upper triangular matrix, i.e., $u_{ij} = 0$ for $i > j$, so that U has the form

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & & \vdots \\ \vdots & & \ddots & u_{n-1,n} \\ 0 & \cdots & 0 & u_{nn} \end{pmatrix}.$$

1. Solve the following linear system by hand:

$$\begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

   Do you observe an iterative pattern?
   Hint: Start with the last equation/row.

2. Let $b \in \mathbb{R}^n$ be a given vector. Then derive a general (iterative) formula to obtain $x \in \mathbb{R}^n$, which solves

$$Ux = b.$$

   What requirements have to be put on the diagonal entries u_{ii} of U, so that the system has a unique solution?
   Hint: Write out the system as

$$\begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & & \vdots \\ \vdots & & \ddots & u_{n-1,n} \\ 0 & \cdots & 0 & u_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix}$$

   and start with the last row to determine x_n. Then use x_n to determine x_{n-1} from the second but last row. Continue this procedure until you reach the first row to determine x_1 with the help of the previously determined values $x_2, \ldots, x_n$.

3. Find a similar formula for lower triangular matrices $L = (\ell_{ij})_{ij} \in \mathbb{R}^{n \times n}$, for which $\ell_{ij} = 0$ for $i < j$.

Remark: A diagonal matrix is a special case of a triangular matrix.

Solution:

1.

$$\begin{array}{r} \text{(I)} \\ \text{(II)} \\ \text{(III)} \end{array} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{array}{rcl} \text{(III)} & \Rightarrow & \frac{1}{2}x_3 = 1 \quad \Rightarrow \quad x_3 = 2 \\ \text{(II)} & \Rightarrow & 2x_2 + x_3 = 0 \quad \Rightarrow \quad x_2 = -1 \\ \text{(I)} & \Rightarrow & x_1 - x_2 + x_3 = 3 \quad \Rightarrow \quad x_1 = 0 \end{array}$$

$$\Rightarrow \quad x = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$$

2. First note that a triangular matrix is invertible if and only if the diagonal entries are nonzero (see the formula below or note that $\det(U) = \prod_i u_{ii}$). Now, from considering the i-th row (equation)

$$u_{ii}x_i + u_{i,i+1}x_{i+1} + \cdots + u_{in}x_n = b_i,$$

we obtain a representation for x_i given by

$$\begin{aligned} x_n &= \frac{b_n}{x_n} \\ x_i &= \underbrace{\frac{1}{u_{ii}}}_{[\text{assume} \quad u_{ii} \neq 0]} \left(b_i - \sum_{j=i+1}^{n} u_{ij}x_j \right), \quad \text{for } i < n, \end{aligned}$$

where all the x_j for $j \in \{i+1, \dots, n\}$ in the sum are computed in previous steps.

3. In the same way, by simply changing the indices in the sum, we find a formula for lower triangular matrices given by

$$\begin{aligned} x_1 &= \frac{b_1}{x_1} \\ x_i &= \underbrace{\frac{1}{\ell_{ii}}}_{[\text{assume} \quad \ell_{ii} \neq 0]} \left(b_i - \sum_{j=1}^{i-1} \ell_{ij}x_j \right), \quad \text{for } i > 1, \end{aligned}$$

where the x_j for $j \in \{1, \dots, i-1\}$ in the sum are determined in previous steps.