Applied Linear Algebra in Data Analysis

Gaussian Elimination

Sivakumar Balasubramanian

Department of Bioengineering
Christian Medical College, Bagayam
Vellore 632002

Solving linear equations: Gaussian Elimination

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n &= b_1 : E_1 \\
a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n &= b_2 : E_2 \\
a_{31}x_1 + a_{32}x_2 \ldots + a_{3n}x_n &= b_3 : E_3 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 \ldots + a_{mn}x_n &= b_m : E_m
\end{aligned}
$$

- Gaussian elimination is a systematic way of simplifying the above equations to an equivalent system that can be easily solved.
- Three simple operations are repeatedly performed:
  - Interchanging of equations E_i and E_j.
  - Replacing equation E_i by αE_i, $\alpha \neq 0$.
  - Replacing equation E_j by $E_j + \alpha E_i$, $\alpha \neq 0$.
- These three operations do not change the solution of the given linear system.

Solving linear equations: Gaussian Elimination

$$\textbf{Augmented matrix}: \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

- We can work with the augmented matrix instead of the equations.
- Gaussian elimination is carried out on the entire matrix.
- The matrix is simplified to a point, from where one can easily:
  - find out the nature of the solutions for the system of equations; and
  - find the solution (with a bit of extra work), if they exist.

Solving linear equations: Gaussian Elimination

$$\left.\begin{array}{rcrcrcr} x_1 & + & 2x_2 & - & x_3 & = & 1 \\ 2x_1 & + & 3x_2 & + & 4x_3 & = & 4 \\ -2x_1 & - & 4x_2 & + & x_3 & = & -3 \end{array}\right\} \longrightarrow \left[\begin{array}{rrr|r} 1 & 2 & -1 & 1 \\ 2 & 3 & 4 & 4 \\ -2 & -4 & 1 & -3 \end{array}\right]$$

Gaussian Elimination

$$\left[\begin{array}{rrr|r} \underline{1} & 2 & -1 & 1 \\ 2 & 3 & 4 & 4 \\ -2 & -4 & 1 & -3 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} \underline{1} & 2 & -1 & 1 \\ 0 & -1 & 6 & 2 \\ -2 & -4 & 1 & -3 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} \underline{1} & 2 & -1 & 1 \\ 0 & \underline{-1} & 6 & 2 \\ 0 & 0 & \underline{-1} & -1 \end{array}\right]$$

Now, we can perform **back substitution** on this triangularized system of linear equations,

$$x_3 = 1; \quad x_2 = 4; \quad x_1 = -6$$

We can continue the simplification process through the **Gauss-Jordan** method.

Solving linear equations: Gauss-Jordan Method

Continue the elimination upwards until all elements, except the ones in the main diagonal, are zero.

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & -1 & 6 & 2 \\ 0 & 0 & -1 & -1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & -1 & 0 & -4 \\ 0 & 0 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -6 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 1 \end{array}\right] \Longrightarrow x_1 = -6; \quad x_2 = 4; \quad x_3 = 1;$$

Everything worked out well without any problems. What can go wrong here?

Try solving the these systems, $\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 2 & 3 & 4 & 4 \\ -2 & -4 & 2 & -3 \end{array}\right]$,

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 2 & 3 & 4 & 4 \\ -2 & -4 & 2 & -2 \end{array}\right]$$

What is the difference between these two systems?

Solving linear equations: Rectangular systems and Row Echelon Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n &= b_1 : E_1 \\ a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n &= b_2 : E_2 \\ a_{31}x_1 + a_{32}x_2 \ldots + a_{3n}x_n &= b_3 : E_3 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 \ldots + a_{mn}x_n &= b_m : E_m \end{aligned} \longrightarrow \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

Consider the following example,

$$\left[\begin{array}{ccccc|c} \underline{1} & -2 & 1 & 0 & 1 & 1 \\ 2 & -4 & 1 & -1 & -2 & 2 \\ -1 & 2 & 1 & 1 & 2 & -1 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} \underline{1} & -2 & 1 & 0 & 1 & 1 \\ 0 & 0 & \underline{-1} & -1 & -4 & 0 \\ 0 & 0 & 2 & 1 & 3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} \underline{1} & -2 & 1 & 0 & 1 & 1 \\ 0 & 0 & \underline{-1} & -1 & -4 & 0 \\ 0 & 0 & 0 & \underline{-1} & -5 & 0 \end{array}\right]$$

Solving linear equations: Rectangular systems and Row Echelon Form

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right] \longrightarrow \begin{bmatrix} \underline{\ast} & \ast & \ast & \ast & \ast & \ast & \ast \\ 0 & 0 & \underline{\ast} & \ast & \ast & \ast & \ast \\ 0 & 0 & 0 & \underline{\ast} & \ast & \ast & \ast \\ 0 & 0 & 0 & 0 & \underline{\ast} & \ast & \ast \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{\ast} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Things to notice about the echelon form:

- If a particualr row consists entirely of zeros, then all rows below that row also contain entirely of zeros.
- If the first non-zero entry in the i^{th} row occurs in the j^{th} position, then all elements below the i^{th} row are zero from columns 1 to j.

Columns containing pivot are called the ***basic columns***.

> **Rank of a matrix A** is defined at the number of basic columns in the row echelon form of the matrix **A**.

Solving linear equations: Reduced Row Echelon Form

$$\begin{bmatrix} \underline{*} & * & * & * & * & * & * \\ 0 & 0 & \underline{*} & * & * & * & * \\ 0 & 0 & 0 & \underline{*} & * & * & * \\ 0 & 0 & 0 & 0 & \underline{*} & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{*} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{\text{Gauss-Jordan}} \begin{bmatrix} \underline{1} & * & 0 & 0 & 0 & * & 0 \\ 0 & 0 & \underline{1} & 0 & 0 & * & 0 \\ 0 & 0 & 0 & \underline{1} & 0 & * & 0 \\ 0 & 0 & 0 & 0 & \underline{1} & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \underline{1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\left[\begin{array}{ccccc|c} \underline{1} & -2 & 1 & 0 & 1 & 1 \\ 0 & 0 & \underline{-1} & -1 & -4 & 0 \\ 0 & 0 & 0 & \underline{-1} & -5 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} \underline{1} & -2 & 0 & -1 & -3 & 1 \\ 0 & 0 & \underline{1} & 1 & 4 & 0 \\ 0 & 0 & 0 & \underline{-1} & -5 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} \underline{1} & -2 & 0 & 0 & 2 & 1 \\ 0 & 0 & \underline{1} & 0 & -1 & 0 \\ 0 & 0 & 0 & \underline{1} & 5 & 0 \end{array}\right]$$

- All non-basic columns can be represented as a linear combination of the basic columns.
- A non-basic columns is a linear combination of only the columns before it.
- Scaling factors for each basic comlumns is determined by the corresponding elements of the non-basic columns.

The reduced row echelon form reveals structure in the original matrix A.

Solving linear equations: Homogenous Systems

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n &= 0 \\ a_{31}x_1 + a_{32}x_2 \ldots + a_{3n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 \ldots + a_{mn}x_n &= 0 \end{aligned} \longrightarrow \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ a_{31} & a_{32} & \cdots & a_{3n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & 0 \end{array}\right]$$

Consider the following case,

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & 0 & 1 & 0 \\ 2 & -4 & 1 & -1 & -2 & 0 \\ -1 & 2 & 1 & 1 & 2 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rrrrr|r} 1 & -2 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 5 & 0 \end{array}\right]$$

$$\begin{aligned} x_1 - 2x_2 + 2x_5 &= 0 \\ x_3 - x_5 &= 0 \\ x_4 + 5x_5 &= 0 \end{aligned} \longrightarrow \begin{aligned} x_1 &= 2x_2 - 2x_5 \\ x_3 &= x_5 \\ x_4 &= -5x_5 \end{aligned} \longrightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 - 2x_5 \\ x_2 \\ x_5 \\ 5x_5 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 0 \\ 1 \\ -5 \\ 1 \end{bmatrix}$$

Solving linear equations: Homogenous Systems

- $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 0 \\ 1 \\ -5 \\ 1 \end{bmatrix}$ represents the general solution of the system of equations.

- In general, any system $[\mathbf{A} \mid \mathbf{0}]$ with $rank\,(\mathbf{A}) = r$ and $r < n$ has the general solution of the form,
$$\mathbf{x} = x_{f_1}\mathbf{h}_1 + x_{f_2}\mathbf{h}_2 + \ldots + x_{f_{n-r}}\mathbf{h}_{n-r}$$
where, the variables $x_{f_1}, x_{f_2}, \ldots, x_{f_{n-r}}$ are called the **free variables**.

- Free variables are the one corresponding to the non-basic columns; the variable variables corresponding to the basic columns are the **basic variables**.

- When does a homogenous system have a unique solution solution?
$\longrightarrow rank\,(\mathbf{A}) = n.$

Solving linear equations: Non-homogenous Systems

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 \ldots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 \ldots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 \ldots + a_{3n}x_n &= b_3 \quad \longrightarrow [\mathbf{A} \mid \mathbf{b}] \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 \ldots + a_{mn}x_n &= b_m
\end{aligned}$$

Consider the following case,

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & 0 & 1 & 1 \\ 2 & -4 & 1 & -1 & -2 & 2 \\ -1 & 2 & 1 & 1 & 2 & -1 \end{array}\right] \longrightarrow \left[\begin{array}{rrrrr|r} 1 & -2 & 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 5 & 0 \end{array}\right]$$

Solving linear equations: Non-homogenous Systems

$$\left[\begin{array}{rrrrr|r} 1 & -2 & 1 & 0 & 1 & 1 \\ 2 & -4 & 1 & -1 & -2 & 2 \\ -1 & 2 & 1 & 1 & 2 & -1 \end{array}\right] \longrightarrow \left[\begin{array}{rrrrr|r} 1 & -2 & 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 5 & 0 \end{array}\right]$$

$$\begin{aligned} x_1 - 2x_2 + 2x_5 &= 1 & \qquad x_1 &= 1 + 2x_2 - 2x_5 \\ x_3 - x_5 &= 0 \longrightarrow & x_3 &= x_5 \\ x_4 + 5x_5 &= 0 & x_4 &= -5x_5 \end{aligned}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 + 2x_2 - 2x_5 \\ x_2 \\ x_5 \\ 5x_5 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 0 \\ 1 \\ -5 \\ 1 \end{bmatrix}$$

The general solution of a non-homogenous system is sum of the particular solution and the general solution of the associated homogenous system.

Solving linear equations: Non-homogenous Systems

- The general solution for $[A \mid 0]$ with $rank\,(A) = r$,

$$\mathbf{x} = \mathbf{p} + x_{f_1}\mathbf{h}_1 + x_{f_2}\mathbf{h}_2 + \cdots + x_{f_{n-r}}\mathbf{h}_{n-r}$$

  where, $\mathbf{p}$ is the particular solution and $x_{f_1}, x_{f_2}, \ldots, x_{f_{n-r}}$ are the free variables.
- When do we have a unique solution to this system? $\longrightarrow$ $rank\,(\mathbf{A}) = n$.
- What about the case when there are no solutions? When does that happen? $\longrightarrow$ *When the system is not consistent.*

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ a_{31} & a_{32} & \cdots & a_{3n} & b_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right] \longrightarrow \left[\begin{array}{cccccc|c} 1 & * & 0 & 0 & 0 & * & c_1 \\ 0 & 0 & 1 & 0 & 0 & * & c_2 \\ 0 & 0 & 0 & 1 & 0 & * & c_3 \\ 0 & 0 & 0 & 0 & 1 & * & c_4 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & c_m \end{array}\right]$$

  There is a problem when $c_m \neq 0$
- The augmented matrix $[\mathbf{A} \mid \mathbf{b}]$ has the same number of basic columns as $\mathbf{A}$.
- $[\mathbf{A} \mid \mathbf{b}] \to [\mathbf{E} \mid \mathbf{c}]$: $\mathbf{c}$ is a non-basic column.
- $rank\,(\mathbf{A}) = rank\,([\mathbf{A} \mid \mathbf{b}])$

LU Factorization of a Matrix

- A major theme of matrix algebra is to decompose matrices into simpler components that provide insights into the nature of the matrix.

- A full rank square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ can be decomosed into the product of a lower triangular and an upper triangular matrix.

- Matrices associated with the three elementary operations:

Inter-changing rows 2 and 4	**Scaling row 2**	**Adding a multiple of row 2 to row 3**
$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \alpha & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

LU Factorization of a Matrix

- Consider the case: $\mathbf{A} = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 0 & 2 \\ 4 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & -2 \\ 0 & 0 & -1 \end{bmatrix} = \mathbf{LU}$

- **LU** factorization can be done only when no zero pivot is encountered during the Guassian elimination process.

- $\mathbf{Ax} = \mathbf{b}$ becomes $\mathbf{LUx} = \mathbf{b}$: This is decomposed into two triangular systems, $\mathbf{Ux} = \mathbf{y}, \quad \mathbf{Ly} = \mathbf{b}$. First solve $\mathbf{Ly} = \mathbf{b}$ and then solve $\mathbf{Ux} = \mathbf{y}$

- Properties:
  - Diagonal elements of **L** are 1, and **U** are not equal to zero.
  - **U** is the final result of Guassian elimination, and **L** is the matrix that reverses this process.
- Uses of the **LU** factorization:
  - Solving $\mathbf{Ax} = \mathbf{b}_i$ for several $\mathbf{b}_i$s. **LU** need to be calculated only once.
  - Factorization requires no extra space.

$\mathbf{PA} = \mathbf{LU}$ Factorization of a Matrix

- Consider the case: $\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \neq \mathbf{LU}$

- It turns out the second pivot become zero after the first elimination step, so **LU** factorization cannot be done on **A**.

- The following however fixes this issue,

$$\mathbf{PA} = \mathbf{LU}$$

  where, **P** is the permunation matrix, which is the elementary matrix for row exchanges.

- In the current example, the following allows matrix factorization.

$$\mathbf{PA} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{LU}$$