Computational Data Mining

Part 3: Linear Algebra

Linear Systems and Least Squares

Instructor: Mahdi Vasighi

Department of Computer Science and Information Technology,
Institute for Advanced Studies in Basic Sciences, Zanjan, Iran

Introduction to linear systems

we briefly review some facts about the solution of linear systems of equations,

$$
\begin{aligned}
a_{11}\boldsymbol{x_1} + a_{12}\boldsymbol{x_2} + a_{13}\boldsymbol{x_3} &= \boldsymbol{y_1} \\
a_{21}\boldsymbol{x_1} + a_{22}\boldsymbol{x_2} + a_{23}\boldsymbol{x_3} &= \boldsymbol{y_2} \\
a_{31}\boldsymbol{x_1} + a_{32}\boldsymbol{x_2} + a_{33}\boldsymbol{x_3} &= \boldsymbol{y_3}
\end{aligned}
$$

$$
\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}
\begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \\ \boldsymbol{x_3} \end{bmatrix}
=
\begin{bmatrix} \boldsymbol{y_1} \\ \boldsymbol{y_2} \\ \boldsymbol{y_3} \end{bmatrix}
$$

$$\mathrm{Ax} = \mathrm{y}$$

where $A \in R^{n\times n}$ is square and nonsingular.

Introduction to linear systems

Definition (**Triangular Matrices**)

An $n \times n$ matrix is said to be upper triangular if $a_{ij} = 0$ for $i > j$ and lower triangular if $a_{ij} = 0$ for $i < j$. Also A is said to be triangular if it is either upper triangular or lower triangular.

Example: $\begin{bmatrix} 1 & 0 & 0 \\ 7 & 3 & 0 \\ 2 & 4 & 3 \end{bmatrix}$ $\begin{bmatrix} 1 & 2 & 7 \\ 0 & 3 & 4 \\ 0 & 0 & 3 \end{bmatrix}$

Definition (**Diagonal Matrices**)

An $n \times n$ matrix is diagonal if $a_{ij} = 0$ whenever $i \neq j$.

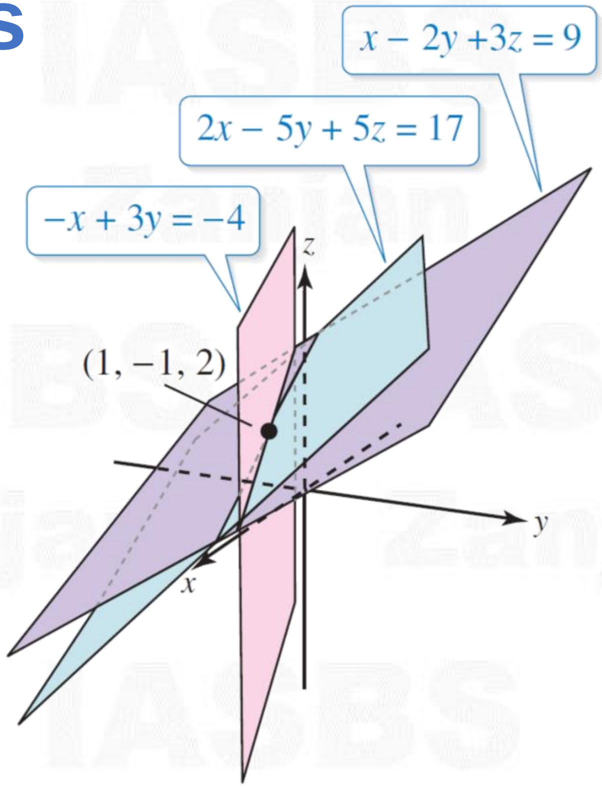
Example: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

Introduction to linear systems

$$\begin{aligned} x - 2y + 3z &= 9 \\ -x + 3y &= -4 \\ 2x - 5y + 5z &= 17 \end{aligned}$$

$$\begin{bmatrix} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{bmatrix}$$

Augmented matrix



These kind of linear systems can be solved using Gaussian elimination

Operations That Produce Equivalent Systems

Each of these operations on a system of linear equations produces an *equivalent* system.

1. Interchange two equations.
2. Multiply an equation by a nonzero constant.
3. Add a multiple of an equation to another equation.

Introduction to linear systems

Gaussian elimination

$$\begin{bmatrix} 1 & -2 & 3 & 9 \\ -1 & 3 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{bmatrix}$$

Adding the 1st row to the 2nd produces a new 2nd row.

$$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 2 & -5 & 5 & 17 \end{bmatrix}$$

Adding −2 times the 1st row to the 3rd row produces a new 3rd row

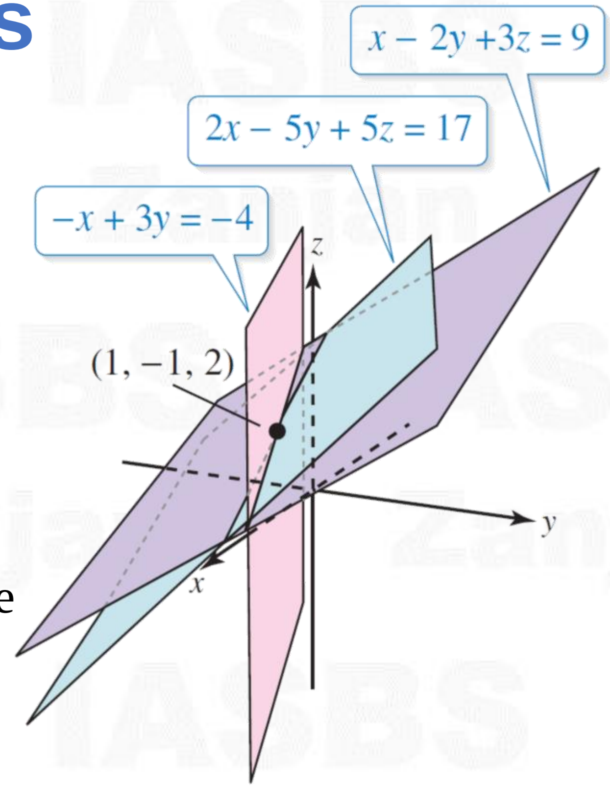
$$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & -1 & -1 & -1 \end{bmatrix}$$

Adding the 2nd equation to the 3rd row produces a new 3rd row.

$$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 2 & 4 \end{bmatrix}$$

Multiplying the 3rd row by ½ produces a new 3rd row.

$$\begin{bmatrix} 1 & -2 & 3 & 9 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$



$$\begin{aligned} x - 2y + 3z &= 9 \\ -x + 3y &= -4 \\ 2x - 5y + 5z &= 17 \end{aligned}$$

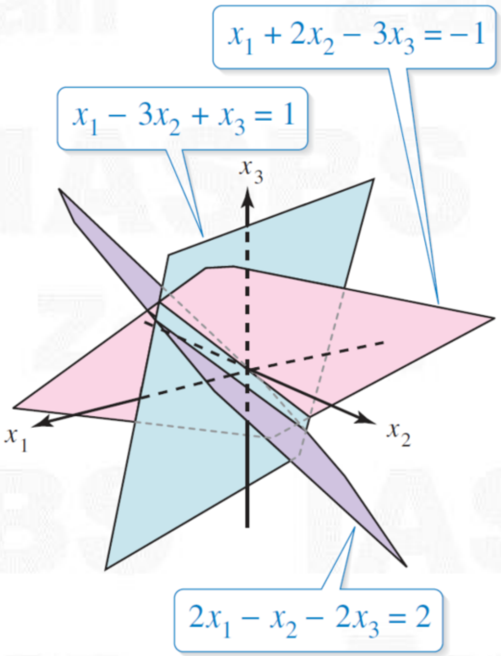
Gaussian Elimination with Partial Pivoting

Introduction to linear systems

A system of equations is considered overdetermined if there are more equations than unknowns:
The matrix $A \in R^{m \times n}$ is rectangular with $m > n$

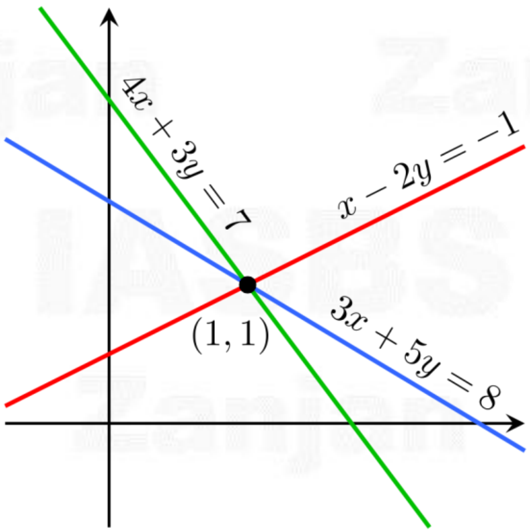
- A linear system is **inconsistent** if it has no solution otherwise it is said to be **consistent**

- The equations of a linear system are **independent** if none of the equations can be derived algebraically from the others.

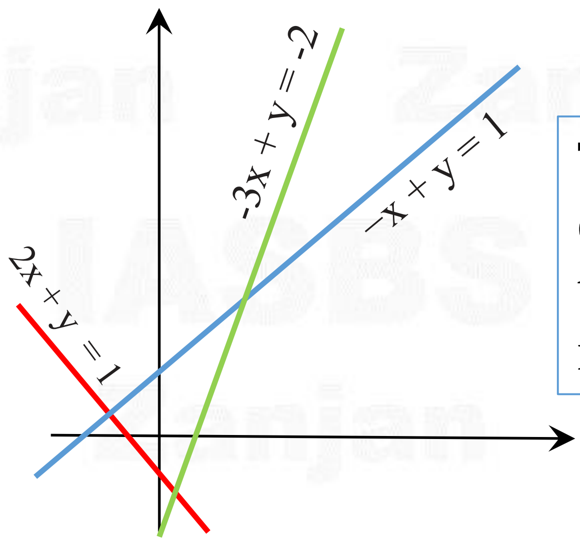


Introduction to linear systems

x − 2y = -1
3x + 5y = 8
4x + 3y = 7



− x + y = 1
2x + y = 1
-3x + y = -2



$$\begin{bmatrix} 1 & -2 & -1 \\ 3 & 5 & 8 \\ 4 & 3 & 7 \end{bmatrix} \quad \text{Rank} = 2$$

$$\begin{bmatrix} -1 & 1 & 1 \\ 2 & 1 & 1 \\ -3 & 1 & -2 \end{bmatrix} \quad \text{Rank} = 3$$

The number of independent equations in a system equals the rank of the augmented matrix of the system

If a system has more independent equations than unknowns, it is inconsistent and has no solutions.

Introduction to linear systems

Let $A \in R^{n\times n}$ and assume that A is nonsingular. Then for any right-hand-side b, the linear system $Ax = b$ has a unique solution.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} \boldsymbol{x_1} \\ \boldsymbol{x_2} \\ \boldsymbol{x_3} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

A square matrix $A \in R^{n\times n}$ with rank n is called nonsingular and has an inverse A^{-1} satisfying.

LU Decomposition

The observation of linear systems ($Ax=b$) involving triangular coefficient matrices are easier to deal with.

Definition (**LU factorization**)

- If the $n \times n$ matrix A can be written as the product of a lower triangular matrix L and an upper triangular matrix U, then we can write:

$$A = LU$$

where L is a lower triangular matrix and U is an upper triangular matrix.

LU Decomposition

Definition (**an Elementary Matrix**)
An $n \times n$ matrix is an elementary matrix when it can be obtained from the identity matrix I_n by a single elementary row operation.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & ½ & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$R_2 \leftrightarrow R_3$ $\qquad$ $(-2)R_1 + R_3 \rightarrow R_3$ $\qquad$ $½R_2 \rightarrow R_2$

Definition (**Row Equivalence**)
Let **A** and **B** be $m \times n$ matrices. Matrix **B** is **row-equivalent** to **A** when there exists a finite number of elementary matrices $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_k$ such that $\mathbf{B} = \mathbf{E}_k\mathbf{E}_{k-1} \ldots \mathbf{E}_2\mathbf{E}_1\mathbf{A}$.

LU Decomposition

$$A = LU$$

Begin by row reducing **A** to upper triangular form while keeping track of the elementary matrices used for each row operation.

Example:

$$\mathbf{A} \underset{\mathbf{E}_1}{\rightarrow} \underset{\mathbf{E}_2}{\rightarrow} \mathbf{U}$$

$$\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{U}$$

$$\mathbf{E}_2^{-1}\mathbf{E}_2\mathbf{E}_1\mathbf{A} = \mathbf{E}_2^{-1}\mathbf{U}$$

$$\mathbf{E}_1^{-1}\mathbf{E}_1\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{U}$$

$$\mathbf{A} = \mathbf{E}_1^{-1}\mathbf{E}_2^{-1}\mathbf{U}$$

$$\mathbf{A} = \mathbf{L}\mathbf{U}$$

L is the product of the inverses of the elementary matrices used in the row reduction.

LU Decomposition

Example:

$$\mathbf{A} = \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 2 & -10 & 2 \end{bmatrix} \quad E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \quad E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

$R_3 + (-2)R_1 \rightarrow R_3$

$$\begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 0 & -4 & 2 \end{bmatrix} \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \quad E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix}$$

$R_3 + (4)R_2 \rightarrow R_3$

$$\begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 14 \end{bmatrix} = \mathbf{U} \qquad E_1^{-1}E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -4 & 1 \end{bmatrix} = \mathbf{L}$$

LU Decomposition

Once an LU-factorization of a matrix A was obtained, the system of n linear equations in n variables **Ax = b** can be solve very efficiently in two steps:

1. let **y = Ux** and solve **Ly = b** for **y**.
2. Solve **Ux = y** for x.

Example:

$$x_1 - 3x_2 = -5$$
$$x_2 + 3x_3 = -1$$
$$2x_1 - 10x_2 + 2x_3 = -20$$

$$\mathbf{A} = \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 2 & -10 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -4 & 1 \end{bmatrix} \begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 14 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -3 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 14 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \\ -20 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} -5 \\ -1 \\ -14 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & -4 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \\ -14 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

LU Decomposition

Any Question?