Classic Ax=b problem

- A is nxn matrix
- X is a column vector
- b is a column vector nx1
- Example of simultaneous linear equations.

Classic Ax=b problem

- What if we can break into two matrices which are upper and lower triangular matrix.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

Classic Ax=b problem

- What if we can break into two matrices which are upper and lower triangular matrix.

$$\left(\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}\right) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Classic Ax=b problem

- What if we can break into two matrices which are upper and lower triangular matrix.

$$\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \left(\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \right) = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Classic Ax=b problem

- What if we can break into two matrices which are upper and lower triangular matrix.

$$\begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

$$\begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

Matrix Decomposition

- A matrix decomposition is way of reducing a matrix into its constituent parts
- It is to simply complex matrices into simple matrices to perform operations on decomposed matrices
- It is called matrix factorization

LU(Lower Upper Matrix Decomposition)

$$S = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix}$$

LU Decomposition – Example 1

$$\begin{matrix} 1 & 2 & 4 \\ 3 & 8 & 14 \\ 2 & 6 & 13 \end{matrix}$$

LU Decomposition – Example 2

$$\begin{matrix} 3 & 1 & 6 \\ -6 & 0 & -16 \\ 0 & 8 & -17 \end{matrix}$$

LU Decomposition – Example 2

$$\begin{matrix} 3 & 1 & 6 \\ -6 & 0 & -16 \\ 0 & 8 & -17 \end{matrix}$$

$$\begin{bmatrix} 3 & 1 & 6 \\ -6 & 0 & -16 \\ 0 & 8 & -17 \end{bmatrix} = \begin{bmatrix} U_{11} & U_{12} & U_{13} \\ L_{21}U_{11} & L_{21}U_{12} + U_{22} & L_{21}U_{13} + U_{23} \\ L_{31}U_{11} & L_{31}U_{12} + L_{32}U_{22} & L_{31}U_{13} + L_{32}U_{23} + U_{33} \end{bmatrix}$$

and comparing elements row by row we see that

$$\begin{array}{lll} U_{11} = 3, & U_{12} = 1, & U_{13} = 6, \\ L_{21} = -2, & U_{22} = 2, & U_{23} = -4 \\ L_{31} = 0 & L_{32} = 4 & U_{33} = -1 \end{array}$$

and it follows that

$$\begin{bmatrix} 3 & 1 & 6 \\ -6 & 0 & -16 \\ 0 & 8 & -17 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 6 \\ 0 & 2 & -4 \\ 0 & 0 & -1 \end{bmatrix}$$

is an LU decomposition of the given matrix.

Matrix Calculus

- Before we get into the zone of Calculus, let us refresh few terminologies:

- **Scalar:**
  - We refer to real numbers are scalars
  - For example, if Matrix Y = 3X (Here 3 is referred as Scalar)

Matrix Calculus

- **Scalar Matrix:**
- Scalar Matrix is a square matrix (n x n) such that
  - All off – diagonal elements of the matrix are 0
  - All on-diagonal elements of the matrix are equal

$$A = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}$$

Matrix Calculus

- **Trace of a Matrix:**

- Trace of a matrix (n x n) is sum of all the elements of the matrix
  - tr(X) = tr(x_{ij}) =$\sum_{i=1}^{n} x_{ii}$
  - X = $\begin{bmatrix} 2 & 3 & 1 \\ 5 & 4 & 5 \\ -2 & 1 & 6 \end{bmatrix}$
  - tr(X) = tr(x_{ij}) = (2 + 3 + 1 + 5 + 4 + 5 -2 + 1 + 6) = 25

Derivates of Vector Functions

- Let x, y are two vectors of the orders n and m respectively

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix},$$

- Say, each element of y_j is a function of x_i i.e. $y_j = f(x_i)$
- or simply put, $y = f(x)$

Derivates of Vector w.r.t. Vector

- Let x, y are two vectors of the orders n and m respectively then,
  - $\frac{\delta y}{\delta x}$ is a n x m matrix with each element of y differentiated with each element of x

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} \stackrel{\text{def}}{=} \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

Derivates of Vector w.r.t. Vector - Example

- Given y = $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, x = $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$
- And $y_1 = x_1^2 - x_3^2$ & $y_2 = 3x_1^2 + 2x_2 - x_3^3$
- $\frac{\delta y_1}{\delta x_1} = 2x_1$
- $\frac{\delta y_1}{\delta x_2} = 0$
- $\frac{\delta y_1}{\delta x_3} = -2x_3$

Derivates of Vector w.r.t. Vector - Example

- And $y_1 = x_1^2 – x_3^2$ & $y_2 = 3x_1^2 + 2x_2 – x_3^3$

- $\frac{\delta y_2}{\delta x_1} = 6x_1$

- $\frac{\delta y_2}{\delta x_2} = 2$

- $\frac{\delta y_2}{\delta x_3} = -3x_3^2$

Derivates of Vector w.r.t. Vector - Example

- Given y = $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$, x = $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$
- And $y_1 = x_1^2 - x_3^2$ & $y_2 = 3x_1^2 + 2x_2 - x_3^3$
- $\frac{\delta y}{\delta x} = \begin{bmatrix} 2x_1 & 6x_1 \\ 0 & 2 \\ -2x_3 & -3x_3^2 \end{bmatrix}$

Derivates of Vector Functions

- If A is an m x n matrix, then f(x) = Ax

Such that f(x) = Ax = ($\sum_j (a_{1j} * x_j)$, $\sum_j (a_{2j} * x_j)$, …$\sum_j (a_{mj} * x_j)$)’

$$\frac{\partial f}{\partial \boldsymbol{x}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \boldsymbol{A}.$$

Derivates of Vector w.r.t. Vector - Example

- Given A = $\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}$, x = $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

- And f = Ax = $\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} * \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ = $\begin{bmatrix} 2x_1 + 3x_2 \\ 4x_1 + 5x_2 \end{bmatrix}$

Derivates of Vector w.r.t. Vector - Example

- Here f = $\begin{bmatrix} 2x_1 + 3x_2 \\ 4x_1 + 5x_2 \end{bmatrix}$, x = $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
- $\frac{\delta f_1}{\delta x_1} = 2; \ \frac{\delta f_1}{\delta x_2} = 3$
- $\frac{\delta f_1}{\delta x_1} = 4; \ \frac{\delta f_2}{\delta x_2} = 5$
- $\boldsymbol{\frac{\delta f}{\delta x}}$ **=** $\begin{bmatrix} \mathbf{2} & \mathbf{3} \\ \mathbf{4} & \mathbf{5} \end{bmatrix}$ **= A**

Derivates of Matrix w.r.t. Scalar

- If the elements of (m x n) matrix A are functions of a scalar,
  - $A = (a_{ij}) = (a_{ij}(x))$, i.e. a_{ij} – a function of x
  - Derivate of A is simply the derivate of x wrt to individual elements of A

Derivates of Matrix w.r.t. Scalar - Example

- If the elements of (m x n) matrix A are functions of a scalar,

- A = $\begin{bmatrix} x^3 & x^2 & 2x \\ 5x^2 & -x & 4x^5 \\ -2x^5 & -x^2 & 2x^3 \end{bmatrix}$, Find $\frac{\delta A}{\delta x}$?

Derivates of Matrix w.r.t. Scalar - Example

$$A = \begin{bmatrix} x^3 & x^2 & 2x \\ 5x^2 & -x & 4x^5 \\ -2x^5 & -x^2 & 2x^3 \end{bmatrix}$$

- $\frac{\delta A}{\delta x} = \begin{bmatrix} \frac{\delta x^3}{\delta x} & \frac{\delta x^2}{\delta x} & \frac{\delta 2x}{\delta x} \\ \frac{\delta 5x^2}{\delta x} & \frac{\delta -x}{\delta x} & \frac{\delta 4x^5}{\delta x} \\ \frac{\delta -2x^5}{\delta x} & \frac{\delta -x^2}{\delta x} & \frac{\delta 2x^3}{\delta x} \end{bmatrix} = \begin{bmatrix} 3x^2 & 2x & 2 \\ 10x & -1 & 20x^4 \\ -10x^4 & -2x & 6x^2 \end{bmatrix}$