BITS Pilani
Pilani Campus

BITS Pilani presentation

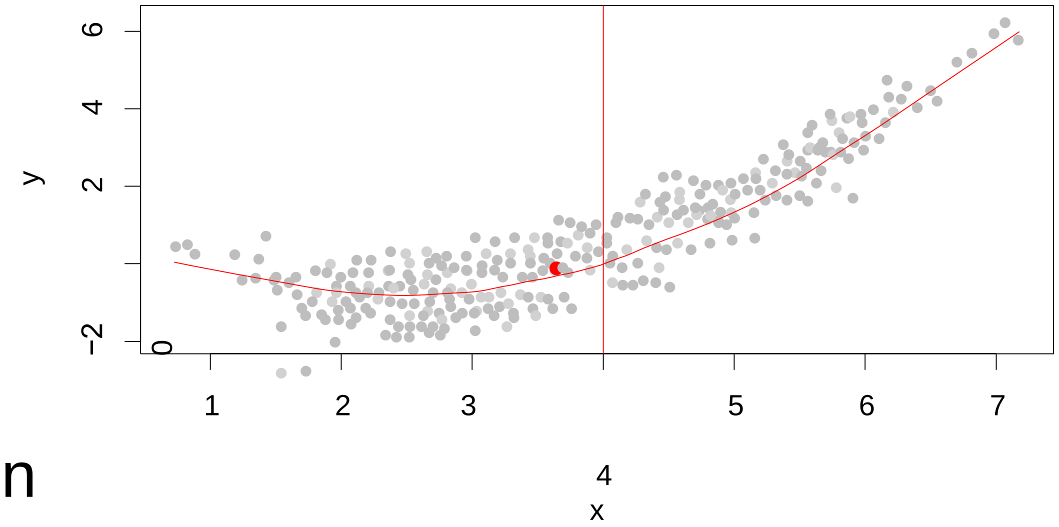
innovate achieve lead

BITS Pilani
Pilani Campus

SS ZC416 Mathematical Foundations for Data Science
Linear Algebra

Motivation
Why Learn Matrices?

Motivation – Linear Regression



Consider Linear Regression

We want to predict Y for a given value of x

Is there an ideal f (X)?

- In particular, what is a good value for f (X) at any selected value of X, say X = 4?
  There can be many Y values at X = 4
  A good value is f (4) = E(Y |X = 4), the expected value of Y given X = 4.

This ideal f (x) = E(Y |X = x) is called the regression function

The Linear Model

$$f(x) = E(Y \mid X = x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p$$

We collect data

#	X1	X2	X3	Y
1	230.1	37.8	69.2	22.1
2	44.5	39.3	45.1	10.4
3	17.2	45.9	69.3	9.3
4	151.5	41.3	58.5	18.5
5	180.8	10.8	58.4	12.9
6	8.7	48.9	75	7.2

Given this set of observations, we estimates the β's: $\hat{Y} = \hat{f} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_p X_p$

If $X^T X$ is non-singular, then the system of normal equations has a unique solution:

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T Y$$

Matrix Operations

Matrix & Matrix Notations

- A matrix is a rectangular array of numbers enclose in brackets.
- These numbers are called entries or elements of the matrix.
- In a matrix horizontal lines are called rows and vertical lines are called columns.
- If a matrix has m rows and n columns, the dimension is said to be $m \times n$
- The matrix may be denoted by A_{mxn}

Vectors

- Matrix having just a single row or column are called vectors. Its entries are called components of the vectors.
- Matrix having just one row is called row vector
- Matrix having just one column is called a column vector
- In this course, when we mention a vector, we will assume it is a column vector unless stated otherwise

Square Matrix

m = n

Matrix Operations

- Matrix identity

  $[A] = [B]$ if and only if $a_{ij} = b_{ij}$ for all i and j

- Matrix Addition and Subtraction

  $[C] = [A] + [B] \quad \Rightarrow \quad C_{ij} = A_{ij} + B_{ij}$

  $[C] = [A] - [B] \quad \Rightarrow \quad C_{ij} = A_{ij} - B_{ij}$

Examples

$A=\begin{bmatrix}1 & 0 & 4\\3 & 1 & 2\end{bmatrix}$ A + B? A – B? C + D? D + C? C + (D + E)? (C + D) + E

$B = \begin{bmatrix}1 & 2\\4 & 5\\7 & 8\end{bmatrix}$

$C = \begin{bmatrix}1 & -1 & 0\\-2 & 1 & 2\\0 & 1 & -2\end{bmatrix}$

$D = \begin{bmatrix}0 & 1 & 0\\-1 & 2 & 2\\0 & 2 & -1\end{bmatrix}$

$E=\begin{bmatrix}0 & -2 & 1\\1 & 1 & -1\\2 & 2 & -2\end{bmatrix}$

Properties

- Commutative
  [A] + [B] = [B] + [A]
  [A] - [B] = -[B] + [A]

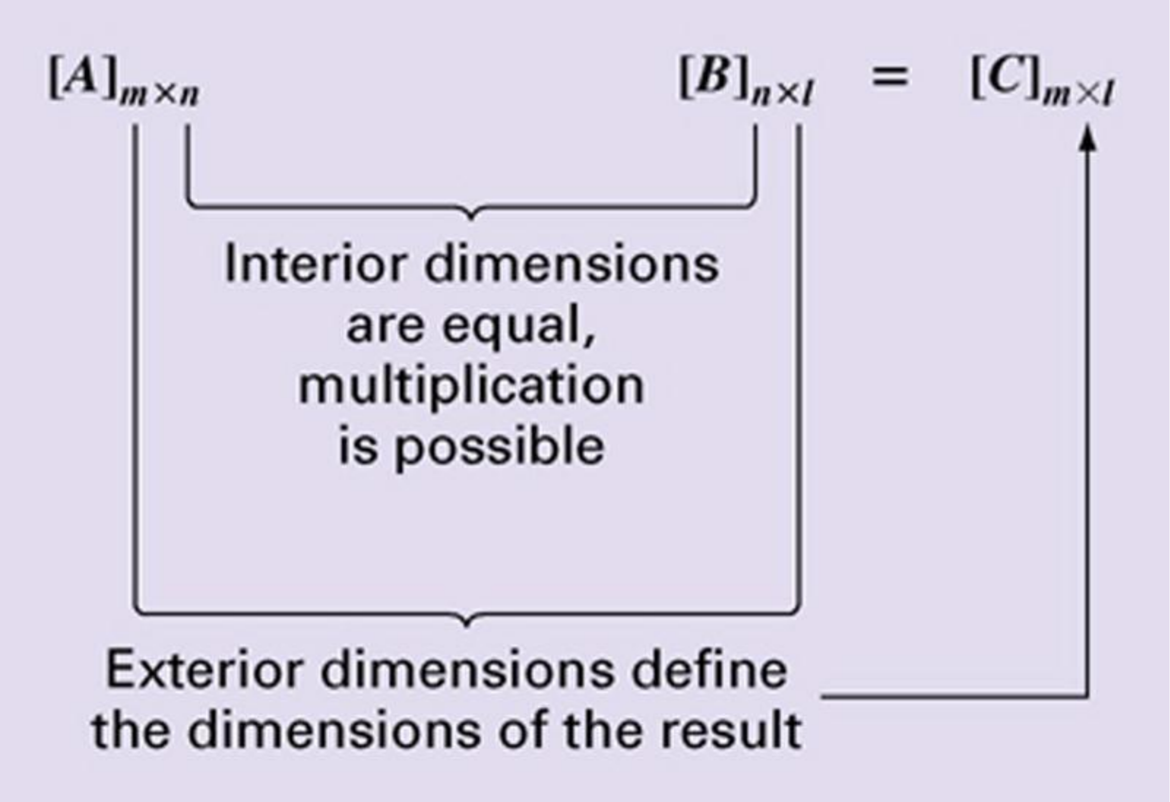
- Associative
  ([A] + [B]) + [C] = [A] + ([B] + [C])
  ([A] + [B]) - [C] = [A] + ([B] - [C])
  ([A] - [B]) + [C] = [A] + (-[B] + [C])

Multiplication of Matrix by a Scalar

A=$\begin{bmatrix} 1 & 0 & 4 \\ 3 & 1 & 2 \end{bmatrix}$ v=$\begin{bmatrix} 2 & 3 & -1 \end{bmatrix}$ c = 2, d = -1

Matrix Multiplication

Matrix multiplication can be performed only if the inner dimensions are equal



Product of Two Matrices

The **product C** = **AB** (in this order) of an $m \times n$ matrix **A** = $[a_{jk}]$ times an $r \times p$ matrix **B** = $[b_{jk}]$ is defined if and only if $r = n$ and is then the $m \times p$ matrix **C** = $[c_{jk}]$ with entries

(1)
$$c_{jk} = \sum_{l=1}^{n} a_{jl} b_{lk} = a_{j1}b_{1k} + a_{j2}b_{2k} + \cdots + a_{jn}b_{nk} \qquad \begin{array}{l} j = 1, \cdots, m \\ k = 1, \cdots, p. \end{array}$$

Examples

$A=\begin{bmatrix}1 & 0 & 4\\3 & 1 & 2\end{bmatrix}$

$B = \begin{bmatrix}1 & 2\\4 & 5\\7 & 8\end{bmatrix}$

$C = \begin{bmatrix}1 & -1 & 0\\-2 & 1 & 2\\0 & 1 & -2\end{bmatrix}$

$D = \begin{bmatrix}0 & 1 & 0\\-1 & 2 & 2\\0 & 2 & -1\end{bmatrix}$

$E=\begin{bmatrix}0 & -2 & 1\\1 & 1 & -1\\2 & 2 & -2\end{bmatrix}$

$A \times B$? $C \times D$? $B \times A$? $D \times C$?

Suppose $v = \begin{bmatrix}1\\-1\\0\end{bmatrix}$ What are the possible multiplications with v?

Matrix Multiplication Properties

- Associative

$$([A][B])[C] = [A]([B][C])$$

- Distributive

$$[A] \times ([B] + [C]) = [A] \times [B] + [A] \times [C]$$

$$([A] + [B]) \times [C] = [A] \times [C] + [B] \times [C]$$

- Not generally commutative $[A][B] \neq [B][A]$

Matrix Multiplication

- In general [A] [B] ≠ [B] [A]
- AB = 0, does not necessarily imply A=0 or B=0 or BA=0.
- AC = AD does not necessarily imply C=D

C = $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ & D = $\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ & A = $\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$

$C \times D$, $D \times C$, $A \times C$, $A \times D$

Special Matrices

- Diagonal matrix
- Identity matrix

$$I_{3x3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \qquad A = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = 2\ I_{3x3}$$

Lower and Upper Triangular Matrices

Upper Triangular Matrix: C = $\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & -2 \end{bmatrix}$

Lower Triangular Matrix = $\begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & 0 \\ 0 & 2 & -1 \end{bmatrix}$

Transpose & Trace

$$B = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & -1 & 0 \\ -2 & 1 & 2 \\ 0 & 1 & -2 \end{bmatrix}$$

Trace(C) = Sum of the diagonal entries = 1 + 1 + (-2) = 0

$$B^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \end{bmatrix} \qquad C^T = \begin{bmatrix} 1 & -2 & 0 \\ -1 & 1 & 1 \\ 0 & 2 & -2 \end{bmatrix}$$

Symmetric and Skew-Symmetric Matrices

$$C = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 2 \\ 0 & 2 & -2 \end{bmatrix}$$

$$D = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 2 & 2 \\ 0 & -2 & -1 \end{bmatrix}$$

Symmetric: $C^T = C$

Skew-Symmetric: $D^T = -D$

innovate achieve lead

BITS Pilani
Pilani Campus

Thank you!!