



Machine Learning

Linear Algebra
review (optional)

Matrices and
vectors

Matrix: Rectangular array of numbers:



$$\rightarrow \boxed{\mathbb{R}^{4 \times 2}}$$



$$\boxed{\mathbb{R}^{2 \times 3}}$$

Dimension of matrix: number of rows x number of columns

Matrix Elements (entries of matrix)

$$A = \begin{bmatrix} 1402 & 191 \\ 1371 & 821 \\ 949 & 1437 \\ 147 & 1448 \end{bmatrix}$$

A_{ij} = " i, j entry" in the i^{th} row, j^{th} column.

$$A_{11} = 1402$$

$$A_{12} = 191$$

$$A_{32} = 1437$$

$$A_{41} = 147$$

~~A_{43}~~ = Undefined (error)

Vector: An $n \times 1$ matrix.

$$\textcircled{y} = \begin{bmatrix} \textcircled{460} \\ \textcircled{232} \\ \textcircled{315} \\ 178 \end{bmatrix}$$

$n = 4$

← 4-dimensional vector.

~~$\mathbb{R}^{3 \times 2}$~~

\mathbb{R}^4

$y_i = i^{th}$ element

$$y_1 = 460$$

$$y_2 = 232$$

$$y_3 = 315$$

→ A, B, C, X

a, b, x, y

1-indexed vs 0-indexed:

$y[1]$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$$

1-indexed

$y[0]$

$$y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

0-indexed



Machine Learning

Linear Algebra review (optional)

Addition and scalar multiplication

Matrix Addition

$$\begin{array}{c} \downarrow \quad \downarrow \\ \rightarrow \begin{bmatrix} \textcircled{1} & 0 \\ \textcircled{2} & 5 \\ \textcircled{3} & 1 \end{bmatrix} + \begin{bmatrix} \textcircled{4} & 0.5 \\ \textcircled{2} & 5 \\ \textcircled{0} & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0.5 \\ 4 & 10 \\ 3 & 2 \end{bmatrix} \\ \text{3} \times \text{2} \quad \text{3} \times \text{2} \quad \text{3} \times \text{2} \end{array}$$

$$\begin{array}{c} \rightarrow \begin{bmatrix} 1 & 0 \\ 2 & 5 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 0.5 \\ 2 & 5 \end{bmatrix} = \\ \text{3} \times \text{2} \quad \text{2} \times \text{2} \end{array}$$

error

Scalar Multiplication

← real number

$$3 \times \begin{bmatrix} 1 & 0 \\ 2 & 5 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 6 & 15 \\ 9 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 5 \\ 3 & 1 \end{bmatrix} \times 3$$

3x2 3x2

$$\begin{bmatrix} 4 & 0 \\ 6 & 3 \end{bmatrix} / 4 = \frac{1}{4} \begin{bmatrix} 4 & 0 \\ 6 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{3}{2} & \frac{3}{4} \end{bmatrix}$$

Combination of Operands

Scalar multiplication

$$3 \times \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} / 3$$

Scalar division

matrix subtraction / vector subtraction

matrix addition / vector addition

$$= \begin{bmatrix} 3 \\ 12 \\ 6 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ \frac{2}{3} \end{bmatrix}$$
$$= \begin{bmatrix} 2 \\ 12 \\ 10\frac{1}{3} \end{bmatrix}$$

3x1 matrix
3-dimensional vector



Machine Learning

Linear Algebra review (optional)

Matrix-vector multiplication

Example

$$\begin{bmatrix} 1 & 3 \\ 4 & 0 \\ 2 & 1 \end{bmatrix}_{3 \times 2} \begin{bmatrix} 1 \\ 5 \end{bmatrix}_{2 \times 1} = \begin{bmatrix} 16 \\ 4 \\ 7 \end{bmatrix}_{3 \times 1} \text{ matrix}$$

$$1 \times 1 + 3 \times 5 = 16$$

$$4 \times 1 + 0 \times 5 = 4$$

$$2 \times 1 + 1 \times 5 = 7$$

Details:

$$\underline{A} \times \underline{x} = \underline{y}$$

\underline{A} is an $m \times n$ matrix (m rows, n columns).
 \underline{x} is an $n \times 1$ matrix (n-dimensional vector).
 \underline{y} is an m -dimensional vector.

→ To get \underline{y}_i , multiply \underline{A} 's i^{th} row with elements of vector \underline{x} , and add them up.

Example

$$\begin{bmatrix} 1 & 2 & 1 & 5 \\ 0 & 3 & 0 & 4 \\ -1 & -2 & 0 & 0 \end{bmatrix}_{3 \times 4} \begin{matrix} \downarrow \\ \begin{bmatrix} 1 \\ 3 \\ 2 \\ 1 \end{bmatrix}_{4 \times 1} \end{matrix} = \begin{bmatrix} 14 \\ 13 \\ -7 \end{bmatrix}_{3 \times 1} = \begin{bmatrix} 14 \\ 13 \\ -7 \end{bmatrix}$$

$$\left. \begin{array}{l} 1 \times 1 + 2 \times 3 + 1 \times 2 + 5 \times 1 = 14 \\ 0 \times 1 + 3 \times 3 + 0 \times 2 + 4 \times 1 = 13 \\ -1 \times 1 + (-2) \times 3 + 0 \times 2 + 0 \times 1 = -7 \end{array} \right\}$$

House sizes:

- 2104
- 1416
- 1534
- 852

Matrix

4x2

1	2104
1	1416
1	1534
1	852

X

$h_0(x)$

2x1

Vector

-40
0.25

=

4x1 matrix

$-40 \times 1 + 0.25 \times 2104$
$-40 \times 1 + 0.25 \times 1416$

$h_0(1416)$

Prediction = Data Matrix \otimes Parameters

4x1

for $i = 1:1000$,
prediction(i) = ...

$$h_{\theta}(x) = -40 + 0.25x$$



Machine Learning

Linear Algebra review (optional)

Matrix-matrix multiplication

Example

$$\begin{array}{l} \begin{bmatrix} 1 & 3 & 2 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 0 & 1 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} 11 & 10 \\ 9 & 14 \end{bmatrix} \\ \textcircled{2 \times 3} \quad \textcircled{3 \times 2} \\ \begin{bmatrix} 1 & 3 & 2 \\ 4 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 11 \\ 9 \end{bmatrix} \\ \begin{bmatrix} 1 & 3 & 2 \\ 4 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 10 \\ 14 \end{bmatrix} \end{array}$$

Handwritten green annotations show the calculation of the 2x2 result matrix from the two 2x1 products:

- The first column of the result is $\begin{bmatrix} 11 \\ 9 \end{bmatrix}$, which is the sum of the first column of the first product ($\begin{bmatrix} 11 \\ 9 \end{bmatrix}$) and the first column of the second product ($\begin{bmatrix} 10 \\ 14 \end{bmatrix}$).
- The second column of the result is $\begin{bmatrix} 10 \\ 14 \end{bmatrix}$, which is the sum of the second column of the first product ($\begin{bmatrix} 11 \\ 9 \end{bmatrix}$) and the second column of the second product ($\begin{bmatrix} 10 \\ 14 \end{bmatrix}$).

Details:

$$\underline{A} \times \underline{B} = \underline{C}$$

$m \times n$ matrix
(m rows,
 n columns)

$n \times o$ matrix
(n rows,
 o columns)

$m \times o$
matrix

The i^{th} column of the matrix C is obtained by multiplying A with the i^{th} column of B . (for $i = 1, 2, \dots, o$)

Example

$$\overset{2 \times 2}{\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}} \overset{2 \times 2}{\begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix}} =$$

$$\overset{2 \times 2}{\begin{bmatrix} 9 & 7 \\ 15 & 12 \end{bmatrix}}$$

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

$$= \begin{bmatrix} 1 \times 0 + 3 \times 3 \\ 2 \times 0 + 5 \times 3 \end{bmatrix} = \begin{bmatrix} 9 \\ 15 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 \times 1 + 3 \times 2 \\ 2 \times 1 + 5 \times 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 12 \end{bmatrix}$$

House sizes:

$$\begin{Bmatrix} \frac{2104}{1416} \\ \frac{1534}{852} \end{Bmatrix}$$

Matrix

$$\begin{bmatrix} 1 & \frac{2104}{1416} \\ 1 & \frac{1534}{852} \\ 1 & \frac{1534}{852} \\ 1 & \frac{1534}{852} \end{bmatrix} \times$$

Matrix

$$\begin{bmatrix} -40 \\ 0.25 \end{bmatrix} \begin{bmatrix} 200 \\ 0.1 \end{bmatrix} \begin{bmatrix} -150 \\ 0.4 \end{bmatrix} =$$

$$\begin{bmatrix} 486 \\ 314 \\ 344 \\ 173 \end{bmatrix} \begin{bmatrix} 410 \\ 342 \\ 353 \\ 285 \end{bmatrix} \begin{bmatrix} 692 \\ 416 \\ 464 \\ 191 \end{bmatrix}$$

Prediction
of first
 h_θ

Predictions
of 2nd
 h_θ

Have 3 competing hypotheses:

1. $h_\theta(x) = -40 + 0.25x$

2. $h_\theta(x) = 200 + 0.1x$

3. $h_\theta(x) = -150 + 0.4x$



Machine Learning

Linear Algebra review (optional)

Matrix multiplication properties

$$3 \times 5 = 5 \times 3$$


"Commutative"

Let A and B be matrices. Then in general,
 $A \times B \neq B \times A$. (not commutative.)

E.g.

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


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


\neq

$$\begin{array}{l} A \times B \\ m \times n \quad \times \quad n \times m \\ \hline A \times B \text{ is } m \times m \\ B \times A \text{ is } n \times n \end{array}$$

$$\underline{3 \times 5 \times 2}$$

$$3 \times 10 = 30 = 15 \times 2$$

$$3 \times (5 \times 2) = (3 \times 5) \times 2$$

"Associative"

$$\begin{array}{l} A \times (B \times C) \leftarrow \\ \underline{(A \times B)} \times C \leftarrow \end{array}$$

$$A \times B \times C.$$

Let $D = B \times C$. Compute $A \times D$.

Let $E = A \times B$. Compute $E \times C$.

$A \times (B \times C)$
 $(A \times B) \times C$
 Some
 answer.

1 is identity

$$1 \times z = z \times 1 = z$$

for any z

Identity Matrix

Denoted I (or $I_{n \times n}$).

Examples of identity matrices:

$[1]$
 1×1

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

2×2

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3×3

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

4×4

Informally:

$$\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}$$

For any matrix A ,

$$A \cdot I = I \cdot A = A$$

$m \times n$ $n \times n$ $m \times m$ $m \times n$ $m \times n$

$I_{n \times n}$

Note:

$AB \neq BA$ in general

$$AI = IA \checkmark$$



Machine Learning

Linear Algebra
review (optional)

Inverse and
transpose

$$\underline{1 = \text{"identity"}}$$

$$3 \underbrace{(3^{-1})}_{\frac{1}{3}} = 1$$

$$12 \times \underbrace{(12^{-1})}_{\frac{1}{12}} = 1$$

$$0 \underbrace{(0^{-1})}_{\text{undefined}}$$

Not all numbers have an inverse.

Matrix inverse: \swarrow square matrix
(#rows = #columns) A^{-1}

If A is an $m \times m$ matrix, and if it has an inverse,

$$\rightarrow \underline{A(A^{-1})} = \underline{A^{-1}A} = \underline{I}.$$

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

e.g. $\underbrace{\begin{bmatrix} 3 & 4 \\ 2 & 16 \end{bmatrix}}_A \underbrace{\begin{bmatrix} 0.4 & -0.1 \\ -0.05 & 0.075 \end{bmatrix}}_{A^{-1}} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{A^{-1}A} = I_{2 \times 2}$

Matrices that don't have an inverse are "singular" or "degenerate"

Matrix Transpose

Example:

$$\underline{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 5 & 9 \end{bmatrix}_{2 \times 3}$$
$$\underline{B} = \underline{A}^T = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 0 & 9 \end{bmatrix}_{3 \times 2}$$

Let A be an $m \times n$ matrix, and let $B = A^T$.

Then B is an $n \times m$ matrix, and

$$\underline{B}_{ij} = \underline{A}_{ji}.$$

$$B_{12} = A_{21} = 2$$

$$B_{32} = 9 \quad A_{23} = 9.$$