Linear Algebra review (optional)

Matrices and vectors

Machine Learning

Matrix: Rectangular array of numbers:

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

4×2 matrix

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

$$2 \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

3

2×3 matrix

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

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 ith row, j^{th} column.

$A_{11} = 1402$

$A_{12} = 191$

$A_{32} = 1437$

$A_{41} = 147$

$A_{43} =$ undefined (error)

Vector: An n x 1 matrix.

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

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

n = 4

← 4-dimensional vector. $\mathbb{R}^4$

$y_i = i^{th}$ element

$y_1 = 460$

$y_2 = 232$

$y_3 = 315$

A, B, C, X

a, b, x, y

1-indexed vs 0-indexed:

$y[i]$

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

1-indexed

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

$y[0]$

0-indexed

Andrew Ng

Linear Algebra review (optional)

Addition and scalar multiplication

Machine Learning

Matrix Addition

$$\begin{bmatrix} 1 & 0 \\ 2 & 5 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 0.5 \\ 2 & 5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0.5 \\ 4 & 10 \\ 3 & 2 \end{bmatrix}$$

3 × 2
matrix

3 × 2

3 × 2

$$\begin{bmatrix} 1 & 0 \\ 2 & 5 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 0.5 \\ 2 & 5 \end{bmatrix} = \text{error}$$

3 × 2

2 × 2

Scalar Multiplication

real number

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

3×2

3×2

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

Combination of Operands

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

Scalar multiplication

Scalar division

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

matrix subtraction / vector subtraction

matrix addition / vector addition

$$= \begin{bmatrix} 2 \\ 12 \\ 10\frac{1}{3} \end{bmatrix}$$

3x1 matrix

3-dimensional vector

Andrew Ng

Linear Algebra review (optional)

Matrix-vector multiplication

Machine Learning

Example

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

$3 \times 2 \qquad 2 \times 1 \qquad\qquad 3 \times 1 \text{ matrix}$

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

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

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

Details:

$$A \quad \times \quad x \quad = \quad y$$

m x n matrix
(m rows,
n columns)

n x 1 matrix
(n-dimensional
vector)

m-dimensional
vector

To get y_i, multiply A's i^{th} row with elements of vector x, and add them up.

Example

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

3×4 4×1 3×1

$1 \times 1 + 2 \times 3 + 1 \times 2 + 5 \times 1 = 14$

$0 \times 1 + 3 \times 3 + 0 \times 2 + 4 \times 1 = 13$

$-1 \times 1 + (-2) \times 3 + 0 \times 2 + 0 \times 1 = -7$

House sizes:

→ 2104
→ 1416
→ 1534
→ 852

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

$h_\theta(x)$

4×2

Matrix

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

2×1

Vector

$$\times \begin{bmatrix} -40 \\ 0.25 \end{bmatrix} =$$

$h_\theta(2104)$

4×1 matrix

$$\begin{bmatrix} -40 \times 1 + 0.25 \times 2104 \\ -40 \times 1 + 0.25 \times 1416 \\ \\ \end{bmatrix}$$

$h_\theta(1416)$

prediction = DataMatrix ⊗ Parameters

4×1

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

```
1    % Initialize a 3 by 2 matrix
2    A = [1, 2; 3, 4;5, 6]
3
4    % Initialize a 2 by 1 matrix
5    B = [1; 2]
6
7    % We expect a resulting matrix of (3 by 2)*(2 by 1) = (3 by 1)
8    mult_AB = A*B
9
10   % Make sure you understand why we got that result
```

ear Algebra
iew (optional)

Matrix-matrix
multiplication

Machine Learning

Example

$$\begin{bmatrix} 1 & 3 & 2 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 11 & 10 \\ 9 & 14 \end{bmatrix}$$

2×3 3×2 2×2

$$\begin{bmatrix} 1 & 3 & 2 \\ 4 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 11 \\ 9 \end{bmatrix}$$

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

Details:

$$A \qquad \times \qquad B \qquad = \qquad C$$

m x n matrix
(m rows,
n columns)

n x o matrix
(n rows,
o columns)

$n \times 1$

m x o
matrix

The ith column of the matrix C is obtained by multiplying A with the ith column of B. (for i = 1,2,...,o)

Andrew Ng

Example

$$\underset{\underset{\textstyle 2\times 2}{}}{\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix}} \underset{\underset{\textstyle 2\times 2}{}}{\begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix}} = \underset{\underset{\textstyle 2\times 2}{}}{\begin{bmatrix} 9 & 7 \\ 15 & 12 \end{bmatrix}}$$

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

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

House sizes:

$$\left\{ \begin{array}{c} \underline{2104} \\ \underline{1416} \\ \underline{1534} \\ \underline{852} \end{array} \right.$$

Have 3 competing hypotheses:

1. $h_\theta(x) = -40 + 0.25x$
2. $h_\theta(x) = 200 + 0.1x$
3. $h_\theta(x) = -150 + 0.4x$

Matrix

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

$\times$

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 hθ

Predictions of 2nd hθ

Linear Algebra review (optional)

Matrix multiplication properties

```
% Initialize random matrices A and B
A = [1,2;4,5]
B = [1,1;0,2]

% Initialize a 2 by 2 identity matrix (Diagonal matrix)
I = eye(2)

% The above notation is the same as I = [1,0;0,1]

% What happens when we multiply I*A ?
IA = I*A

% How about A*I ?
AI = A*I

% Compute A*B
AB = A*B

% Is it equal to B*A?
BA = B*A

% Note that IA = AI but AB != BA
```

Machine Learning

$$3 \times 5 = 5 \times 3 \qquad \text{``Commutative''}$$

Let A and B be matrices. Then in general,

$$A \times B \neq B \times A. \text{ (not commutative.)}$$

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

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

$\neq$

$$A \times B$$
$$m \times n \qquad n \times m$$

$A \times B$ is $m \times m$

$B \times A$ is $n \times n$

$3 \times 5 \times 2$

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

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

"Associative"

$A \times (B \times C)$ ⇐

$(A \times B) \times C$ ⇐

$A \times B \times C.$

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.

Identity Matrix

1 is identity

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

for any z

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

Examples of identity matrices:

Informally:

$[1]$
1×1

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

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
3 x 3

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
4 x 4

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

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

$AI = IA$ ✓

Linear Algebra review (optional)

Inverse and transpose

```
% Initialize matrix A
A = [1,2,0;0,5,6;7,0,9]

% Transpose A
A_trans = A'

% Take the inverse of A
A_inv = inv(A)
% Inverse of A (In matlab, inv(A); In Octave, pinv(A))

% What is A^(-1)*A?
A_invA = inv(A)*A
```

Machine Learning

$1 = $ "identity"

$3 \boxed{(3^{-1})} = 1$

$\quad\quad\quad \frac{1}{3}$

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

$\quad\quad\quad \frac{1}{12}$

$0 \, (0^{-1})$ — undefined

Not all numbers have an inverse.

Matrix inverse:

square matrix
(#rows = #columns)

A^{-1}

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

If A is an m x m matrix, and if it has an inverse,

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

e.g. $\begin{bmatrix} 3 & 4 \\ 2 & 16 \end{bmatrix}$ $\begin{bmatrix} 0.4 & -0.1 \\ -0.05 & 0.075 \end{bmatrix}$ $= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_{2\times2}$

$\quad\quad\quad\quad\quad 2\times2 \quad\quad\quad\quad\quad A^{-1} \quad\quad\quad\quad\quad\quad\quad A^{-1}A$

$\quad\quad\quad\quad\quad\quad A$

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

Matrix Transpose

Example:

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

2×3

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

3×2

Let A be an m x n matrix, and let $B = A^T$.
Then B is an n x m matrix, and

$$B_{ij} = A_{ji}.$$

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

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