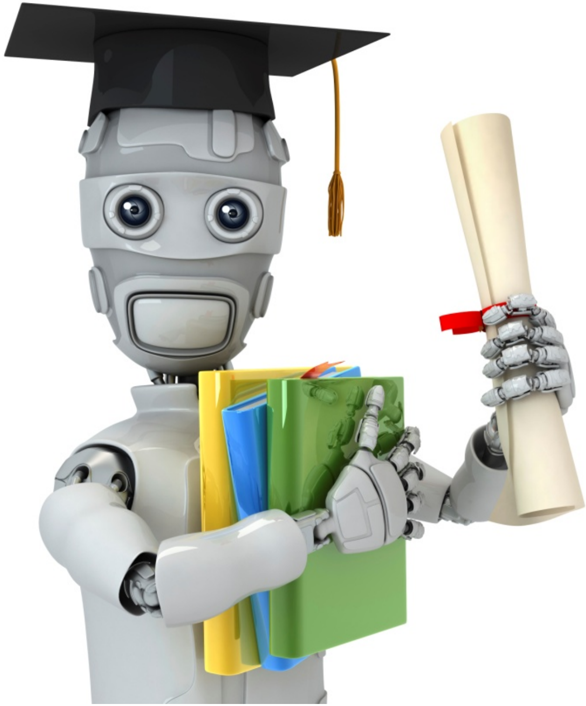


Machine Learning

Linear Algebra review (optional)

Matrices and vectors

Matrix: Rectangular array of numbers:

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

4×2 matrix

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

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

$$3$$

2×3 matrix

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

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.

$A_{11} = 1402$

$A_{12} = 191$

$A_{32} = 1437$

$A_{41} = 147$

$\cancel{A_{43}}$ = undefined (error)

Andrew Ng

Vector: An n x 1 matrix.

n= 4

$$y = \begin{bmatrix} 460 \\ 232 \\ 315 \\ 178 \end{bmatrix} \leftarrow \text{4-dimensional vector.} \quad \mathbb{R}^4$$

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

$y_i = i^{th}$ element

$y_1 = 460$

$y_2 = 232$

$y_3 = 315$

$\boxed{A, B, C, X}$ $\boxed{a, b, x, y}$

1-indexed vs 0-indexed:

y[1]

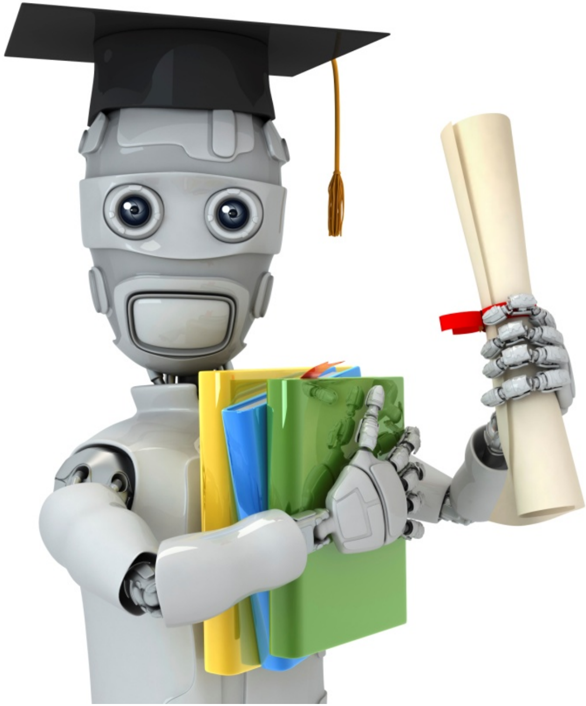
$$y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$$

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

$$\underbrace{\begin{bmatrix} 1 & 0 \\ 2 & 5 \\ 3 & 1 \end{bmatrix}}_{\substack{3\times 2 \\ \text{matrix}}} + \underbrace{\begin{bmatrix} 4 & 0.5 \\ 2 & 5 \\ 0 & 1 \end{bmatrix}}_{3\times 2} = \underbrace{\begin{bmatrix} 5 & 0.5 \\ 4 & 10 \\ 3 & 2 \end{bmatrix}}_{3\times 2}$$

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

Scalar Multiplication

real number

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

3×2 $\quad$ 3×2

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

Combination of Operands

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

Scalar multiplication

Scalar division

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

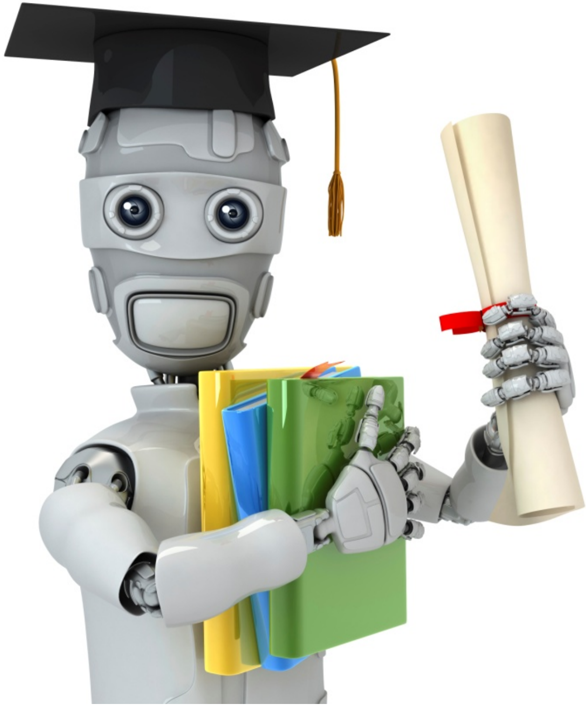
matrix subtraction / vector subtraction

matrix addition / vector addition

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

3x1 matrix

3-dimensional vector



Machine Learning

Linear Algebra review (optional)

Matrix-vector multiplication

Example

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

3×1 matrix

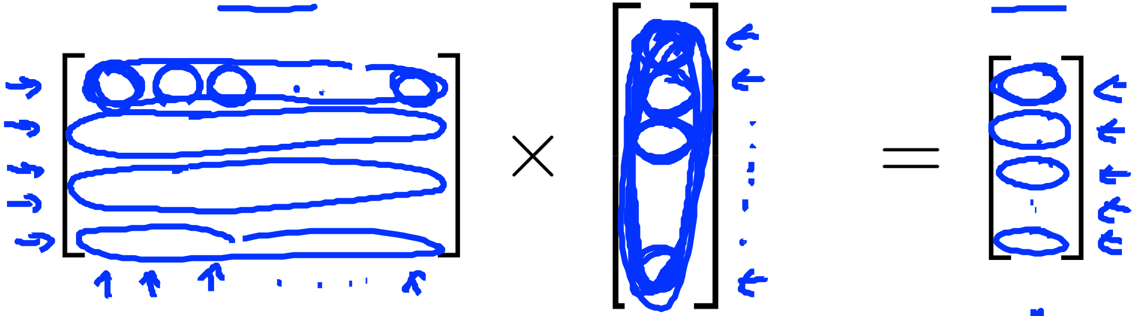
$1\times 1 + 3\times 5 = 16$

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

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

Details:

$$A \times x = y$$



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.

Andrew Ng

Example

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

$\boxed{3} \times 4$ 4×1 3×1

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

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

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

Andrew Ng

House sizes:

→ 2104

→ 1416

→ 1534

→ 852

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

$h_\theta(x)$

4x2

Matrix

2x1

Vector

4x1 matrix

$h_\theta(2104)$

$$\begin{bmatrix} 1 & 2104 \\ 1 & 1416 \\ 1 & 1534 \\ 1 & 852 \end{bmatrix} \times \begin{bmatrix} -40 \\ 0.25 \end{bmatrix} = \begin{bmatrix} -40 \times 1 + 0.25 \times 2104 \\ -40 \times 1 + 0.25 \times 1416 \\ \\ \\ \end{bmatrix}$$

$h_\theta(1416)$

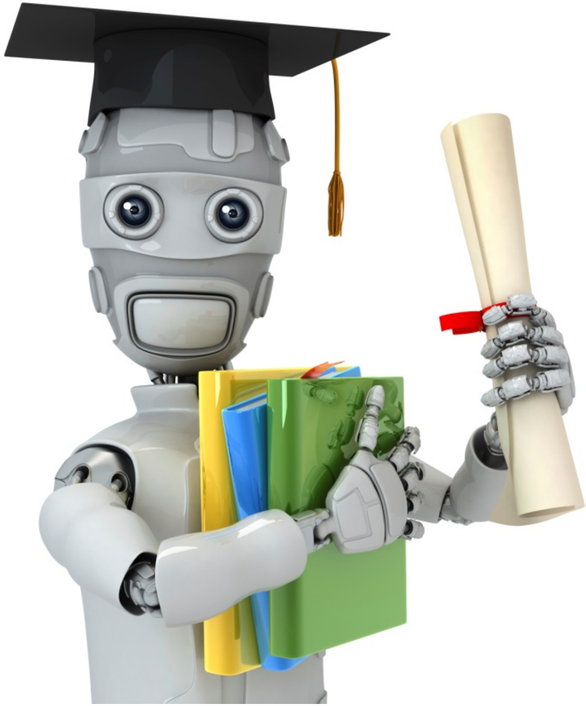
$\boxed{\text{prediction}}$ = DataMatrix $\times$ parameters

4x1

for i = 1:1000,

prediction (i) = ...

Andrew Ng



Machine Learning

Linear Algebra review (optional)

Matrix-matrix multiplication

Example

$$\underbrace{\begin{bmatrix} 1 & 3 & 2 \\ 4 & 0 & 1 \end{bmatrix}}_{2\times 3} \underbrace{\begin{bmatrix} \boxed{\begin{matrix}1\\0\\5\end{matrix}} & \boxed{\begin{matrix}3\\1\\2\end{matrix}} \end{bmatrix}}_{3\times 2} = \underbrace{\begin{bmatrix} 11 & 10 \\ 9 & 14 \end{bmatrix}}_{2\times 2}$$

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

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

Details:

$$A \times B = C$$

$$\left[\quad\right] \times \left[\quad\right] = \left[\quad\right]$$

n × 1

A	B	C
m x n matrix (m rows, n columns)	n x o matrix (n rows, o columns)	m x 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,…,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:

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

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

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

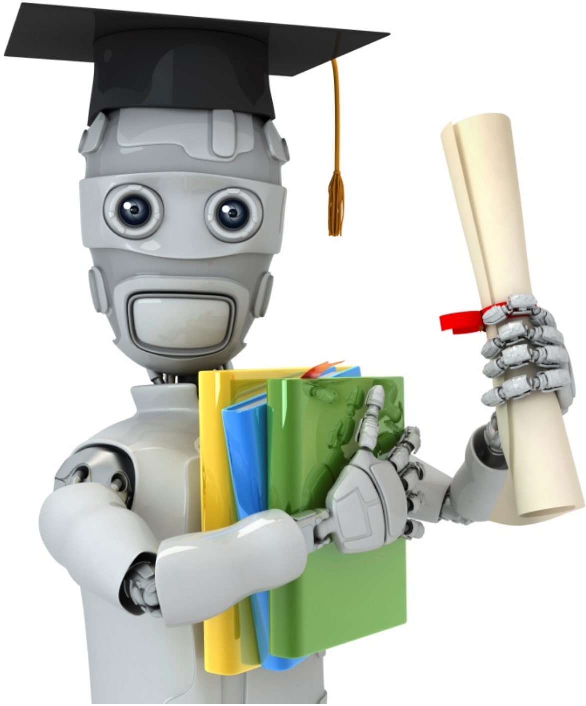
$\times$

Matrix

$$\begin{bmatrix} -40 & 200 & -150 \\ 0.25 & 0.1 & 0.4 \end{bmatrix} = \begin{bmatrix} 486 & 410 & 692 \\ 314 & 342 & 416 \\ 344 & 353 & 464 \\ 173 & 285 & 191 \end{bmatrix}$$

Prediction of first h_θ

Predictions of 2^{nd} h_θ



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

$\neq$

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

$\underset{m \times n}{A} \times \underset{n \times m}{B}$

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

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

$3 \times 5 \times 2$

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

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

$\boxed{\text{"Associative"}}$

$A \times (B \times C)$

$(A \times B) \times C$

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

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

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

$A \times (B \times C)$

$(A \times B) \times C$

Same answer.

Identity Matrix

1 is identity

$\boxed{1 \times z = z \times 1 = z}$

for any z

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

Examples of identity matrices:

$[1]$

1×1

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

2 x 2   3 x 3   4 x 4

Informally:

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

For any matrix A,

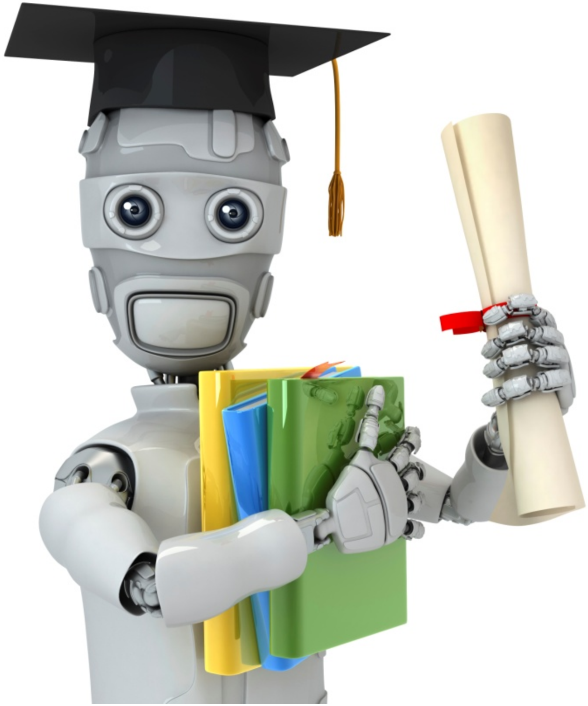
$$\underset{m\times n}{A} \cdot \underset{n\times n}{\boxed{I}} = \underset{m\times m}{\boxed{I}} \cdot \underset{m\times n}{A} = \underset{m\times n}{A}$$

$I_{n \times n}$

Note:

$AB \neq BA$ in general

$AI = IA$ ✓



Machine Learning

Linear Algebra review (optional)

Inverse and transpose

$1 =$ "identity"

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

$\frac{1}{3}$

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

$\frac{1}{12}$

$0 \, (0^{-1})$ undefined

Not all numbers have an inverse.

Matrix inverse:

square matrix (#rows = #columns)

A^{-1}

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

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

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

E.g.

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

$A^{-1}A$

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

Andrew Ng

Matrix Transpose

Example: $A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 5 & 9 \end{bmatrix}$ 2x3 $\qquad B = A^T = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 0 & 9 \end{bmatrix}$ 3x2

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

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

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

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