

Matrices:

Consider the matrix:

$$\begin{bmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{bmatrix}$$

Above matrix can be represented as

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

This symbol means that the matrix is an element of the family of matrices having 4 rows and 2 columns.

Vectors:

Vectors are matrices with n rows but only one column.

Vector: An $n \times 1$ matrix.

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

$n = 4$

← 4-dimensional vector.

$$\mathbb{R}^4$$

$y_i = i^{\text{th}}$ element

$$\begin{aligned} y_1 &= 460 \\ y_2 &= 232 \\ y_3 &= 315 \end{aligned}$$

1-indexed vs 0-indexed:

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

1-indexed

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

0-indexed

MATRIX ADDITION:

Matrix Addition

$$\begin{array}{l} \downarrow \downarrow \\ \rightarrow \begin{bmatrix} 1 & 0 \\ 2 & 5 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 0.5 \\ 2 & 5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 0.5 \\ 4 & 10 \\ 3 & 2 \end{bmatrix} \\ \text{3x2 matrix} \quad \text{3x2} \quad \text{3x2} \end{array}$$

$$\begin{array}{l} \rightarrow \begin{bmatrix} 1 & 0 \\ 2 & 5 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 0.5 \\ 2 & 5 \end{bmatrix} = \text{error} \\ \text{3x2} \quad \text{2x2} \end{array}$$

their sum is not well-defined.

Andrew Ng

Scalar Multiplication

← real number

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

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

Finally, for a slightly

Andrew Ng

Combination of Operands

$$\begin{aligned}
 & \text{Scalar multiplication} \rightarrow 3 \times \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} / 3 \quad \text{Scalar division} \\
 & = \begin{bmatrix} 3 \\ 12 \\ 6 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ \frac{2}{3} \end{bmatrix} \quad \begin{array}{l} \text{matrix subtraction /} \\ \text{vector subtraction} \end{array} \\
 & = \begin{bmatrix} 2 \\ 12 \\ 10\frac{1}{3} \end{bmatrix} \quad \begin{array}{l} \text{matrix addition /} \\ \text{vector addition} \end{array} \\
 & \quad \quad \quad \begin{array}{l} 3 \times 1 \text{ matrix} \\ 3\text{-dimensional vector} \end{array}
 \end{aligned}$$

Andrew Ng

Matrix vector Multiplication:

Example

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

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

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

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

So that's how you multiply a matrix and a vector.

Andrew Ng

Example

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

3×2 2×1 3×1 matrix

$1 \times 1 + 3 \times 5 = 16$
 $4 \times 1 + 0 \times 5 = 4$
 $2 \times 1 + 1 \times 5 = 7$

at what just happened and what

Andrew Ng

Predicting house prices using matrix vector multiplication:

House sizes:

- 2104
- 1416
- 1534
- 852

Matrix

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

4×2

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

$h_{\theta}(x)$

2×1 Vector

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

4×1 matrix

$h_{\theta}(2104)$

$h_{\theta}(1416)$

$h_{\theta}(1534)$

$h_{\theta}(852)$

$\text{prediction} = \text{Data Matrix} \times \text{Parameters}$

4×1

for $i = 1, \dots, 1000$,
 prediction(i) = ...

say we have 1000 houses & want to predict their price w/o using matrices are task becomes tedious.

Multiplying two MATRICES:

Example

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

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

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

Andrew Ng

Eg 2:

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

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

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

Predicting Housing prices using 3 equations simultaneously:

House sizes:

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

Have 3 competing hypotheses:

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

Matrix

$$\begin{bmatrix} 1 & 2104 \\ 1 & 1416 \\ 1 & 1534 \\ 1 & 852 \end{bmatrix} \times \begin{bmatrix} -40 \\ 0.25 \end{bmatrix} \begin{bmatrix} 200 \\ 0.1 \end{bmatrix} \begin{bmatrix} -150 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 486 & 410 & 692 \\ 314 & 342 & 416 \\ 344 & 353 & 464 \\ 173 & 285 & 191 \end{bmatrix}$$

Prediction of first h_{θ}

Predictions of 2nd h_{θ}

Andrew Ng

Properties of Matrix Multiplication:

1. Matrix multiplication is not commutative

$$3 \times 5 \neq 5 \times 3 \quad \text{"Commutative"}$$

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

E.g.

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

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

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

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

Matrix multiplication is Associative:

$$\underline{3 \times 5 \times 2} \quad 3 \times (5 \times 2) = (3 \times 5) \times 2$$

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

"Associative"

Let A, B, C be 3 matrices

$$A \times (B \times C) \quad \leftarrow \text{same}$$

$$(A \times B) \times C \quad \leftarrow \text{same}$$

$$A \times B \times C$$

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

Identity matrix:

Identity Matrix

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

Examples of identity matrices:

$$[1] \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

1×1 2×2 3×3 4×4

1 is identity.

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

for any z

Informally:

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

Informal representations of identity matrix

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:

$$AB \neq BA \text{ in general}$$

$$AI = IA \checkmark$$

Inverse of a Matrix:

1 = "identity."

$$3 \left[\frac{1}{3} \right] = 1$$

$$12 \times \left(\frac{1}{12} \right) = 1$$

Not all numbers have an inverse.

$$0 \left(\frac{1}{0} \right) \text{ undefined}$$

Matrix inverse:

square matrix
(#rows = #columns)

A^{-1}

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

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

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

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

$A \qquad A^{-1} \qquad A^{-1}A$

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

Andrew Ng

Transpose of a Matrix:

In transpose, rows become columns and columns become rows.

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

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

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

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

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

in fact concludes our linear algebra review.

Andrew Ng

Multivariate Linear Regression:

$$\rightarrow h_{\theta}(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_n x_n$$

For convenience of notation, define $x_0 = 1$. ($x_0^{(i)} = 1$)

$$x = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^{n+1} \quad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{bmatrix} \in \mathbb{R}^{n+1}$$

$$h_{\theta}(x) = \theta_0 x_0 + \theta_1 x_1 + \dots + \theta_n x_n$$

$\downarrow = 1$

$$= \boxed{\theta^T x}$$

Transpose of θ \times x

Matrix $\theta^T x$

Matrix x

Multivariate linear regression. \leftarrow

$$\text{Transpose of } \theta = [\theta_0, \theta_1, \theta_2, \dots, \theta_n]$$

We assume $x_0 = 1$ so that we can form a matrix of x of $n+1$ terms from x_0 to x_{n+1} .

Multiple Features

Note: $[7:25 - \theta^T]$ is a 1 by $(n+1)$ matrix and not an $(n+1)$ by 1 matrix

Linear regression with multiple variables is also known as "multivariate linear regression".

We now introduce notation for equations where we can have any number of input variables.

$x_j^{(i)}$ = value of feature j in the i^{th} training example
 $x^{(i)}$ = the input (features) of the i^{th} training example
 m = the number of training examples
 n = the number of features

feature means columns & here j represents columns whereas i represents rows.
 so x represents 4th row

The multivariable form of the hypothesis function accommodating these multiple features is as follows:

$$h_{\theta}(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots + \theta_n x_n$$

In order to develop intuition about this function, we can think about θ_0 as the basic price of a house, θ_1 as the price per square meter, θ_2 as the price per floor, etc. x_1 will be the number of square meters in the house, x_2 the number of floors, etc.

Using the definition of matrix multiplication, our multivariable hypothesis function can be concisely represented as:

$$h_{\theta}(x) = \theta_0 \quad \theta_1 \quad \dots \quad \begin{matrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{matrix} = \theta^T x$$

whereas x_1 represents data there in 1st column of 4th row.

Remark: Note that for convenience reasons in this course we assume $x_0^{(i)} = 1$ for $(i \in 1, \dots, m)$. This allows us to do matrix operations with θ and x . Hence making the two vectors ' θ ' and $x^{(i)}$ match each other element-wise (that is, have the same number of elements: $n+1$).]