linear algebra basics for machine learning

matrices and vectors

a matrix is a rectangular array of numbers:

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

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

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

the dimensions of a matrix are donated as **rows^n x columns^n**

Which of the following statements are true? Check all that apply.

- [x] $\begin{bmatrix} 1 & 2 \\ 4 & 0 \\ 0 & 1 \end{bmatrix}$ is a 3×2 matrix.

Correct Response

- [ ] $\begin{bmatrix} 0 & 1 & 4 & 2 \\ 3 & 4 & 0 & 9 \end{bmatrix}$ is a 4×2 matrix.

Correct Response

- [x] $\begin{bmatrix} 0 & 4 & 2 \\ 3 & 4 & 9 \\ 5 & -1 & 0 \end{bmatrix}$ is a 3×3 matrix.

Correct Response

- [x] $\begin{bmatrix} 1 & 2 \end{bmatrix}$ is a 1×2 matrix.

Correct Response

matrix elements (entries of a matrix):

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix} \quad \rightarrow \quad A_{ij} = i,j \text{ entry in the } i^{th} \text{ row}, j^{th}, \text{column}$$

for example: $A_{11} = 1,\ A_{12} = 2,\ A_{32} = 6,$ and $A_{41} = 7;\ A_{43} = \text{undefined (error)}$

a vector is an **n x 1** matrix:

$$y = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \quad \rightarrow \quad 4 - \text{dimensional vector}; \quad \mathbb{R}^4$$

$$y_i = i^{th} \text{ element} \ \rightarrow \ y_1 = 1,\ y_2 = 2,\ y_3 = 3$$

1-indexed vs 0-indexed

$$y = \begin{bmatrix} \boldsymbol{y_1} \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \qquad y = \begin{bmatrix} \boldsymbol{y_0} \\ y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

addition and scalar multiplication

matrix addition only applies when matrices have the same dimensions:

$$\underset{\mathbb{R}^{3\times 2}}{\begin{bmatrix}1 & 0\\ 2 & 5\\ 3 & 1\end{bmatrix}} + \underset{\mathbb{R}^{3\times 2}}{\begin{bmatrix}4 & 0.5\\ 2 & 5\\ 0 & 1\end{bmatrix}} = \underset{\mathbb{R}^{3\times 2}}{\begin{bmatrix}5 & 0.5\\ 2 & 10\\ 3 & 2\end{bmatrix}}$$

$$\underset{\mathbb{R}^{3\times 2}}{\begin{bmatrix}1 & 0\\ 2 & 5\\ 3 & 1\end{bmatrix}} + \underset{\mathbb{R}^{2\times 2}}{\begin{bmatrix}4 & 0.5\\ 2 & 5\end{bmatrix}} = \textbf{error}$$

scalar (real number) matrix multiplication

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

$$\underset{\mathbb{R}^{3\times 2}}{\begin{bmatrix}4 & 0\\ 6 & 3\end{bmatrix}} /4 = \frac{1}{4} \times \underset{\mathbb{R}^{3\times 2}}{\begin{bmatrix}4 & 0\\ 6 & 3\end{bmatrix}} = \underset{\mathbb{R}^{3\times 2}}{\begin{bmatrix}0 & 0\\ \frac{3}{2} & \frac{3}{4}\end{bmatrix}}$$

combination of operands:

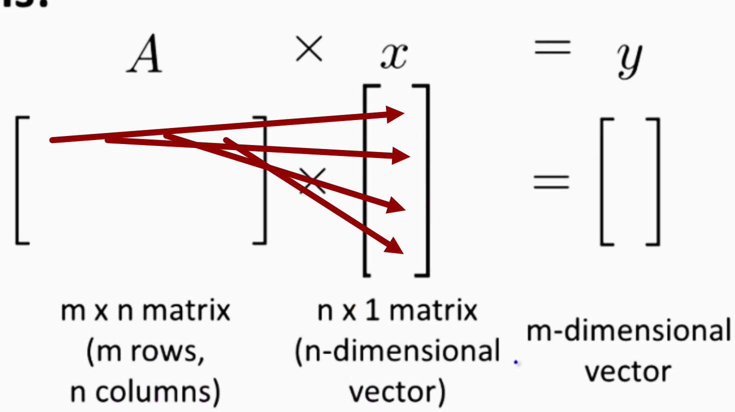
$$3 \times \begin{bmatrix}1\\ 4\\ 2\end{bmatrix} + \begin{bmatrix}0\\ 0\\ 5\end{bmatrix} - \begin{bmatrix}3\\ 0\\ 2\end{bmatrix} /3 = \begin{bmatrix}3\\ 12\\ 6\end{bmatrix} + \begin{bmatrix}0\\ 0\\ 5\end{bmatrix} + \begin{bmatrix}1\\ 0\\ \frac{2}{3}\end{bmatrix} = \begin{bmatrix}2\\ 12\\ 10\frac{1}{3}\end{bmatrix}$$

$$\begin{bmatrix}4\\ 6\\ 7\end{bmatrix} /2 - 3\begin{bmatrix}2\\ 1\\ 0\end{bmatrix} = \begin{bmatrix}2\\ 3\\ 3.5\end{bmatrix} - \begin{bmatrix}6\\ 3\\ 0\end{bmatrix} = \begin{bmatrix}-4\\ 0\\ 3.5\end{bmatrix}$$

matrix vector multiplication

$$\underset{\mathbb{R}^{3\times 2}}{\begin{bmatrix}1 & 3\\ 4 & 0\\ 2 & 1\end{bmatrix}} \times \underset{\mathbb{R}^{2\times 1}}{\begin{bmatrix}1\\ 5\end{bmatrix}} = \begin{bmatrix}1 \times 1 + 3 \times 5 = 16\\ 4 \times 1 + 0 \times 5 = 4\\ 2 \times 1 + 1 \times 5 = 7\end{bmatrix} = \underset{\mathbb{R}^{3\times 1}}{\begin{bmatrix}16\\ 4\\ 7\end{bmatrix}}$$

Details:



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

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

$$\mathbb{R}^{3\times 4} \quad \mathbb{R}^{4\times 1} \longrightarrow \mathbb{R}^{3\times 1}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 5 \\ 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \times 1 + 0 \times 6 + 3 \times 2 = 7 \\ 2 \times 1 + 1 \times 6 + 5 \times 2 = 18 \\ 3 \times 1 + 1 \times 6 + 2 \times 2 = 13 \end{bmatrix} = \begin{bmatrix} 7 \\ 18 \\ 13 \end{bmatrix}$$

$$\mathbb{R}^{3\times 3} \quad \mathbb{R}^{3\times 1} \longrightarrow \mathbb{R}^{3\times 1}$$

applying **matrix vector** multiplication to a function $h_\theta(x)$:

examples:

2104
1416
1534
852

hypothesis:

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

matrix x vector = predicted values $h_\theta(x)$ = $h_\theta(x)$

$$\begin{bmatrix} 1 & 2104 \\ 1 & 1416 \\ 1 & 1534 \\ 1 & 852 \end{bmatrix} \times \begin{bmatrix} -40 \\ 0.25 \end{bmatrix} = \begin{bmatrix} -40 \times 1 + 0.25 \times 2104 = 486 \\ -40 \times 1 + 0.25 \times 1416 = 314 \\ -40 \times 1 + 0.25 \times 1534 = 343.5 \\ -40 \times 1 + 0.25 \times 852 = 173 \end{bmatrix} = \begin{bmatrix} 486 \\ 314 \\ 343.5 \\ 173 \end{bmatrix}$$

$$\mathbb{R}^{4\times 2} \quad \mathbb{R}^{2\times 1} \longrightarrow \mathbb{R}^{4\times 1}$$

matrix matrix multiplication

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

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

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

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

Details:

$$\underset{\substack{\text{m x n matrix}\\\text{(m rows,}\\\text{n columns)}}}{\overset{A}{\begin{bmatrix} \quad & \quad \end{bmatrix}}} \times \underset{\substack{\text{n x o matrix}\\\text{(n rows,}\\\text{o columns)}}}{\overset{B}{\begin{bmatrix} \quad & \quad \end{bmatrix}}} = \underset{\substack{\text{m x o}\\\text{matrix}}}{\overset{C}{\begin{bmatrix} \quad & \quad \end{bmatrix}}}$$

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)

$$\begin{bmatrix} \mathbf{1} & \mathbf{3} \\ \mathbf{2} & \mathbf{5} \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{3} & \mathbf{2} \end{bmatrix} = \begin{bmatrix} \mathbf{9} & \mathbf{4} \\ \mathbf{15} & \mathbf{12} \end{bmatrix}$$

$$\mathbb{R}^{\mathbf{2\times2}} \qquad \mathbb{R}^{\mathbf{3\times2}} \longrightarrow \mathbb{R}^{\mathbf{2\times2}}$$

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1\times0+3\times3=9 \\ 2\times0+5\times3=15 \end{bmatrix} = \begin{bmatrix} \mathbf{9} \\ \mathbf{15} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1\times1+3\times2=7 \\ 2\times1+5\times2=12 \end{bmatrix} = \begin{bmatrix} \mathbf{7} \\ \mathbf{12} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{1} & \mathbf{3} \\ \mathbf{2} & \mathbf{4} \\ \mathbf{0} & \mathbf{5} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{2} & \mathbf{3} \end{bmatrix} = \begin{bmatrix} \mathbf{7} & \mathbf{9} \\ \mathbf{10} & \mathbf{12} \\ \mathbf{10} & \mathbf{15} \end{bmatrix}$$

$$\mathbb{R}^{\mathbf{3\times2}} \qquad \mathbb{R}^{\mathbf{2\times2}} \longrightarrow \mathbb{R}^{\mathbf{2\times2}}$$

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1\times1+3\times2=7 \\ 2\times1+4\times2=10 \\ 0\times1+5\times2=10 \end{bmatrix} = \begin{bmatrix} \mathbf{7} \\ \mathbf{10} \\ \mathbf{10} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1\times0+3\times3=9 \\ 2\times0+4\times3=12 \\ 0\times0+5\times3=15 \end{bmatrix} = \begin{bmatrix} \mathbf{9} \\ \mathbf{12} \\ \mathbf{15} \end{bmatrix}$$

applying **matrix matrix** multiplication to a function $h_\theta(x)$:

examples:

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

hypothesis:

$$h_\theta(x) = -40 + 0.25x$$
$$h_\theta(x) = 200 + 0.10x$$
$$h_\theta(x) = -150 + 0.40x$$

matrix x matrix = predicted values $h_\theta(x)$

$$\begin{bmatrix} \mathbf{1} & \mathbf{2104} \\ \mathbf{1} & \mathbf{1416} \\ \mathbf{1} & \mathbf{1534} \\ \mathbf{1} & \mathbf{852} \end{bmatrix} \times \begin{bmatrix} \mathbf{-40} & \mathbf{200} & \mathbf{-150} \\ \mathbf{0.25} & \mathbf{0.10} & \mathbf{0.40} \end{bmatrix} = \begin{bmatrix} \mathbf{486} & \mathbf{410.4} & \mathbf{692.6} \\ \mathbf{314} & \mathbf{341.6} & \mathbf{416.4} \\ \mathbf{343.5} & \mathbf{353.4} & \mathbf{464.6} \\ \mathbf{173} & \mathbf{285.2} & \mathbf{191.8} \end{bmatrix}$$

$$\mathbb{R}^{\mathbf{4\times2}} \qquad \mathbb{R}^{\mathbf{2\times3}} \longrightarrow \mathbb{R}^{\mathbf{4\times3}}$$

$$\begin{bmatrix} 1 & 2104 \\ 1 & 1416 \\ 1 & 1534 \\ 1 & 852 \end{bmatrix} \times \begin{bmatrix} -40 \\ 0.25 \end{bmatrix} = \begin{bmatrix} 1 \times -40 + 2104 \times 0.25 = 486 \\ 1 \times -40 + 1416 \times 0.25 = 314 \\ 1 \times -40 + 1534 \times 0.25 = 343.5 \\ 1 \times -40 + 852 \times 0.25 = 173 \end{bmatrix} = \begin{bmatrix} \mathbf{486} \\ \mathbf{314} \\ \mathbf{343.5} \\ \mathbf{173} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2104 \\ 1 & 1416 \\ 1 & 1534 \\ 1 & 852 \end{bmatrix} \times \begin{bmatrix} 200 \\ 0.10 \end{bmatrix} = \begin{bmatrix} 1 \times 200 + 2104 \times 0.10 = \\ 1 \times 200 + 1416 \times 0.10 = \\ 1 \times 200 + 1534 \times 0.10 = \\ 1 \times 200 + 852 \times 0.10 = \end{bmatrix} = \begin{bmatrix} \mathbf{410.4} \\ \mathbf{341.6} \\ \mathbf{353.4} \\ \mathbf{285.2} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2104 \\ 1 & 1416 \\ 1 & 1534 \\ 1 & 852 \end{bmatrix} \times \begin{bmatrix} -150 \\ 0.40 \end{bmatrix} = \begin{bmatrix} 1 \times -150 + 2104 \times 0.40 = \\ 1 \times -150 + 1416 \times 0.40 = \\ 1 \times -150 + 1534 \times 0.40 = \\ 1 \times -150 + 852 \times 0.40 = \end{bmatrix} = \begin{bmatrix} \mathbf{691.6} \\ \mathbf{416.4} \\ \mathbf{463.6} \\ \mathbf{190.8} \end{bmatrix}$$

matrix multiplication properties

multiplication is **not commutative** when working with the multiplication of matrices and scalars:

in multiplication of real numbers 3 x 5 is the same as multiplying 5 x 3; this is the property of being mathematically commutative

for matrices A and B; in general $A \times B \neq B \times A$ (not commutative):

$$\left[\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right] \neq \left[\begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{2} & \mathbf{2} \end{bmatrix}\right]$$

if A is a $m \times n$ and B is a $n \times m$ matrix; $A \times B$ is a $m \times m$ and $B \times A$ is an $n \times n$ matrix multiplication of real numbers have an **associative** property and shares this property with the multiplication of matrices:

multiplying real number $3 \times (5 \times 2) = 3 \times 10 = (3 \times 5) \times 2 = 15 \times 2$ equally

for matrices A, B and C; the computation $A \times (B \times C) = (A \times B) \times C$ equally:

let $D = B \times C$; compute $A \times D \rightarrow A \times (B \times C)$

let $E = A \times B$; compute $E \times C \rightarrow (A \times B) \times C$ the results are **associative**

identity matrix

a matrix with the property of 1's along the diagonals and 0's everywhere else:

denoted I or $I_{n \times n}$ with examples as follows:

$$\underset{I_{2\times2}}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \qquad \underset{I_{3\times3}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \qquad \underset{I_{4\times4}}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}$$

for any matrix A, $A \times I = I \times A = A$

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

although matrix matrix multiplication is generally **not commutative**, matrix matrix multiplication is **commutative** when the product involves an identity matrix: $AB \neq BA$ in general, however $AI = IA$ as described in the latter

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

inverse and transpose

$1 = \text{"identity"}$ $\qquad 3(3^{-1}) = \frac{1}{3} = 1$ $\qquad 12 \times (12^{-1}) = 1$

not all numbers have an inverse: $0(0^{-1}) = \text{undefined}$

Matrix inverse:

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

$$AA^{-1} = A^{-1}A = I.$$

an $m \times m$ matrix is also a square matrix (number of rows = number of columns)

$$\underset{A}{\begin{bmatrix} 3 & 4 \\ 2 & 16 \end{bmatrix}} \underset{A^{-1}}{\begin{bmatrix} 0.4 & -0.1 \\ -0.05 & 0.075 \end{bmatrix}} = \underset{A^{-1}A}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} = I_{2\times2}$$

$A^{-1}A$; computation of A^{-1} through `pinv(A)` in octave

matrices that do not have an inverse are "**singular**" or "**degenerate**" matrices

Matrix Transpose

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

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

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

for example: $B_{12} = A_{21} = 2$ and $B_{32} = A_{23} = 9$ e.g. $\begin{bmatrix} 0 & 3 \\ 1 & 4 \end{bmatrix}^T = \begin{bmatrix} 0 & 1 \\ 3 & 4 \end{bmatrix}$