Matrices and Vectors

A matrix is a rectangular array of numbers written between square brackets.

Matrix: Rectangular array of numbers:

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

Example of 4×2 matrix ($\mathbb{R}^{4 \times 2}$). Dimension of matrix: number of rows x number of columns

Matrices are 2-dimensional arrays:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \\ j & k & l \end{bmatrix}$$

The above matrix has four rows and three columns, so it is a 4×3 matrix ($\mathbb{R}^{4 \times 3}$).

A vector is a matrix with one column and many rows:

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix}$$

So vectors are a subset of matrices. The above vector is a 4×1 matrix ($\mathbb{R}^4$).

Notation and terms:

- A_{ij} refers to the element in the i^{th} row and j^{th} column of matrix A.
- A vector with 'n' rows is referred to as an 'n'-dimensional vector.
- v_i refers to the element in the i^{th} row of the vector.
- In general, all our vectors and matrices will be 1-indexed. Note that for some programming languages, the arrays are 0-indexed.
- Matrices are usually denoted by uppercase names while vectors are lowercase.
- "Scalar" means that an object is a single value, not a vector or matrix.
- $\mathbb{R}$ refers to the set of scalar real numbers.
- $\mathbb{R}^n$ refers to the set of n-dimensional vectors of real numbers.

Addition and Scalar Multiplication

Addition and subtraction are **element-wise**, so we simply add or subtract each corresponding element:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} a + w & b + x \\ c + y & d + z \end{bmatrix}$$

Subtracting Matrices:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} - \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} a - w & b - x \\ c - y & d - z \end{bmatrix}$$

To add or subtract two matrices, their dimensions must be **the same**.

In scalar multiplication, we simply multiply every element by the scalar value:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} * x = \begin{bmatrix} a*x & b*x \\ c*x & d*x \end{bmatrix}$$

In scalar division, we simply divide every element by the scalar value:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} /x = \begin{bmatrix} a/x & b/x \\ c/x & d/x \end{bmatrix}$$

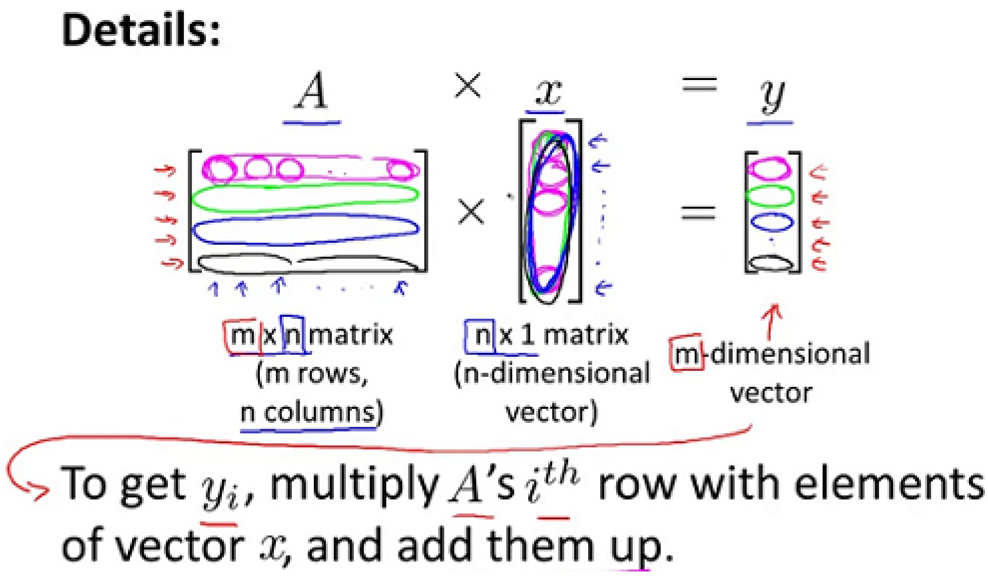
Example: $3 \times \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} - \begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} /3 = \begin{bmatrix} 3 \\ 12 \\ 6 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 2/3 \end{bmatrix} = \begin{bmatrix}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

Matrix-Vector Multiplication

We map the column of the vector onto each row of the matrix, multiplying each element and summing the result.

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a*x+b*y \\ c*x+d*y \\ e*x+b*y \end{bmatrix}$$

The result is a **vector**. The number of **columns** of the matrix must equal the number of **rows** of the vector.



An $m \times n$ matrix multiplied by an $n \times 1$ vector results in an $m \times 1$ vector.

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

(4×3)

Operations:

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

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

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

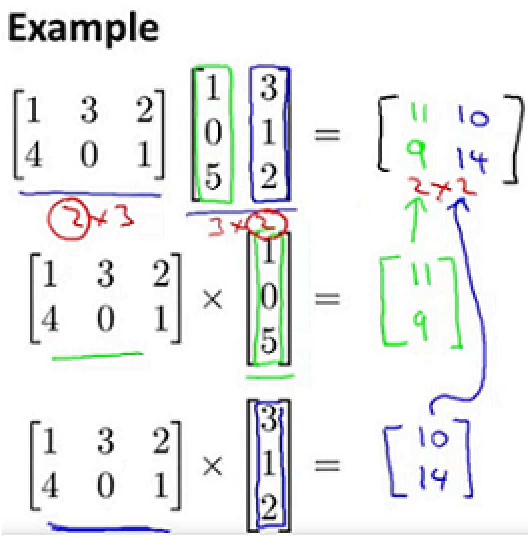
The resulting vector is an n-dimensional vector

Matrix-Matrix Multiplication

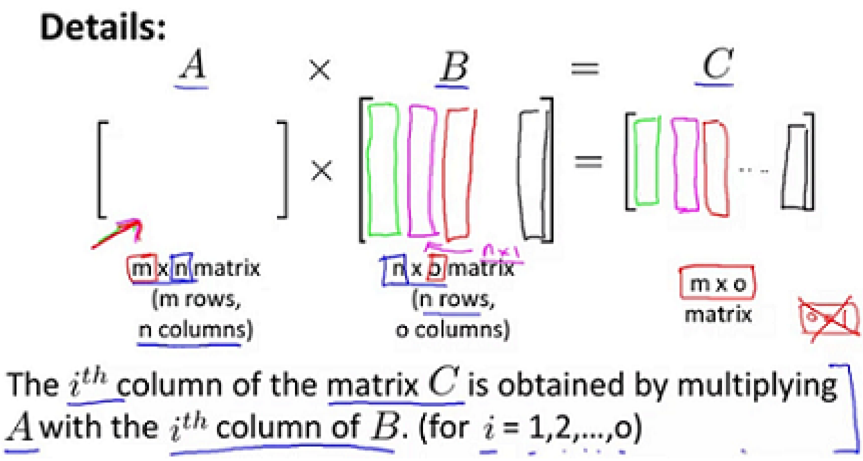
We multiply two matrices by breaking it into several vector multiplications and concatenating the result.

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} * \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} a*w+b*y & a*x+b*z \\ c*w+d*y & c*x+d*z \\ e*w+f*y & e*x+f*z \end{bmatrix}$$

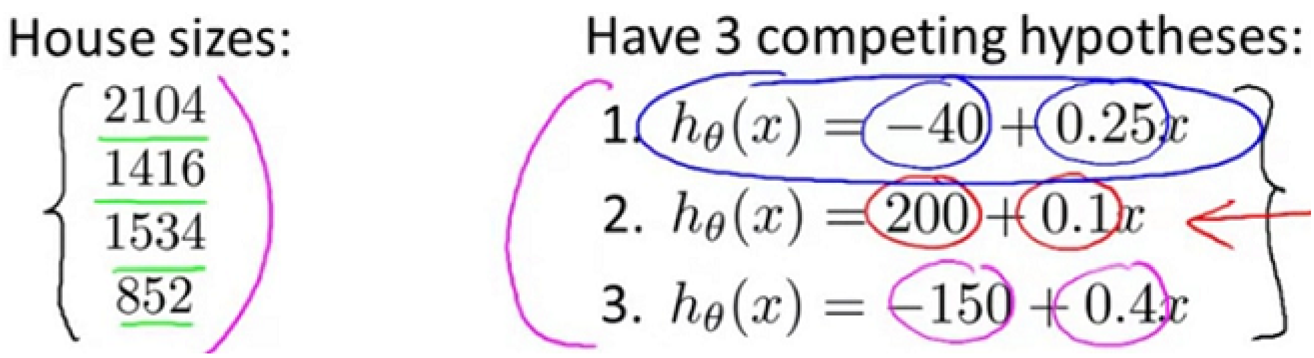
An $m \times n$ matrix multiplied by an $n \times o$ **matrix** results in an $m \times o$ matrix. In the above example, a 3×2 matrix times a 2×2 matrix resulted in a 3×2 matrix.



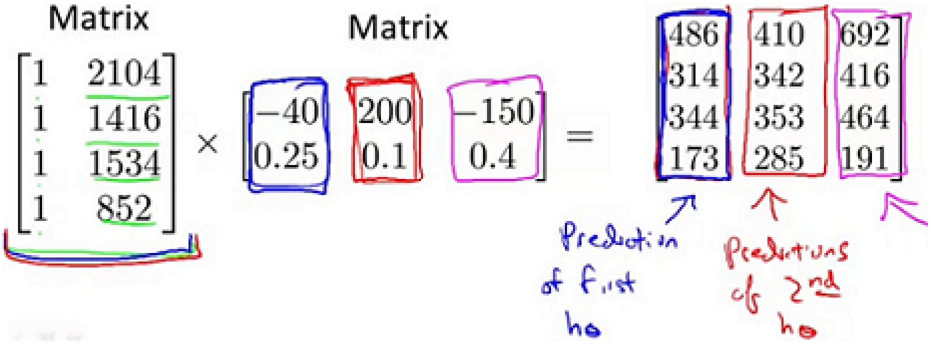
To multiply two matrices, the number of **columns** of the first matrix must equal the number of **rows** of the second matrix.



Example that we can do with matrix-matrix multiplication. Let's say, that we have four houses whose prices we wanna predict and and we have three competing hypotheses.

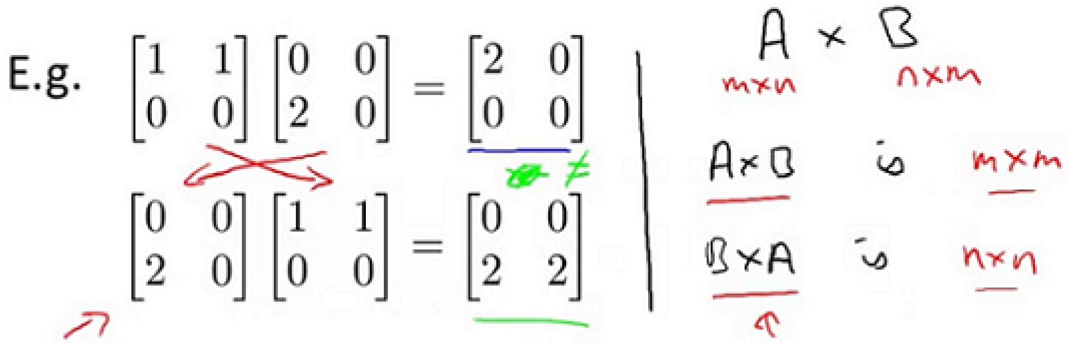


So if we want to apply all three competing hypotheses to all four our houses, it turns out we can do that very efficiently using a matrix-matrix multiplication.

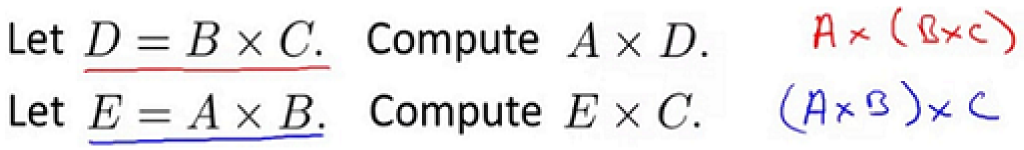


Matrix Multiplication Properties

Matrices are not commutative: $A * B \neq B * A$



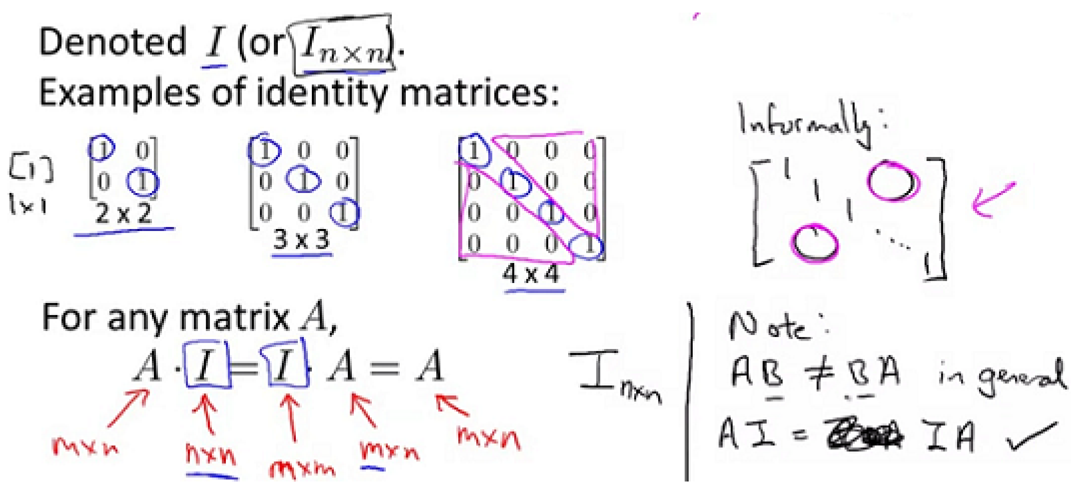
Matrices are associative: $(A * B) * C = A * (B * C)$



The identity matrix, when multiplied by any matrix of the same dimensions, results in the original matrix. It's just like multiplying numbers by 1. The identity matrix simply has 1's on the diagonal (upper left to lower right diagonal) and 0's elsewhere.

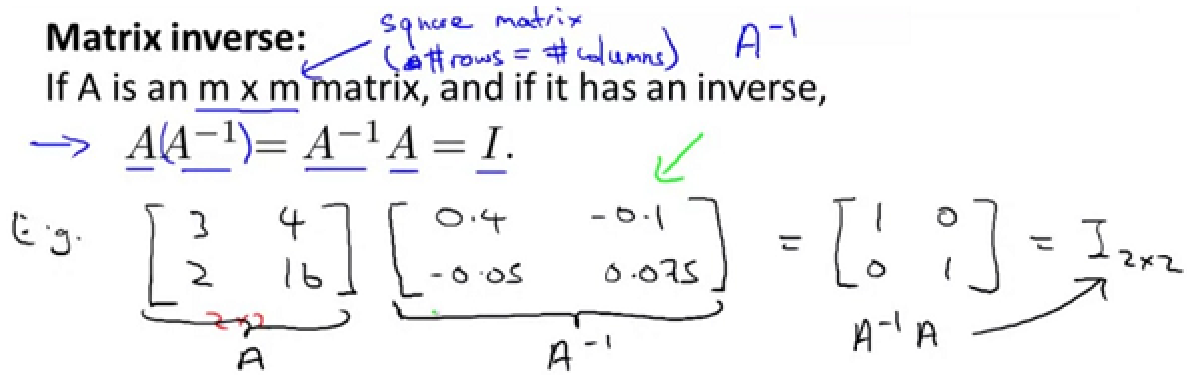
$$\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$$

When multiplying the identity matrix after some matrix (A∗I), the square identity matrix's dimension should match the other matrix's **columns**. When multiplying the identity matrix before some other matrix (I∗A), the square identity matrix's dimension should match the other matrix's **rows**.

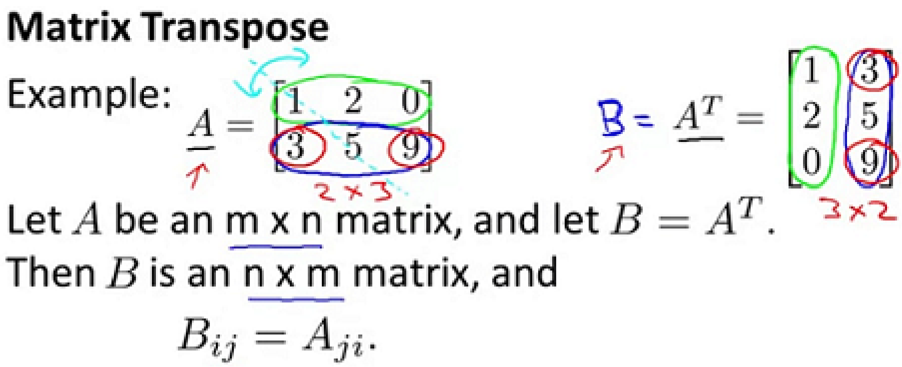


Inverse and Transpose

The **inverse** of a matrix A is denoted A^{-1}. Multiplying by the inverse results in the identity matrix.



A non square matrix does not have an inverse matrix. We can compute inverses of matrices in octave with the $pinv(A)$ function and in Matlab with the $inv(A)$ function. Matrices that don't have an inverse are **singular** or **degenerate**.



The **transposition** of a matrix is like rotating the matrix 90° in clockwise direction and then reversing it. We can compute transposition of matrices in matlab with the $transpose(A)$ function or A':

$$A = \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} A^{-1} = \begin{bmatrix} a & c & e \\ b & d & f \end{bmatrix}$$