VECTORS AND MATRICIES

RECAP OF LAST WEEK.

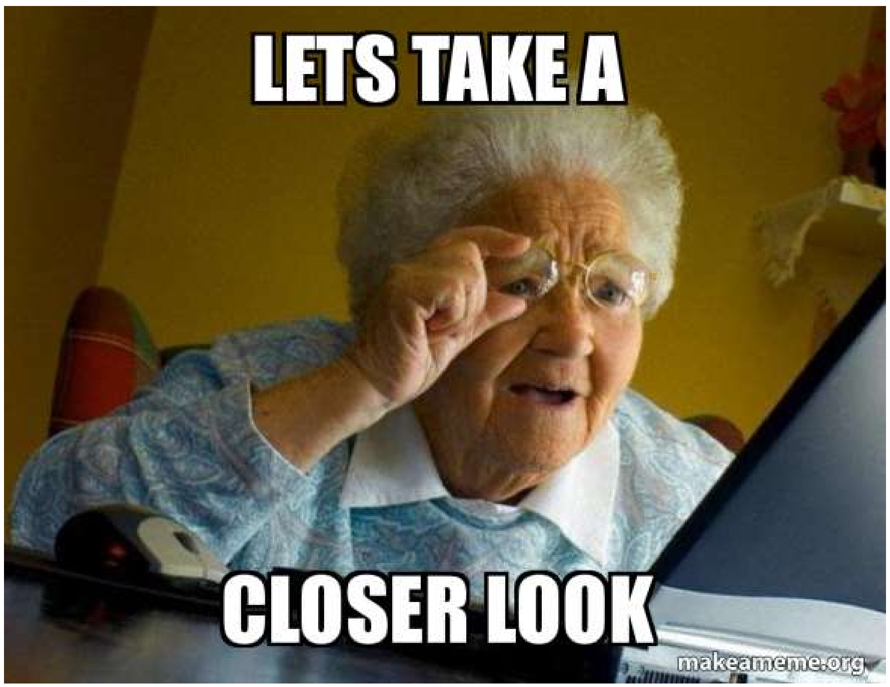
- Different sets of numbers. (Real numbers, Integers, Rational and Irrational Number).
  - Pi = Irrational. (Doesn't terminate)
  - 1/3 = Rational (Ratio of two numbers, that terminates)
- Latex

```
$$ \hat{\overline{y}} = \frac{\sum{y_i}}{n} \approx \overline{y} $$
```

$$\hat{\overline{y}} = \frac{\sum{y_i}}{n} \approx \overline{y}$$

- Set calculation (Union, Intersection, set minus etc.).
  - Order doesn't matter and number of occurrences doesn't matter. (This changes today).
- Refreshed our basic algebra exponents roots and solving inequalities.

RECAP OF LAST WEEK.



- Exercise 1.1)
  - A lot of you struggled with two of these.

(-x^4 * y^2)^2

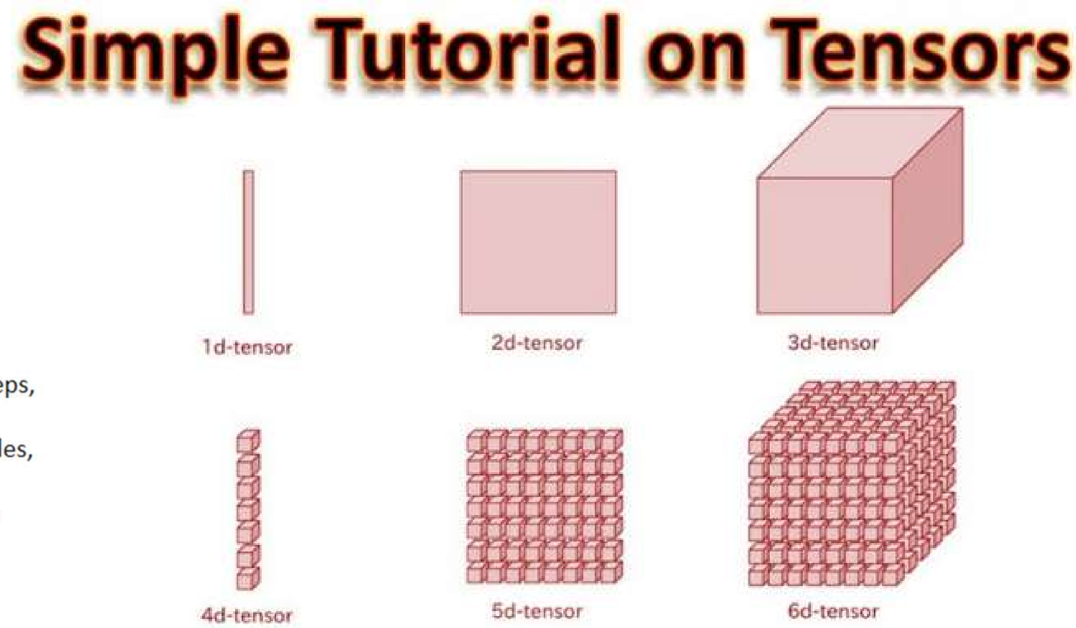
(1/27b^3)^1/3

THE PLAN FOR TODAY.

- Introducing you to the concept of tensor dimensions and dimensions. (They are different).

TENSORS:

- **Dimension and tensors.**
- Tensor dimensions describe the data structure. (scalar, matrix, multi dimensional arrays).



Types of Tensors:

 Vector data—2D tensors of shape (samples, features)

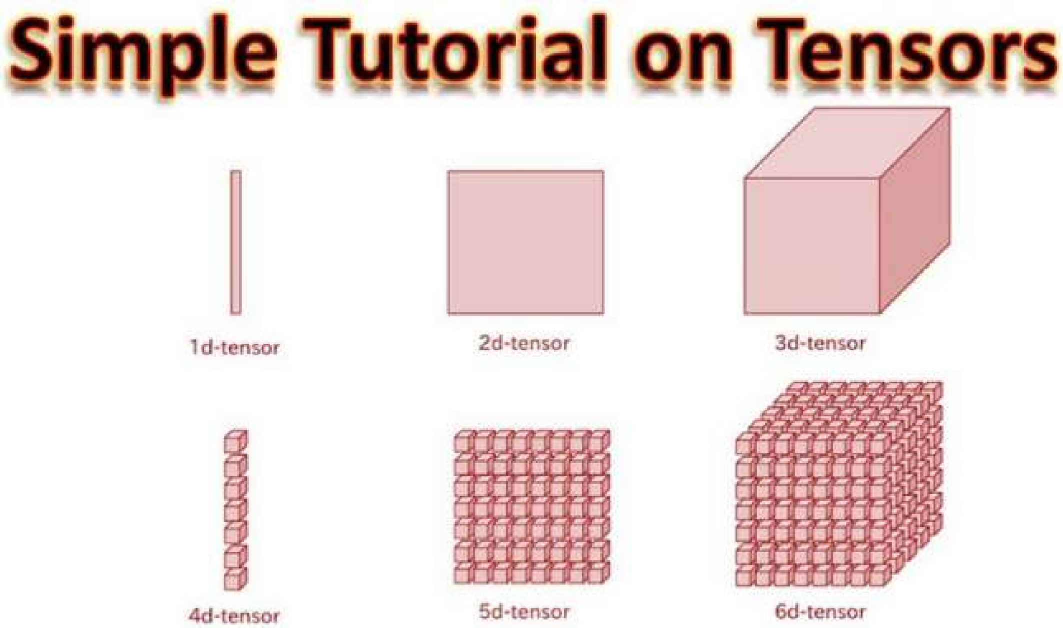
 Timeseries data or sequence data—3D tensors of shape (samples, timesteps, features)

 Images—4D tensors of shape (samples, height, width, channels) or (samples, channels, height, width)

 Video—5D tensors of shape (samples, frames, height, width, channels) or (samples, frames, channels, height, width)

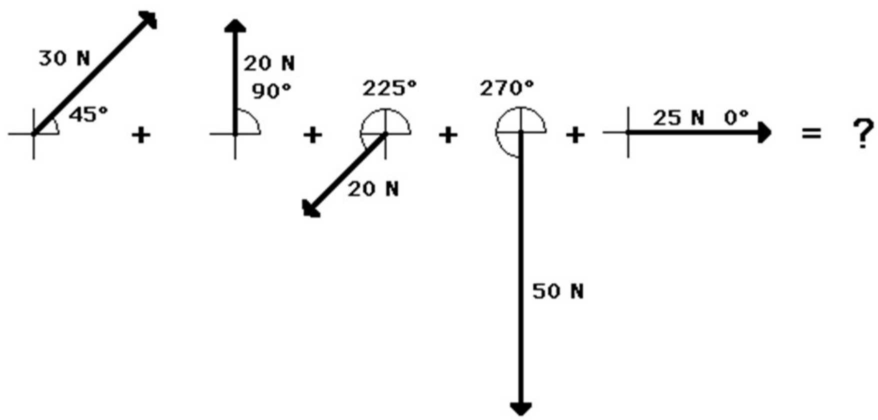
TENSORS:

- **Dimension and tensors.**
- x = [2,7,2]. 1D tensor and a 3D vector.
  - Don't confuse the two.
- A 3D vector is a 1D tensor with 1 axis and 3 dimensions along its one axis.
- A 10x8-D matrix is a 2D tensor. With 10 dimensions along its first axis and 8 along its second axis. (10rows, 8 col).



WHAT IS A VECTOR?

- It depends who you ask.. (not in the everything is relative sense).
- **Physicist.**
  - A vector is an arrow pointing in space.
  - What define it is it **length** and its **direction.**
  - You **can freely move** it around it doesn't have to start in **origo.** (still the same vector).



WHAT IS A VECTOR?

- It depends who you ask.. (not in the everything is relative sense).
- **Data-scientist.**
  - An ordered list of numbers. (unlike sets)
  - Multiple occurrences matters. (unlike sets)

- Let's say we wanted to model **Income ~ Age + Education + Country.**

  We would have a 3D vector for each sample. [Age, Education, Country].

  If we allowed [Age, Education, Country] = [Education, Age, Country] our predictions would be **really random!**

  Bind multiple 3D vectors together and we get a matrix (2D tensors) size: Samples-x-3

VECTORS SUM UP.

- What is their properties. (unlike sets)
  - A collection of values.
  - Order matters.
    - [3,2,1] != [1,2,3]
  - Number of occurrences also matters.

  - Column or Row vector.

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}^{\mathrm{T}} = \begin{bmatrix} x_1 \; x_2 \; \dots \; x_m \end{bmatrix}.$$

  - Transpose a vector.

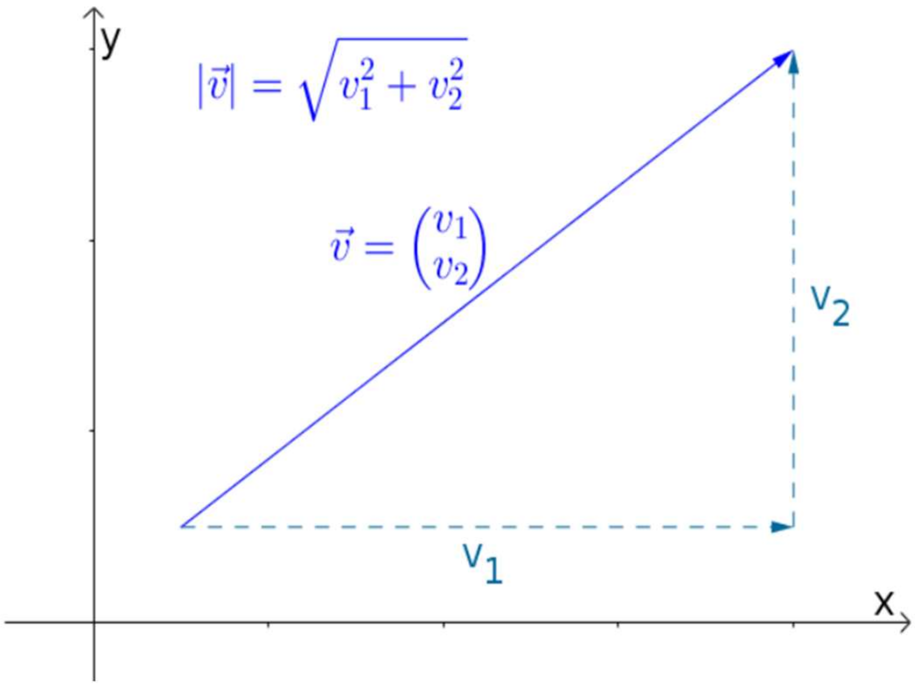
VECTOR LENGTH

The vector length (or magnitude) of $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ is:

$$|\vec{v}| = \left| \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right| = \sqrt{v_1^2 + v_2^2}$$

In three-dimensional space:

$$|\vec{v}| = \left| \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \right| = \sqrt{v_1^2 + v_2^2 + v_3^2}$$



DIFFERENT VECTOR OPERATIONS:

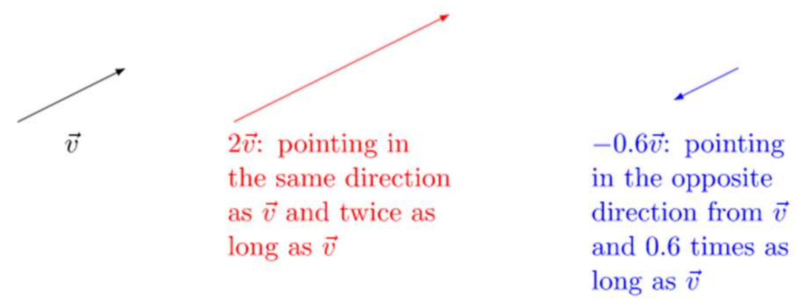
- Scalar Multiplication

- Any scalar operation on a vector is done element wise.

The first key operation is **scalar multiplication**, multiplying a scalar and a vector. If k is a scalar and $\vec{v}$ is a vector, their product $k\vec{v}$ is defined as follows:

- If $k > 0$, then $k\vec{v}$ is the vector pointing in the same direction as $\vec{v}$ that's k times as long as $\vec{v}$.
- If $k = 0$, then $k\vec{v}$ is $\vec{0}$.
- If $k < 0$, then $k\vec{v}$ is the vector pointing in the opposite direction from $\vec{v}$ that's $|k|$ times as long as $\vec{v}$.

For example, if $\vec{v}$ is the vector shown at left below, here's how you'd picture $2\vec{v}$ and $-0.6\vec{v}$:



VECTOR ADDITION

- Must be same dimensional vector.
- 2D not compatible with 3D unless you can make 3D into 2D.
- Same rules apply for subtraction.

The second key operation is **vector addition**, adding one vector to another. Here's how this is defined: if we have two vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$,[2] draw $\vec{v}$ (with its tail anywhere), and then draw $\vec{w}$ with its tail at the head of $\vec{v}$. Then, $\vec{v} + \vec{w}$ is defined to be the vector that goes from the tail of $\vec{v}$ to the head of $\vec{w}$.

two vectors

picturing the sum

$\mathbb{R}$ is the set of real numbers. That is, $\mathbb{R} = \{x : x \text{ is a real number}\}$.

$\mathbb{R}^2$ is the set of pairs of real numbers. That is,
$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) : x \text{ and } y \text{ are real numbers}\}$.

Definition 6.1. For two vectors $\vec{a} = \langle a_1, \ldots, a_n \rangle$ and $\vec{b} = \langle b_1, \ldots, b_n \rangle$ in $\mathbb{R}^n$, [3] the **dot product** $\vec{a} \cdot \vec{b}$ is the scalar $a_1 b_1 + \cdots + a_n b_n$.

Example 6.2. If $\vec{v} = \langle 1, 2, 3 \rangle$ and $\vec{w} = \langle 4, 5, 6 \rangle$, then $\vec{v} \cdot \vec{w} = (1)(4) + (2)(5) + (3)(6) = 32$.

Notice that the dot product is an operation between two vectors in $\mathbb{R}^n$ that produces a scalar. As you'll see in class, the key property of the dot product is this one:

Fact 6.3. *If θ is the angle between $\vec{a}$ and $\vec{b}$, then $\vec{a} \cdot \vec{b} = \|\vec{a}\| \|\vec{b}\| \cos\theta$. (Remember that $\|\vec{a}\|$ denotes the length of $\vec{a}$.)*

Example 6.4. Using this key fact, we can find the angle between $\vec{v} = \langle 1, 2, 3 \rangle$ and $\vec{w} = \langle 4, 5, 6 \rangle$. We already calculated in Example 6.2 that $\vec{v} \cdot \vec{w} = 32$. By the Pythagorean Theorem, $\|\vec{v}\| = \sqrt{1^2 + 2^2 + 3^2} = \sqrt{14}$ and $\|\vec{w}\| = \sqrt{4^2 + 5^2 + 6^2} = \sqrt{77}$. So, the key fact says that, if θ is the angle between $\vec{v}$ and $\vec{w}$, then $32 = (\sqrt{14})(\sqrt{77}) \cos\theta$. Therefore, $\cos\theta = \frac{32}{(\sqrt{14})(\sqrt{77})} = \frac{32}{7\sqrt{22}}$, so $\theta = \arccos\left(\frac{32}{7\sqrt{22}}\right)$.

VECTOR DOT PRODUCT

- Summation of element wise scalar multiplication.
- The output is going to be a scalar that denotes the angle between the two vectors.
- If dot-product = 0 the two vectors are orthogonal.
- Fact 6.3 to isolate for cos(theta) and take the inverse to find the angle.

$$\cos\theta = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\| \|\vec{b}\|}$$

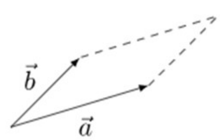
VECTOR CROSS-PRODUCT.

- The output is a new vector orthogonal to both the previous vectors.
- u x v = -v x u (Order matters)

$$\mathbf{u} \times \mathbf{v} = \left[u_2 v_3 - u_3 v_2, u_3 v_1 - u_1 v_3, u_1 v_2 - u_2 v_1\right],$$

- Fun fact. (Length of the resulting vector)

$\|\vec{a} \times \vec{b}\| = \|\vec{a}\|\|\vec{b}\| \sin\theta$, where θ is the angle between $\vec{a}$ and $\vec{b}$. This fact has a nice geometric interpretation: if you draw a parallelogram with $\vec{a}$ and $\vec{b}$ being two of its sides (like in the picture below), then $\|\vec{a} \times \vec{b}\|$ is the area of this parallelogram.



VECTOR ROWS AND COLUMNS (AS MATRIX)

- **Dot-product treated as matrix multiplication.**
- Row Vector **x** Column Matrix (Inner product)
- Columns Matrix **x** Row Matrix (Outer product)

1. $c = ab^T$ (only if a and b have same number of elements)

$$c = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}_{1\times3} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}_{3\times1} = \begin{bmatrix} a_1b_1 + a_2b_2 + a_3b_3 \end{bmatrix}_{1\times1}$$

2. $d = a^T b$

$$d = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}_{3\times1} \begin{bmatrix} b_1 & b_2 & b_3 \end{bmatrix}_{1\times3} = \begin{bmatrix} a_1b_1 & a_1b_2 & a_1b_3 \\ a_2b_1 & a_2b_2 & a_2b_3 \\ a_3b_1 & a_3b_2 & a_3b_3 \end{bmatrix}_{3\times3}$$

MATRIX

- Matrix is a 2D tensor but can have any number of rows and columns.
  - Shape of a matrix) *rows x columns.*

$$\underset{2\times 2}{\mathbf{X}} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

- Order as well as repeat occurrences matters. (unlike sets).

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & \cdots & x_{1(p-1)} & x_{1p} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2(p-1)} & x_{2p} \\ \vdots & \vdots & \ddots & & & \vdots \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ x_{(n-1)1} & x_{(n-1)2} & \cdots & \cdots & x_{(n-1)(p-1)} & x_{(n-1)p} \\ x_{n1} & x_{n2} & \cdots & \cdots & x_{n(p-1)} & x_{np} \end{bmatrix}$$

MATRIX (ADDITION/SUBRATCTION) AND SCALAR.

- Matrix addition and subtraction are performed only for two conformable matrices. (same amount rows and columns).

3.16: Matrix Addition.

$$\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{Y} = \begin{bmatrix} -2 & 2 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{X} + \mathbf{Y} = \begin{bmatrix} 1-2 & 2+2 \\ 3+0 & 4+1 \end{bmatrix} = \begin{bmatrix} -1 & 4 \\ 3 & 5 \end{bmatrix}.$$

- Scalar multiplication and division works the same way as for vectors (element wise).

.18: Scalar Multiplication.

$$\mathbf{X} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad s = 5$$

$$s \times \mathbf{X} = \begin{bmatrix} 5\times 1 & 5\times 2 \\ 5\times 3 & 5\times 4 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 15 & 20 \end{bmatrix}$$

MATRIX MULTIPLICATION

- Matrix compatibility for multiplication.
  - Numbers of Rows in the first has to match the number of columns in the second.
  - Our case $\underset{(k\times n)}{\mathbf{X}}\underset{(n\times p)}{\mathbf{Y}}$ X pre-multiplies Y or Y post-multiplies X (Order matters).
  - Dimensionality of the output matrix. (K x P)

$$\begin{bmatrix} 2 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \\ 5 \end{bmatrix}$$

- Compatible because (2x3 x 3x1) Inner numbers match
- Output: 2x1 (Outer numbers).

HOW DOES THE MAGIC WORK?

- The reason why $\underset{(k\times n)}{\mathbf{X}} \underset{(n\times p)}{\mathbf{Y}}$ has to be true.

What is the prerequisite for a dot-product?

- Want to find: $$\mathrm{XY} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$$
- First row in the X matrix now a vector [x11, x12] dot-product first column in Y matrix [y11, y21].
  - XY[1,1] = x11*y11 + x12 * y 21
- First row X second Columns in Y. [x11, x12] [y12, y22]

  XY[1,2] = x11 * y12 + x12 * y22.

Continue for every combination…

ALL OF THIS IS THE SAME AS SAYING.

$$\underset{(k\times n)}{\mathbf{X}}\underset{(n\times p)}{\mathbf{Y}} = \begin{bmatrix} \sum_{i=1}^{n} x_{1i}y_{i1} & \sum_{i=1}^{n} x_{1i}y_{i2} & \cdots & \sum_{i=1}^{n} x_{1i}y_{ip} \\ \sum_{i=1}^{n} x_{2i}y_{i1} & \sum_{i=1}^{n} x_{2i}y_{i2} & \cdots & \sum_{i=1}^{n} x_{2i}y_{ip} \\ \vdots & & \ddots & \vdots \\ \sum_{i=1}^{n} x_{ki}y_{i1} & \cdots & \cdots & \sum_{i=1}^{n} x_{ki}y_{ip} \end{bmatrix}$$

EXERCISES .

- 3.1
- 3.7 (Look up identity matrix in the chapter)
- 3.10
- 3.11
- 3.13
- 3.22
- Make up versions of the following matrices and explain why they are called so:
  - Square matrix
  - Symmetric matrix
  - Skew-symmetric matrix
  - Diagonal Matrix
  - Identity Matrix
  - J matrix
  - 0 matrix.
- Exercises in the pdf I uploaded for today.

- Matrixmultiplcation.xyz