② Data Structures for Linear Algebra

28.04.2025

What is a tensor?

* Tensor: ML generalization of vectors and matrices to any number of dimensions.

Dimensions	Mathematical Name	Description
0	scalar	magnitude-only
1	vector	array
2	matrix	flat table (e.g. square)
3	3-tensor	3D table (e.g. cube)
n	n-tensor	higher dimensional

What is a scalar?

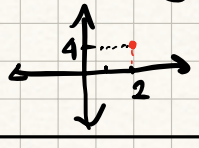
* A scalar:
  * Has no dimensions
  * Is a single number
  * Denoted in lowercase, italics (e.g. x)
  * Should be typed

* 0201_data_structures_for_algebra.py

What is a vector?

* Vectors are one-dimensional array of numbers.
* Denoted in lowercase, italic, bold.
* Arranged in an order, so element can be accessed by its index. (Elements are scalars, so not bold).
* Representing a point in space. $[x_1\ x_2] = [4\ 2]$

⇒



What is a vector transposition?

* Vector transposition:

$$v = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \Rightarrow v^T = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

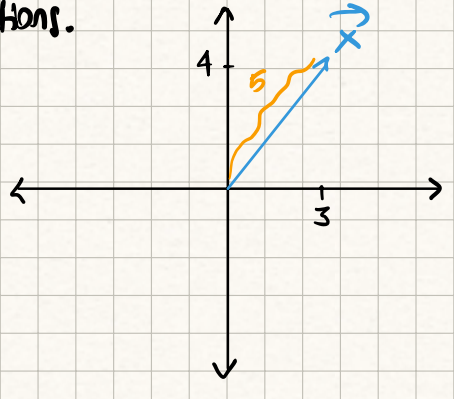
row vector
shape = (1, 3)
rows, columns

column vector
shape = (3, 1)

* 0201_data_structures_for_algebra.py

What is norm of a vector?

* Aside from representing a point, vectors can represent a magnitude and direction from origin to that point. The magnitude is called norm. Norms, particularly, L^1 and L^2, are used to regularize objective functions.



* $\vec{x} = [x_1 \; x_2] = [3 \; 4]$

* L^2 Norm $= \|\vec{x}\|_2 = \sqrt{\sum_i x_i^2}$

* L^2 Norm measures simple (Euclidean) distance from origin. It's the most common norm in ML. So that most of the time it's notated as $\|\vec{x}\|$, instead of $\|x\|_2$

* $\|x\| = \sqrt{3^2 + 4^2} = \sqrt{25} = 5$

* 0201....py

What is a unit vector?

* Special case of vector where its L2 norm is equal to one.
  If $||x|| = 1$, x is "unit vector"

What other norms are there?

* L^1 Norm $= ||x||_1 = \sum_i |x_i|$
  * Another common norm in ML.
  * Varies linearly at all locations whether near or far from the origin.
  * Used whenever difference between zero and non-zero is key.
  * 0201...
* Squared L^2 Norm $= ||x||_2^2 = \sum_i x_i^2$
  * Computationally cheaper than L^2 norm because:
    * Squared L^2 norm equals simply $X^T \cdot X$.
    * Derivative of element x requires that element alone, where as L^2 norm requires x vector.
  * Downside is it grows slowly near origin so it can't be used if distinguishing between zero and near-zero is important.
  * 0201...
* Max Norm (or L^∞ Norm) $= ||x||_\infty = \max_i |x_i|$
  * Returns the absolute value of the largest magnitude element.
  * 0201...

What are basis vectors?

- Basis vectors can be scaled to represent any vector in a given vector space.
- Typically use unit vectors along axes of vector space:
  $i = (1,0)$, $j = (0,1)$ $\Rightarrow$ $\vec{v} = (5,2) = 5 \cdot i + 2 \cdot j$

What are orthogonal vectors?

- $\vec{x}$ and $\vec{y}$ are orthogonal ($\perp$) if $x^T \cdot y = 0$ (and vice versa).
- This means they are at 90° angle to each other (assuming non-zero norms.
- n-dimensional space has max. n mutually orthogonal vectors.

What are orthonormal vectors?

- Orthonormal vectors are orthogonal vectors that have unit norm. (e.g. standard basis vectors)
- 0201...

What are matrices?

- Two-dimensional array of numbers.
- Denoted in uppercase, italics, bold, e.g: X
- Rows x Columns in notation: $X_{2\times3}$, $X_{2,3}$
- Colon represents an entire row or column:
  $X_{:,1}$ = Left column $X_{1,:}$ = Top row
- 0201...

What are n-tensors?

* These are multi-dimensional tables.
* $X_{i,j,k,l}$
* For example, rank 4 tensors are common for images, where each dimension corresponds to:
  1. Number of images in training batch.
  2. Image height in pixels
  3. Image width in pixels
  4. Number of color channels (e.g. 3 for RGB)
* 0201...

Quiz

1. What is the transpose of $X = \begin{bmatrix} 25 \\ 2 \\ -3 \\ -23 \end{bmatrix}$?

$$x^T = [25 \ 2 \ -3 \ -23]$$

2. Using algebraic notation, what are the dimensions of $Y = \begin{bmatrix} 42 & 4 & 7 & 99 \\ -99 & -3 & 17 & 22 \end{bmatrix}$? $= Y_{2x4}$

3. What is the position of the element in Y with the value of 17.

$$17 = Y_{2,3}$$