Arrays: computing with many numbers

Some perspective

- We have so far (mostly) looked at what we can do with single numbers (and functions that return single numbers).
- Things can get much more interesting once we allow not just one, but many numbers together.
- It is natural to view an array of numbers as one object with its own rules.
- The simplest such set of rules is that of a **vector.**

$\underset{\sim}{u}$ $\underset{\approx}{A}$

Demo "Sound as Vectors"

Vectors

A vector is an element of a Vector Space

$$n\text{-vector:} \qquad \boldsymbol{x} = \begin{Bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{Bmatrix} = [x_1 \quad x_2 \cdots x_n]^T$$

Vector space $\mathcal{V}$:

A vector space is a set $\mathcal{V}$ of vectors and a field $\mathcal{F}$ of scalars with two operations:

1) addition: $u + v \in \mathcal{V}$, and $u, v \in \mathcal{V}$

2) multiplication : $\alpha \cdot u \in \mathcal{V}$, and $u \in \mathcal{V}$, $\alpha \in \mathcal{F}$

Vector Space

The addition and multiplication operations must satisfy:

(for $\alpha, \beta \in \mathcal{F}$ and $u, v \in \mathcal{V}$)

Associativity: $u + (v + w) = (u + v) + w$

Commutativity: $u + v = v + u$

Additive identity: $v + 0 = v$

Additive inverse: $v + (-v) = 0$

Associativity wrt scalar multiplication: $\alpha \cdot (\beta \cdot v) = (\alpha \cdot \beta) \cdot v$

Distributive wrt scalar addition: $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$

Distributive wrt vector addition: $\alpha \cdot (u + v) = \alpha \cdot u + \alpha \cdot v$

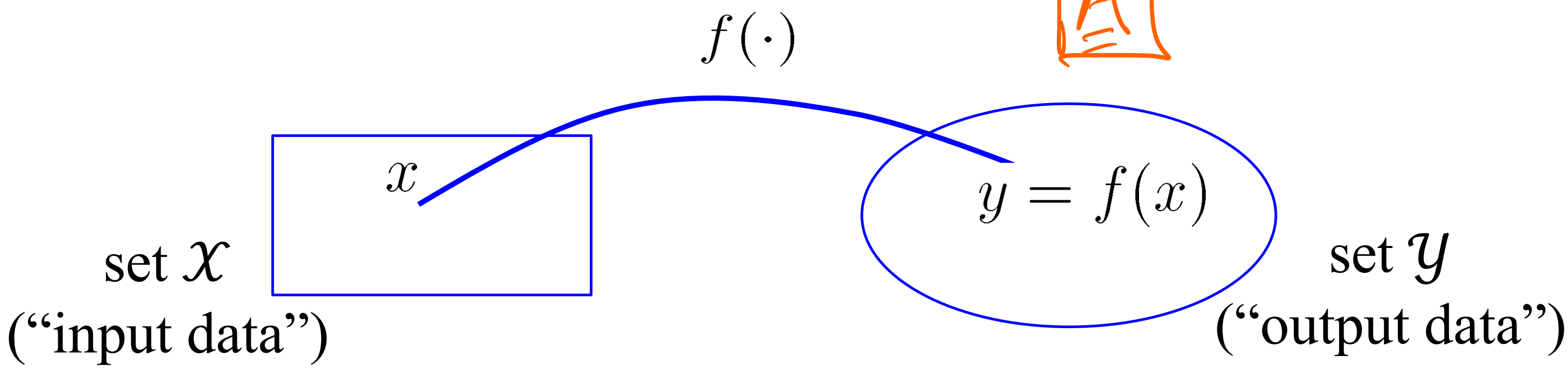
Scalar multiplication identity: $1 \cdot (u) = u$

Linear Functions

$\underset{\sim}{x} \xrightarrow{f(x)} \underset{\sim}{y}$

$\underset{=}{A}$

Function: $f : \mathcal{X} \to \mathcal{Y}$



The function f takes vectors $\boldsymbol{x} \in \mathcal{X}$ and transforms into vectors $\boldsymbol{y} \in \mathcal{Y}$

A function f is a linear function if

$$\left\{\begin{array}{l} (1)\ f(\boldsymbol{u}+\boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v}) \\ (2)\ f(a\boldsymbol{u}) = a f(\boldsymbol{u}) \text{ for any scalar } a \end{array}\right.$$

$\underset{=}{A}\,\underset{\sim}{x} = \underset{\sim}{y}$

Iclicker question

$\underset{\sim}{u}, \underset{\sim}{v}$

$$f(u+v) = \frac{|u+v|}{u+v} \neq \frac{|u|}{u} + \frac{|v|}{v}$$

1) Is

$$f(x) = \frac{|x|}{x}, \; f: \mathcal{R} \to \mathcal{R}$$

a linear function?

A) YES

B) NO

2) Is

$$f(x) = a\,x + b, \; f: \mathcal{R} \to \mathcal{R}, \;\; a, b \in \mathcal{R} \text{ and } a, b \neq 0$$

a linear function? $\underset{\sim}{u}, \underset{\sim}{v}$

$$f(\underset{\sim}{u}+\underset{\sim}{v}) = a(\underset{\sim}{u}+\underset{\sim}{v}) + b$$

$$= a\underset{\sim}{u} + a\underset{\sim}{v} + b$$

A) YES

B) NO

$$\left.\begin{aligned} f(\underset{\sim}{u}) &= a\underset{\sim}{u} + b \\ f(\underset{\sim}{v}) &= a\underset{\sim}{v} + b \end{aligned}\right\} a\underset{\sim}{u} + a\underset{\sim}{v} + 2b$$

$\neq$

Matrices

- $n \times m$-matrix

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{bmatrix}$$

- Linear functions $f(\boldsymbol{x})$ can be represented by a Matrix-Vector multiplication.

- Think of a matrix $\boldsymbol{A}$ as a linear function that takes vectors $\boldsymbol{x}$ and transforms them into vectors $\boldsymbol{y}$

$$\boldsymbol{y} = f(\boldsymbol{x}) \rightarrow \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$$

- Hence we have:

$$\begin{cases} \boldsymbol{A}(\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{A}\,\boldsymbol{u} + \boldsymbol{A}\,\boldsymbol{v} \\ \boldsymbol{A}(\alpha\,\boldsymbol{u}) = \alpha\,\boldsymbol{A}\,\boldsymbol{u} \end{cases}$$

Matrix-Vector multiplication

$n\times 1$ $n\times m$ $m\times 1$

- Recall summation notation for matrix-vector multiplication $\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{x}$
- You can think about matrix-vector multiplication as:

$$y_i = \sum_{j=1}^{m} A_{ij} x_j$$

2 Linear combination of column vectors of **A**

$$\boldsymbol{y} = x_1\,\mathbf{A}[:,1] + x_2\,\mathbf{A}[:,2] + \cdots + x_m\,\mathbf{A}[:,m]$$

1 Dot product of $\boldsymbol{x}$ with rows of **A**

$$\boldsymbol{y} = \begin{pmatrix} \mathbf{A}[1,:]\cdot\boldsymbol{x} \\ \vdots \\ \mathbf{A}[n,:]\cdot\boldsymbol{x} \end{pmatrix}$$

$$y_i = A_{i,:} \cdot \underset{\sim}{x}$$

$A[i,:]$

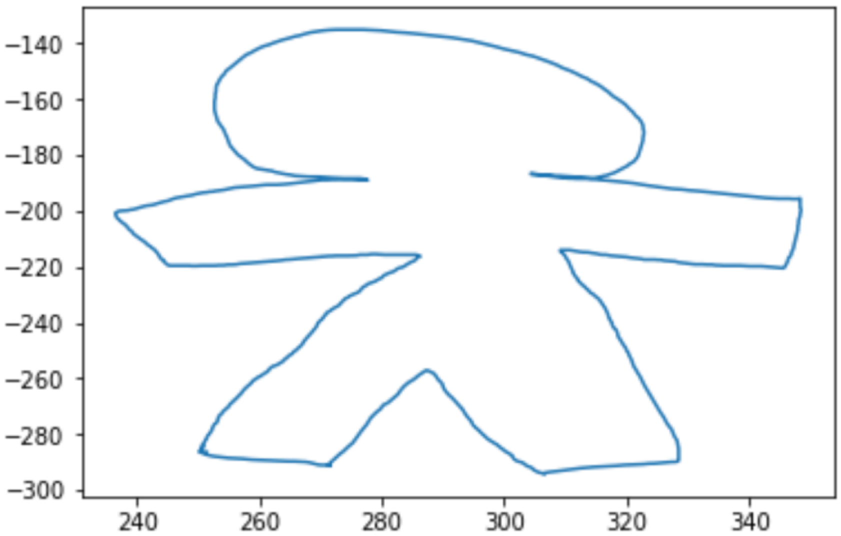
$$\begin{array}{cc} & n\times m \\ \mathbf{A} = & \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{bmatrix} \end{array} \begin{array}{c} m\times 1 \\ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \end{array} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\begin{bmatrix} & & \\ & A_{n\times m} & \\ & & \end{bmatrix} \begin{bmatrix} \\ x_{m\times 1} \\ \\ \end{bmatrix} = \begin{bmatrix} \\ y_{n\times 1} \\ \\ \end{bmatrix}$$

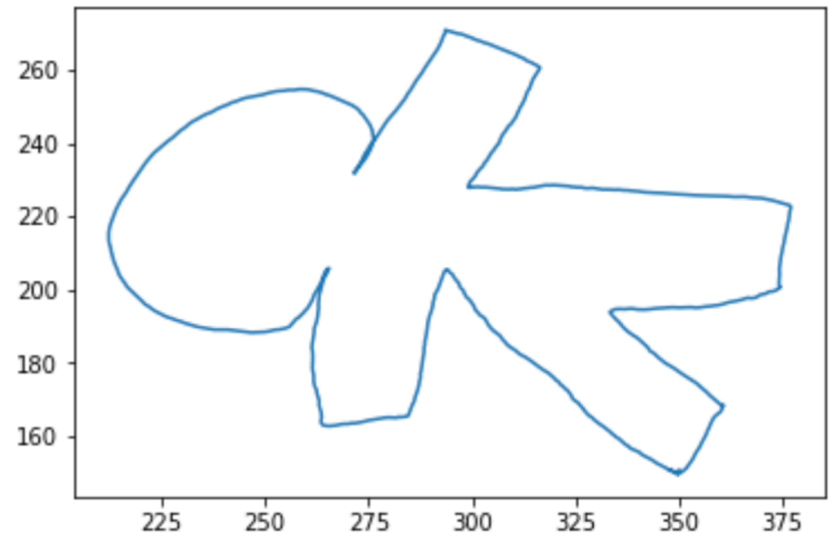
$$y_i = \sum_{j=1}^{m} A_{ij} x_j$$

$$= x_1 \begin{bmatrix} \\ A_{:,1} \\ \\ \end{bmatrix} + x_2 \begin{bmatrix} \\ A_{:,2} \\ \\ \end{bmatrix} + \cdots + x_m \begin{bmatrix} \\ A_{:,m} \\ \\ \end{bmatrix}$$

Matrices operating on data



Data set: $\boldsymbol{x}$



Data set: $\boldsymbol{y}$

Rotation

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x})$$

or

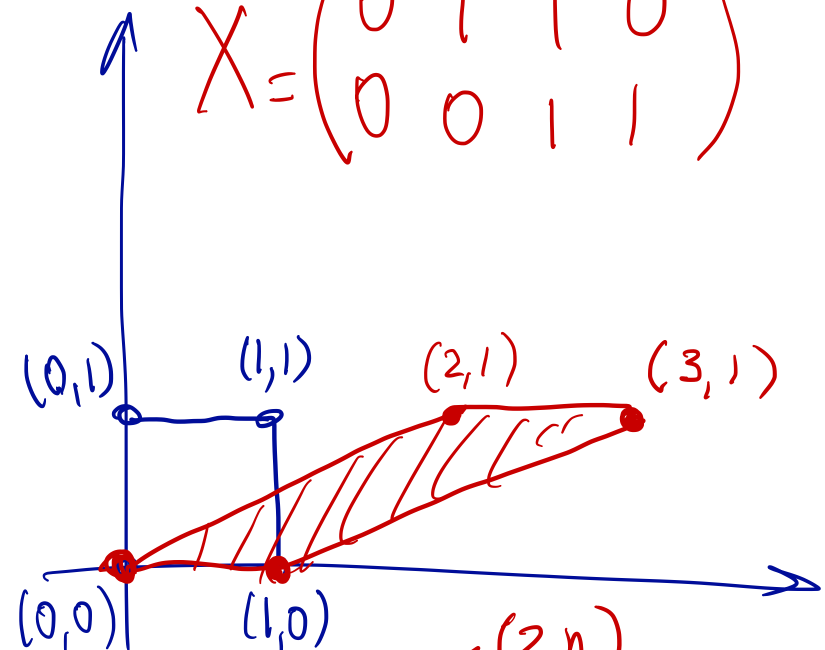
$$\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{x}$$

Example: Shear operator

Matrix-vector multiplication for each vector (point representation in 2D):

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

$$X = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$



$$\underset{\sim}{p_1} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\underset{\sim}{p_2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\underset{\sim}{p_3} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

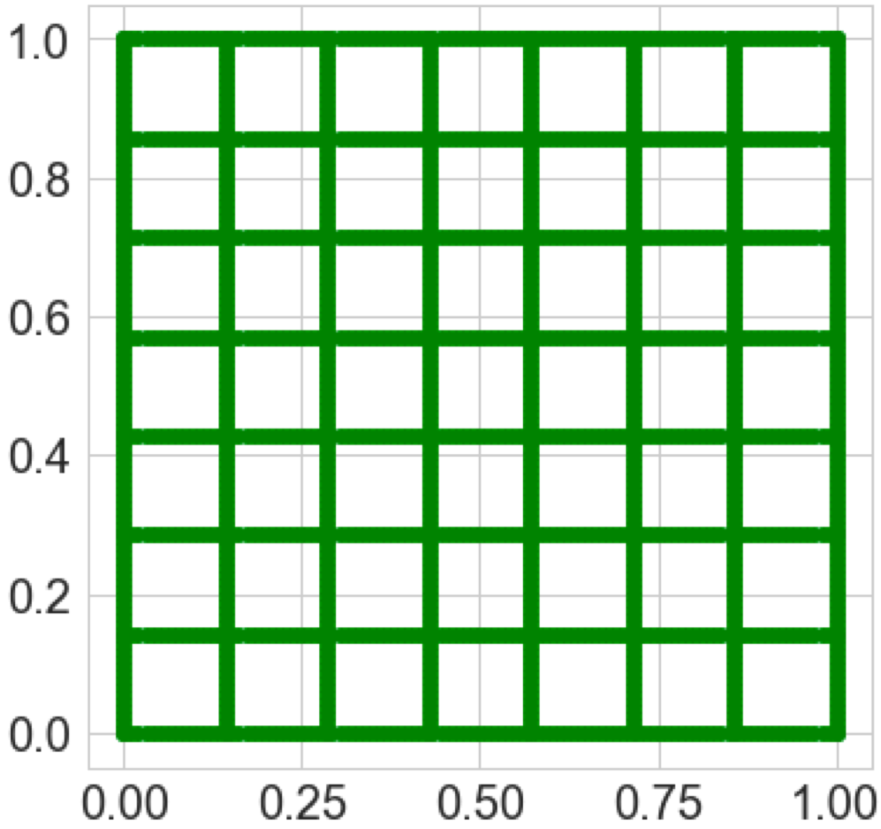
$$\underset{\sim}{p_4} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

(2, n)

$$\underset{\substack{= \\ 2\times 2}}{A}\ \underset{\substack{\sim \\ 2\times n}}{X} = \underset{\substack{= \\ 2\times n}}{Y}$$

Matrices as operators

- **Data**: grid of 2D points
- Transform the data using matrix multiply



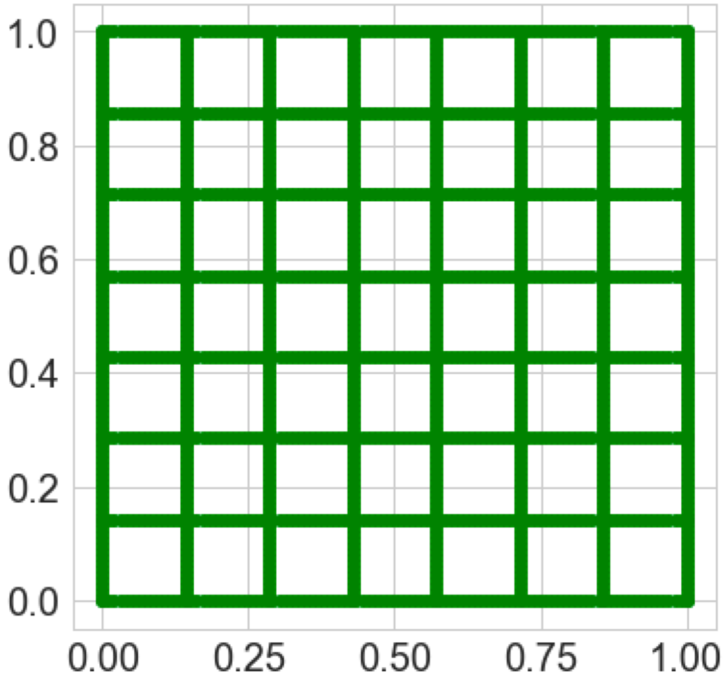
What can matrices do?

1. Shear
2. Rotate
3. Scale
4. Reflect

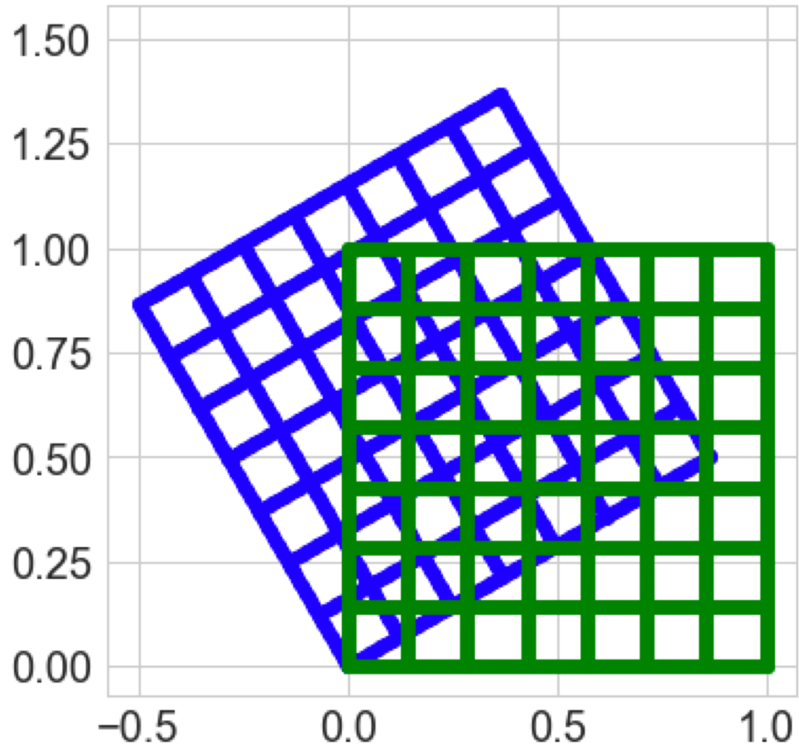
5. Can they translate?

Rotation operator

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

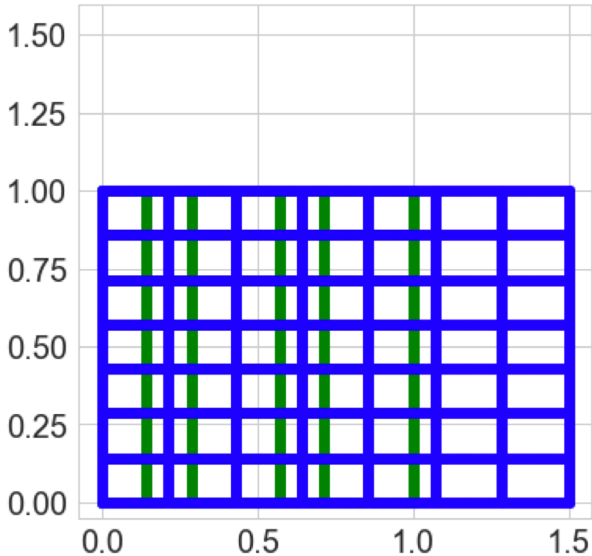


$\theta = \pi/6$



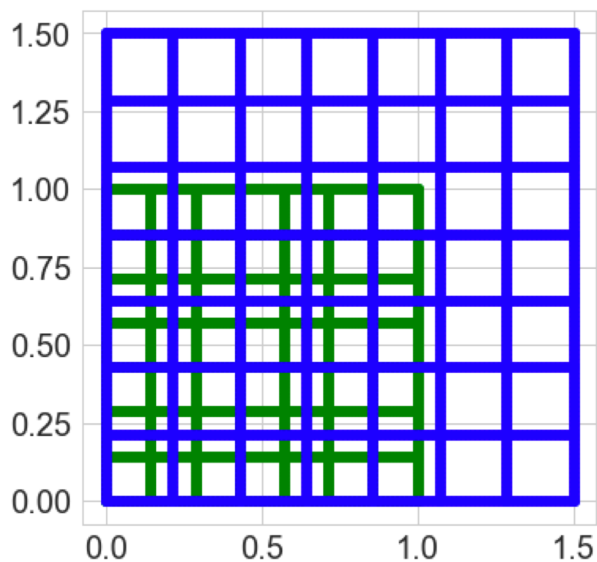
Scale operator

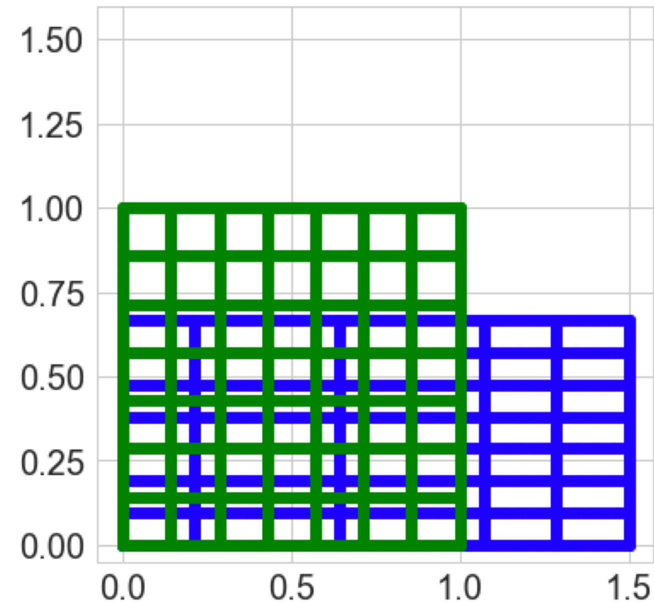
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$



$$\begin{pmatrix} 3/2 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 3/2 & 0 \\ 0 & 3/2 \end{pmatrix}$$



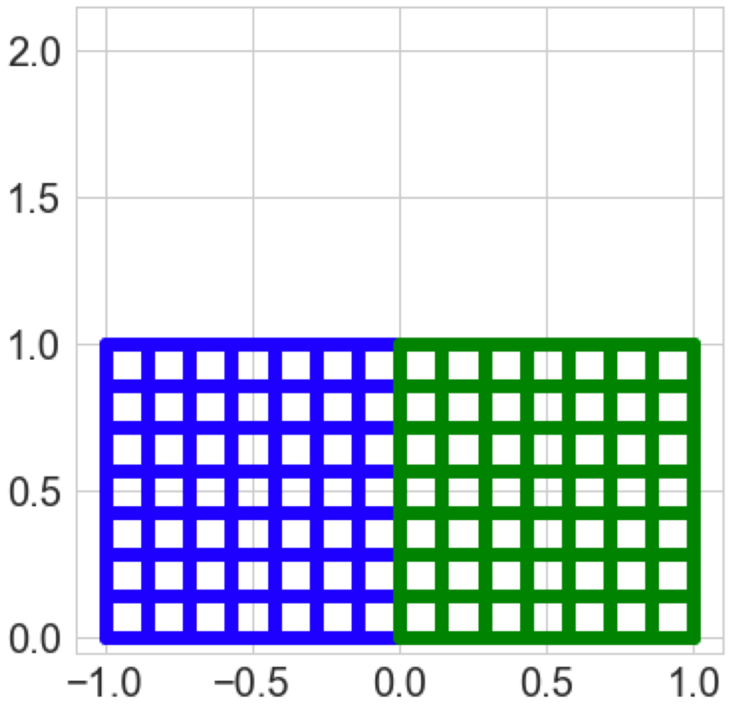


$$\begin{pmatrix} 3/2 & 0 \\ 0 & 2/3 \end{pmatrix}$$

Reflection operator

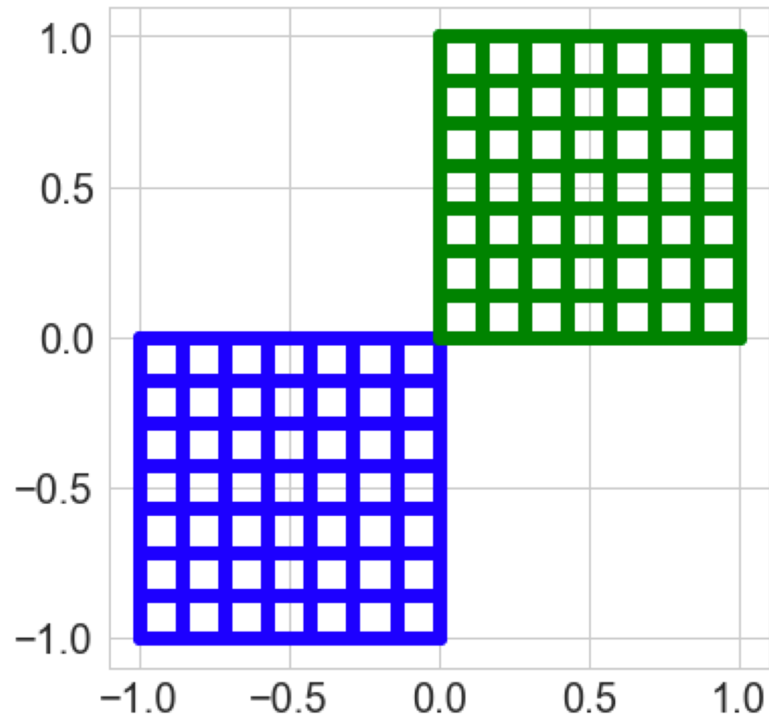
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$



Reflect about y-axis

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

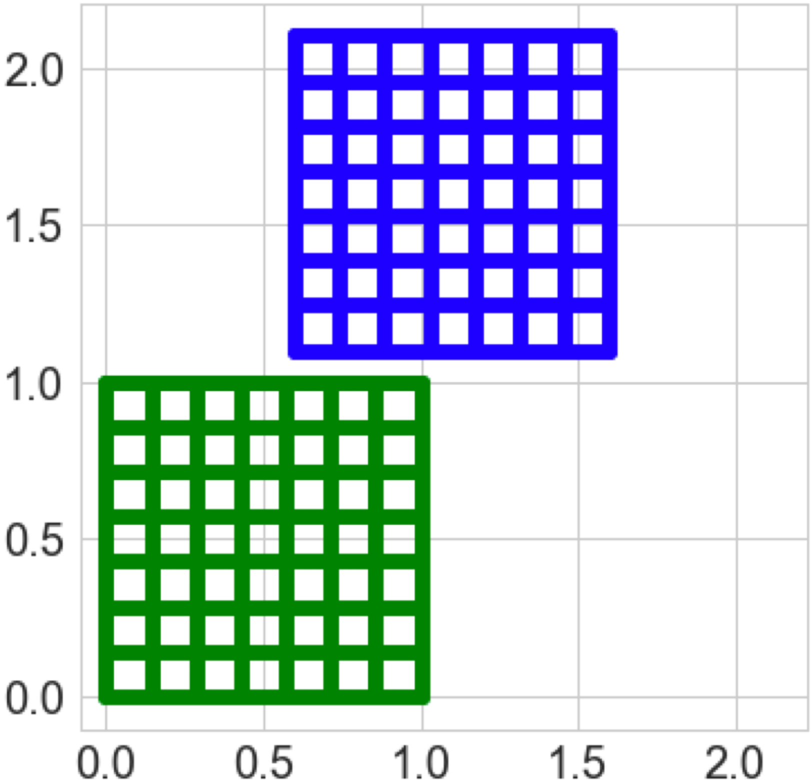


Reflect about x and y-axis

Translation (shift)

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$$

$a = 0.6; b = 1.1$

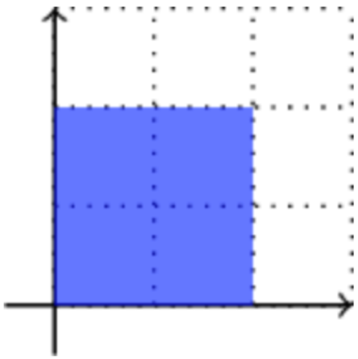


Iclicker question

Images of a brick

1 point

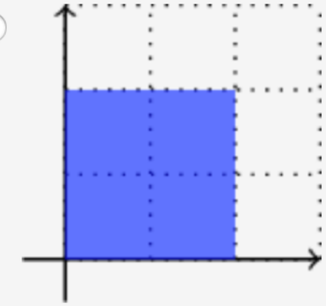
Consider the unit square in the plane:



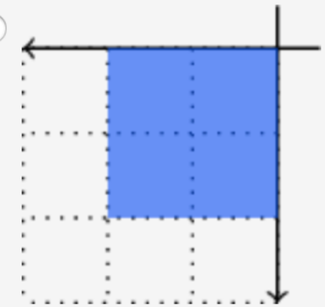
Suppose you take every vector $\mathbf{x}$ corresponding to a point in the unit square and compute $A\mathbf{x}$ for the given matrix A. Which set of points could you obtain?

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

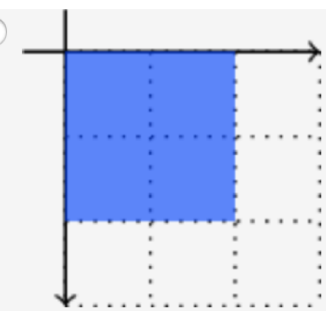
a)



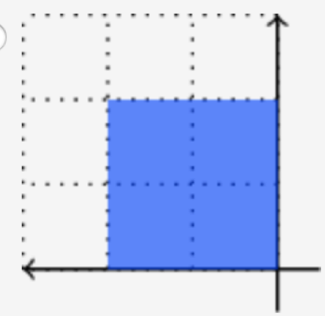
b)



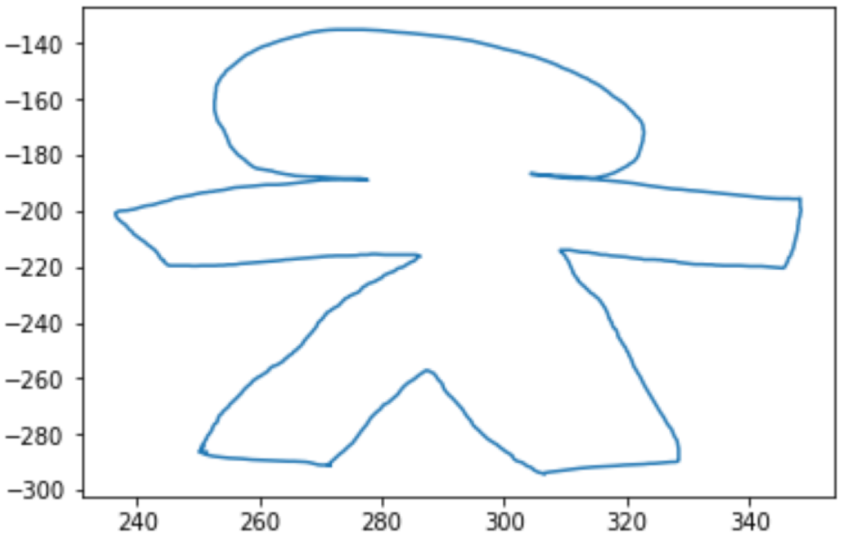
c)

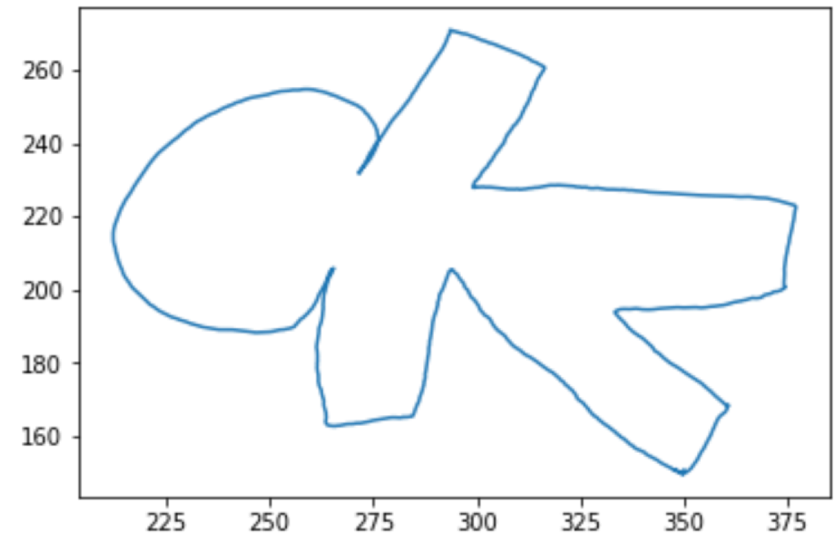


d)



Matrices operating on data





Data set: *A*

Data set: *B*

Rotation

Demo "Matrices for geometry transformation"

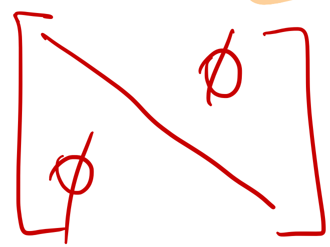
Notation and special matrices

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- Square matrix: $m = n$
- Zero matrix: $A_{ij} = 0$
- Identity matrix $[\boldsymbol{I}] = [\delta_{ij}]$
- Symmetric matrix: $A_{ij} = A_{ji} \quad [\boldsymbol{A}] = [\boldsymbol{A}]^T$
- Permutation matrix:
  - Permutation of the identity matrix
  - Permutes (swaps) rows

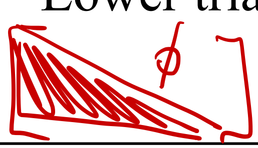
$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & \boxed{1} & 0 \end{pmatrix} \begin{pmatrix} a \\ \boxed{b} \\ c \end{pmatrix} = \begin{pmatrix} c \\ a \\ \boxed{b} \end{pmatrix}$$

- Diagonal matrix: $A_{ij} = 0, \ \forall i, j \mid i \neq j$



- Triangular matrix:

Lower triangular: $L_{ij} = \begin{cases} L_{ij}, i \geq j \\ 0, \quad i < j \end{cases}$



Upper triangular: $U_{ij} = \begin{cases} U_{ij}, i \leq j \\ 0, \quad i > j \end{cases}$

More about matrices

- Rank: the rank of a matrix $\boldsymbol{A}$ is the dimension of the vector space generated by its columns, which is equivalent to the number of linearly independent columns of the matrix.
- Suppose $\boldsymbol{A}$ has shape $m \times n$:
  - $rank(\boldsymbol{A}) \leq \min(m, n)$
  - Matrix $\boldsymbol{A}$ is **full rank**: $rank(\boldsymbol{A}) = \min(m, n)$. Otherwise, matrix $\boldsymbol{A}$ is **rank deficient**.
- Singular matrix: a square matrix $\boldsymbol{A}$ is invertible if there exists a square matrix **B** such that $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$. If the matrix is not invertible, it is called singular.

A^{-1} → does not exist

Iclicker question

What is the value of m that makes the matrix singular?

$$A = \begin{bmatrix} m & 2 \\ 9 & 6 \end{bmatrix}$$

A) 1

B) 3

C) 5

D) 7

$\det(A) = 0 \Rightarrow$ sig.

$m = 3$

Norms

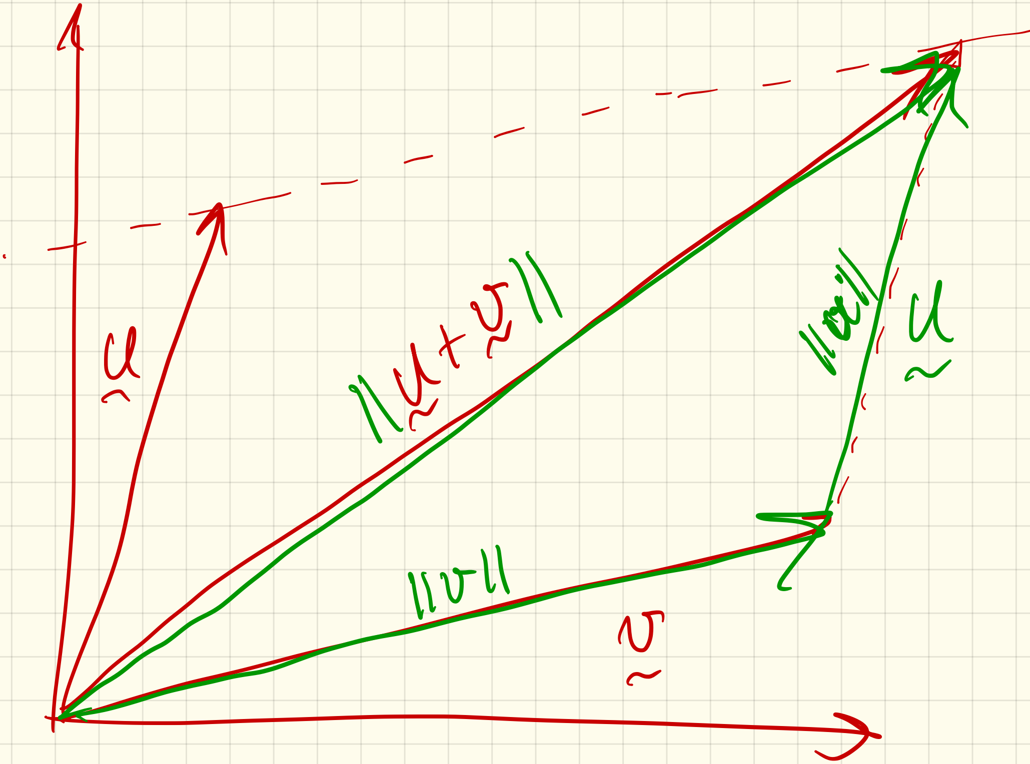
What's a norm?

- A generalization of 'absolute value' to vectors.
- $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}_0^+$, returns a 'magnitude' of the input vector
- In symbols: Often written $\|\mathbf{x}\|$.

Define norm.

A function $\|\mathbf{x}\| : \mathbb{R}^n \to \mathbb{R}_0^+$ is called a norm if and only if

1. $\|\mathbf{x}\| > 0 \Leftrightarrow \mathbf{x} \neq \mathbf{0}$.
2. $\|\gamma \mathbf{x}\| = |\gamma| \, \|\mathbf{x}\|$ for all scalars γ.
3. Obeys triangle inequality $\|\mathbf{x} + \mathbf{y}\| \leqslant \boxed{\|\mathbf{x}\| + \|\mathbf{y}\|}$


$\underset{\sim}{u}$
$\|\underset{\sim}{u}+\underset{\sim}{v}\|$
$\|u\|$
$\underset{\sim}{u}$
$\|v\|$
$\underset{\sim}{v}$

$$\|u+v\| \leq \|u\| + \|v\|$$

Example of Norms

What are some examples of norms?

The so-called *p-norms*:

$$\left\| \begin{pmatrix} x_1 \\ \\ x_n \end{pmatrix} \right\|_p = \left(\sqrt[p]{|x_1|^p + \cdots + |x_n|^p}\right)^{1/p} \quad (p \geqslant 1)$$

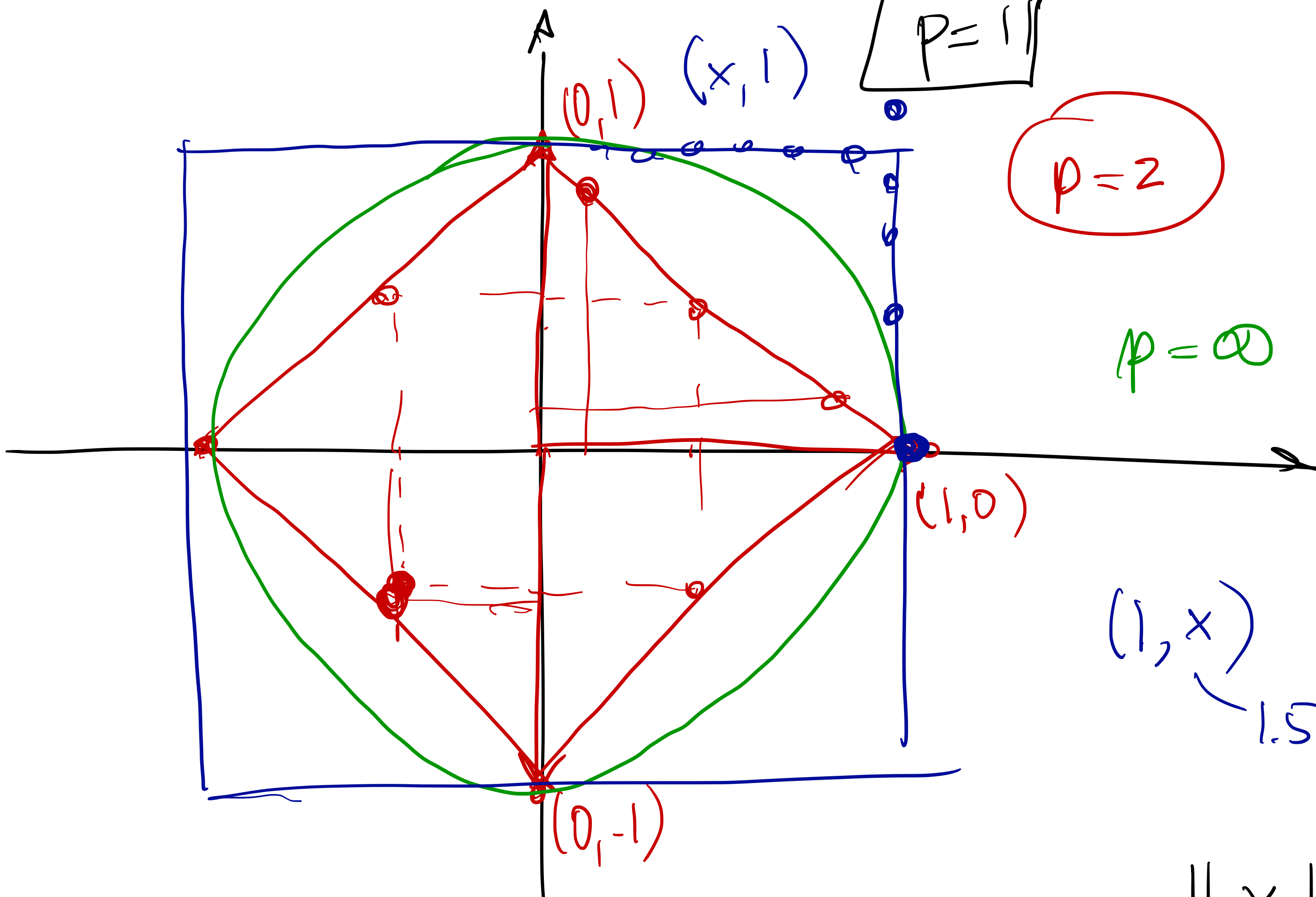
$p = 1, 2, \infty$ particularly important

$\|\underset{\sim}{x}\|_{p=1} = |x_1| + |x_2| + \ldots + |x_n| \longrightarrow$

$\|\underset{\sim}{x}\|_{p=2} = \sqrt{|x_1|^2 + |x_2|^2 + \ldots + |x_n|^2} \longrightarrow$ Euclidean norm

$\|x\|_{p=\infty} = \left(|x_1|^p + |x_2|^p + \ldots + |x_n|^p\right)^{1/p} = \max(|x_i|)$

Unit Ball: Set of vectors $\boldsymbol{x}$ with norm $\|\boldsymbol{x}\| = 1$



$\|x\|_p = 1$

~~$x = (2,1)$ $\|x\| = 3$~~

Demo “Vector Norms”

Norms and Errors

If we're computing a vector result, the error is a vector.
That's not a very useful answer to 'how big is the error'.
What can we do?

Apply a norm! $\left\| \underset{\sim}{x}_{exact} - \underset{\sim}{x}_{approx} \right\| = \|\underset{\sim}{e}_a\|$

How? Attempt 1:

~~Magnitude of error $\neq \|\text{true value}\| - \|\text{approximate value}\|$~~ **WRONG!**

Attempt 2:

$$\text{Magnitude of error} = \|\text{true value} - \text{approximate value}\|$$

$$\|e_r\| = \frac{\|\underset{\sim}{x}_{exact} - \underset{\sim}{x}_{approx}\|}{\|\underset{\sim}{x}_{exact}\|}$$

Absolute and Relative Errors

What are the absolute and relative errors in approximating the location of Siebel center $\underbrace{(40.114, -88.224)}_{\underset{\sim}{x}_{exact}}$ as $\underbrace{(40, -88)}_{\underset{\sim}{x}_{app}}$ using the 2-norm?

Absolute error:

a) *0.2240*

b) *0.3380*

c) *0.2513*

Relative error:

a) 2.59×10^{-3}

b) 2.81×10^{-3}

$$\|e_a\| = \|x_{ap} - x_{exact}\|$$

$$\|e_r\| = \frac{\|e_a\|}{\|x_{exact}\|}$$

Matrix Norms

What norms would we apply to matrices?

- Easy answer: '*Flatten*' matrix as vector, use vector norm. This corresponds to an entrywise matrix norm called the Frobenius norm,

$$\|A\|_F := \sqrt{\sum_{i,j} a_{ij}^2}.$$

Induced Norms

$$\|\underline{\underline{A}}\|_p = ? \qquad \boxed{\|A\|_p = \max_{\|\underset{\sim}{x}\|_p = 1} \|A\underset{\sim}{x}\|_p}$$

$$\left.\begin{array}{l}
\|\underbrace{\underline{\underline{A}}\,\underline{\underline{u_1}}}_{y_1}\|_p = \|y_1\|_p \checkmark \\
\|\underline{\underline{A}}\,\underline{u_2}\|_p = \|y_2\|_p \checkmark \\
\quad\vdots \\
\|A u_m\|_p = \|y_m\|_p \checkmark
\end{array}\right\} \max$$

$$\|u\| = 1$$

Matrix Norms

However, interpreting matrices as linear functions, what we are really interested in is the maximum amplification of the norm of any vector multiplied by the matrix,

$$\|A\| := \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| \, .$$

These are called induced matrix norms, as each is associated with a specific vector norm $\|\cdot\|$.

Matrix Norms

The following are equivalent:

$$\max_{\|x\|\neq 0} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \max_{\|x\|\neq 0} \left\| A \underbrace{\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}}_{y} \right\| \overset{\|y\|=1}{=} \max_{\|\boldsymbol{y}\|=1} \|A\boldsymbol{y}\| = \|A\| \,.$$

Logically, for each vector norm, we get a different matrix norm, so that, e.g. for the vector 2-norm $\|\boldsymbol{x}\|_2$ we get a matrix 2-norm $\|A\|_2$, and for the vector ∞-norm $\|\boldsymbol{x}\|_\infty$ we get a matrix ∞-norm $\|A\|_\infty$.

Induced Matrix Norms

SVD

$$\|\boldsymbol{A}\|_1 = \max_j \sum_{i=1}^{n} |A_{ij}|$$

Maximum absolute column sum of the matrix $\boldsymbol{A}$

note the absolute values here!

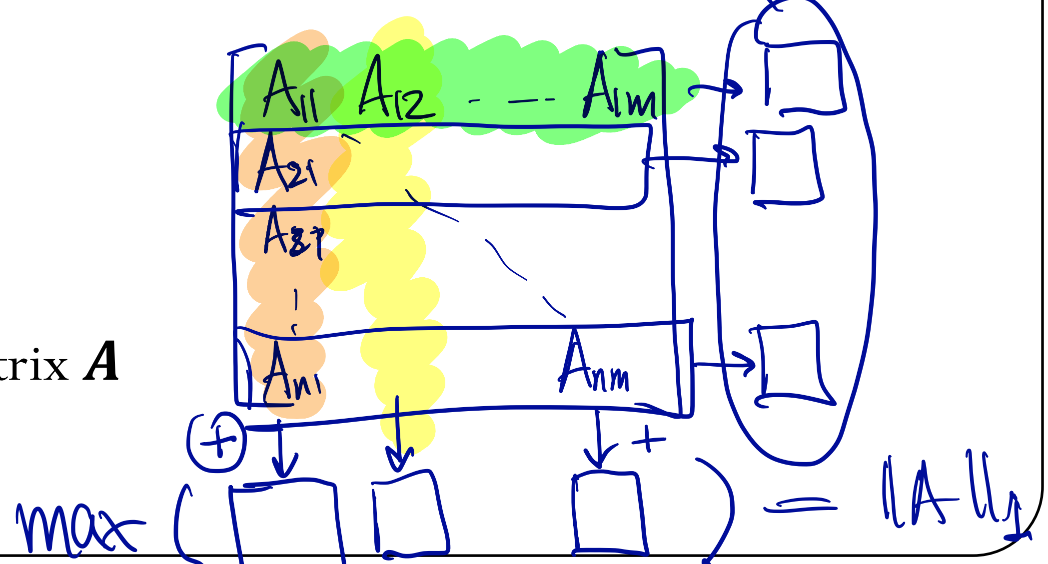
max $\|A\|_\infty$

$$\|\boldsymbol{A}\|_\infty = \max_i \sum_{j=1}^{n} |A_{ij}|$$

Maximum absolute row sum of the matrix $\boldsymbol{A}$

$$\boxed{\|\boldsymbol{A}\|_2 = \max_k \sigma_k}$$

~~σ_k are the singular value of the matrix $\boldsymbol{A}$~~



Properties of Matrix Norms

Matrix norms inherit the vector norm properties:

1. $||A|| > 0 \Leftrightarrow A \neq \mathbf{0}$.
2. $||\gamma A|| = |\gamma| \, ||A||$ for all scalars γ.
3. Obeys triangle inequality $||A + B|| \leqslant ||A|| + ||B||$

But also some more properties that stem from our definition:

1. $||A\mathbf{x}|| \leqslant ||A|| \, ||\mathbf{x}||$
2. $||AB|| \leqslant ||A|| \, ||B||$ (easy consequence)

Both of these are called submultiplicativity of the matrix norm.

Demo “Matrix Norms”

Iclicker question

Determine the norm of the following matrices:

1) $\left\|\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right\|_\infty \begin{array}{l} \to 3 \\ \to 7 \end{array}$

$\|A\|_\infty = 7$

2) $\left\|\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\right\|_1$

$\begin{array}{cc} 4 & 6 \end{array}$

$\|A\|_1 = 6$

a) 3

b) 4

c) 5

d) 6

e) 7

what if matrix was

$A = \begin{pmatrix} -1 & 2 \\ -3 & 4 \end{pmatrix}$?

Norm would be the same!

sum of absolute values $|A_{ij}|$

Iclicker question

Matrix Norm Approximation

Suppose you know that for a given matrix A three vectors $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$ for the vector norm $||\cdot||$,

$||\mathbf{x}|| = 2$, $||\mathbf{y}|| = 1$, $||\mathbf{z}|| = 3$,

and for corresponding induced matrix norm,

$||A\mathbf{x}|| = 20$, $||A\mathbf{y}|| = 5$, $||A\mathbf{z}|| = 90$.

What is the largest lower bound for $||A||$ that you can derive from these values?

a) 90

b) 30

c) 20

d) 10

e) 5

$$\max\left(\frac{||Ax||}{||x||}\right)$$

$$\left(\underbrace{\frac{20}{2}}_{10}, \underbrace{\frac{5}{1}}_{5}, \underbrace{\frac{90}{3}}_{30}\right)$$

Induced Matrix Norm of a Diagonal Matrix

What is the 2-norm-based matrix norm of the diagonal matrix

$$A = \begin{bmatrix} 100 & 0 & 0 \\ 0 & 13 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}?$$

SVD(A)

diagonal entries are the singular values

$SVD(A) \to 100, 13, 0.5$

$\|A\|_{p=2} = 100$

Induced Matrix Norm of an Inverted Diagonal Matrix

What is the 2-norm-based matrix norm of the **inverse** of the diagonal matrix

$$A = \begin{bmatrix} 100 & 0 & 0 \\ 0 & 13 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}?$$

$$A^{-1} = \begin{bmatrix} 1/100 & & \\ & 1/13 & \\ & & 1/0.5 \end{bmatrix}$$

$$\|A^{-1}\|_{p=2} = ? \quad 2 \quad \frac{1}{0.5}$$