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.**

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: $v + (-v) = 0$

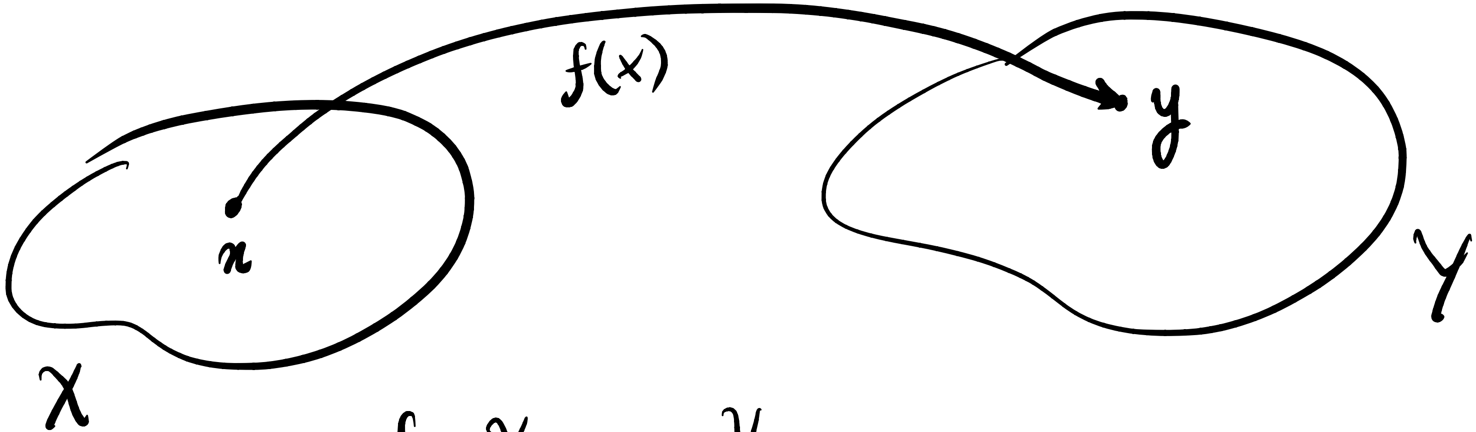
Associativity wrt scalar multiplication: $\alpha \cdot (\beta \cdot v) = (\alpha \cdot \beta) \cdot v$

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



$$f: \mathcal{X} \longrightarrow \mathcal{Y}$$

$y = f(x)$ f takes vectors $x \in \mathcal{X}$ and transforms into vectors $y \in \mathcal{Y}$

A function f is linear if:

$$f(u+v) = f(u) + f(v)$$

$$f(au) = a\,f(u)$$

Iclicker question

1) Is

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

a linear function?

A) YES

B) NO

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

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

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

2) Is

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

a linear function?

A) YES

B) NO

$f(u+v) = a(u+v)+b = au + av + b$ $\neq$

$f(u) = au+b$

$f(v) = av+b$

$f(u)+f(v) = a(u+v)+2b$

Matrices

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

- 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:

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

Matrix-Vector multiplication

- Recall summation notation for matrix-vector multiplication $\boldsymbol{y} = \boldsymbol{A}\,\boldsymbol{x}$

- You can think about matrix-vector multiplication as:

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

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} A[1,:] \cdot x_1 \\ A[2,:] \cdot x_2 \\ \vdots \\ A[n,:] \cdot x_n \end{bmatrix} = y$$

Matrix-Vector multiplication

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

Linear combination of column vectors of **A**

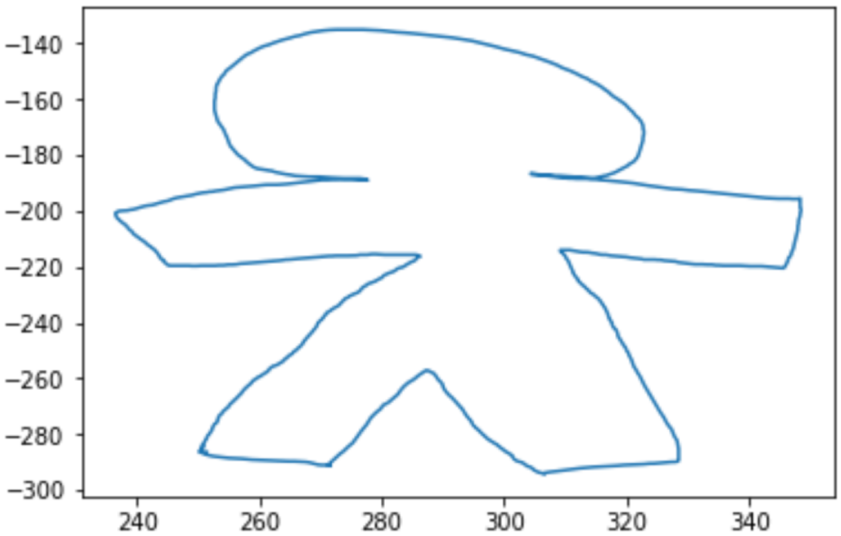
$$y_i = \sum_{j=1}^{m} A_{ij} x_j =$$

$$\underset{\sim}{y} = x_1 \underbrace{A[:,1]} + x_2 A[:,2] + \ldots + x_n A[:,n]$$

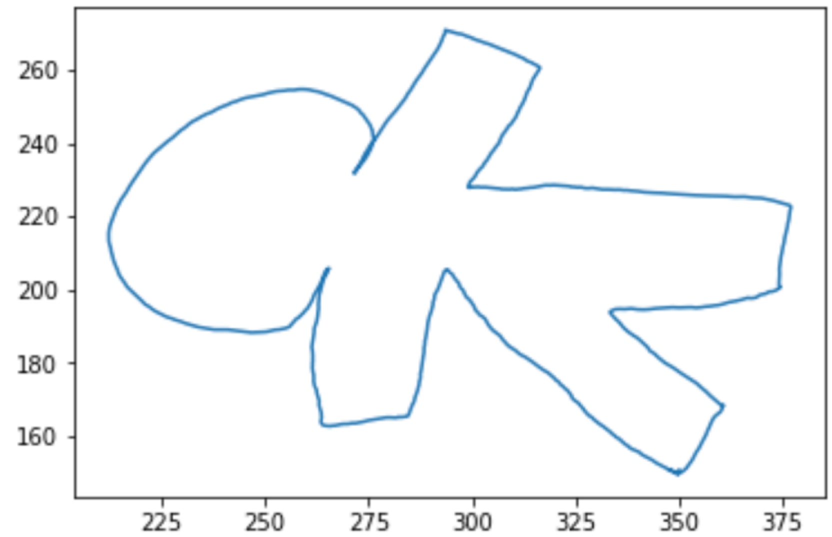
x_1 → scalar $A[:,1]$ → vector

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

Matrices operating on data



Data set: x



Data set: y

Rotation

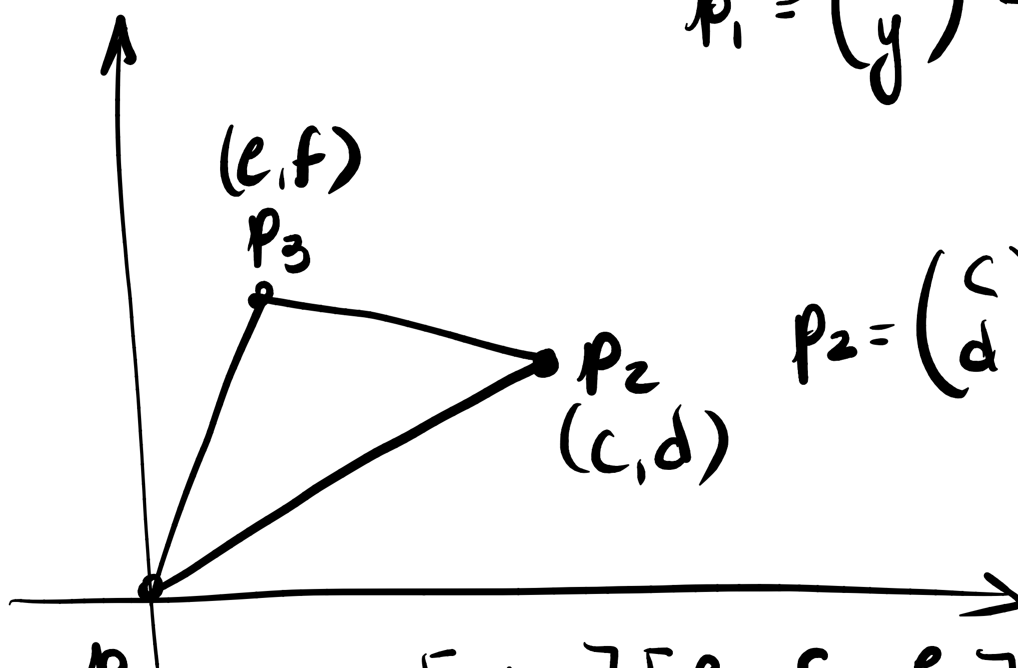
$$y = f(x)$$

or

$$y = A\,x$$

Example: Shear operator

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



$$p_1 = \begin{pmatrix} x \\ y \end{pmatrix} \longrightarrow A p_1 =$$

$$\begin{bmatrix} 2\times 2 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \text{new} \end{bmatrix}$$

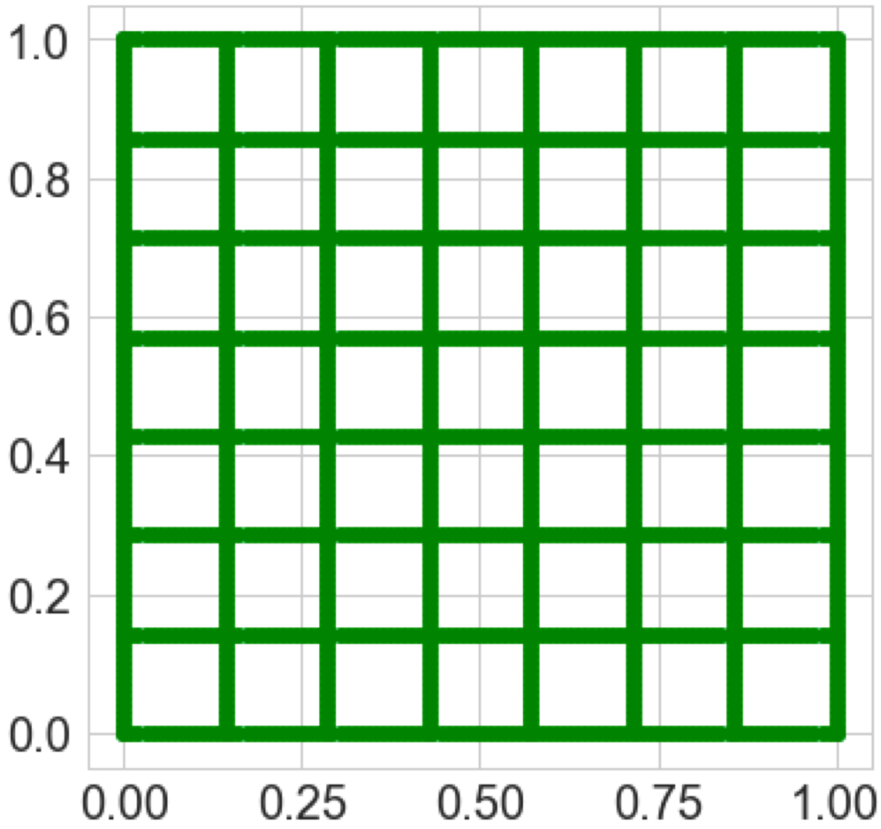
$$p_2 = \begin{pmatrix} c \\ d \end{pmatrix} \longrightarrow \begin{bmatrix} 2\times 2 \end{bmatrix}\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \text{new} \end{bmatrix}$$

$$\begin{bmatrix} \underset{=}{A} \end{bmatrix}\begin{bmatrix} a & c & e \\ b & d & f \end{bmatrix} = \begin{bmatrix} \text{new } p_1 & \text{new } p_2 & \text{new } p_3 \end{bmatrix}$$

2×2 $\downarrow \underset{\sim}{p_1}$ $\downarrow \underset{\sim}{p_2}$ $\downarrow \underset{\sim}{p_3}$

Matrices as operators

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



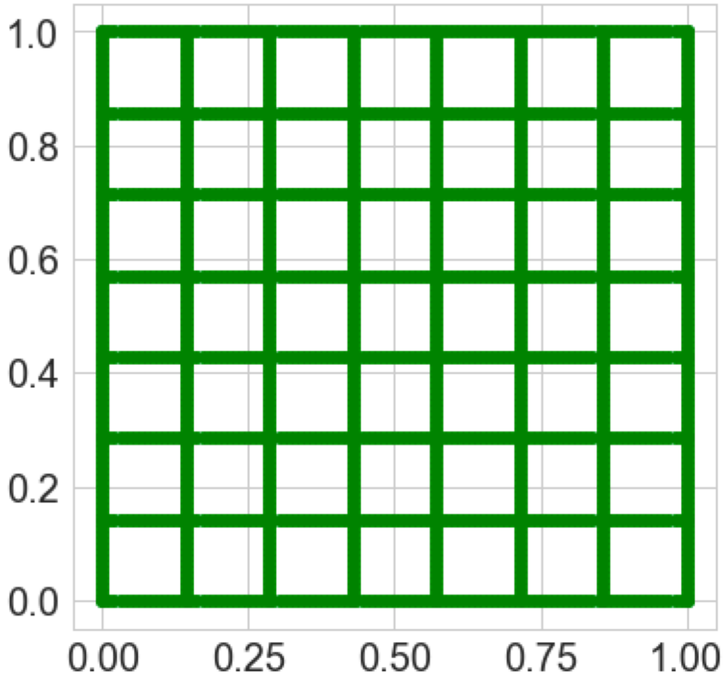
What can matrices do?

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

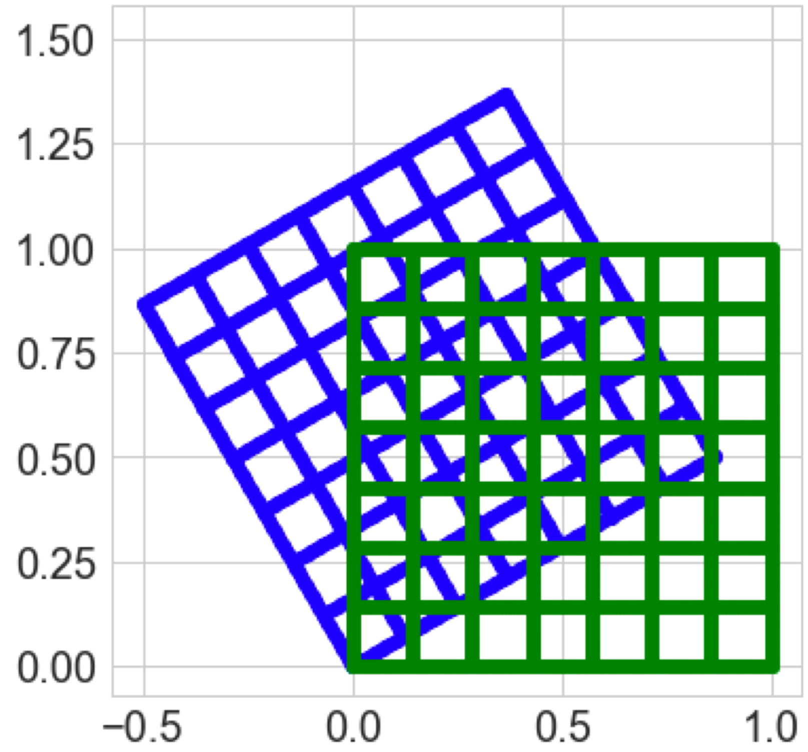
5. Can they translate?

Rotation operator

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



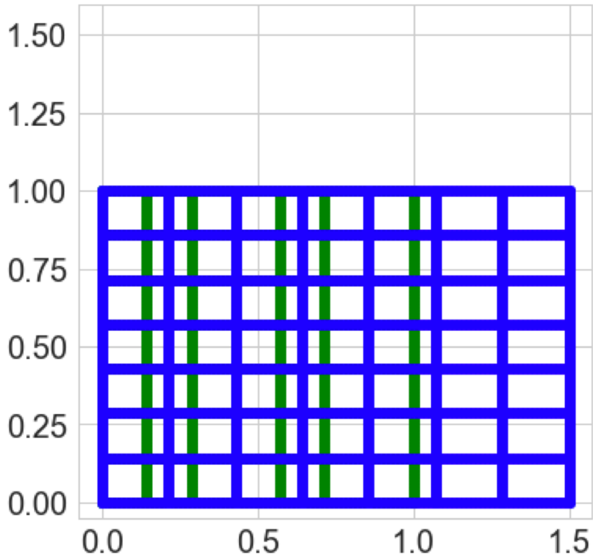
$\theta = \pi/6$



Demo "Matrix transformations"

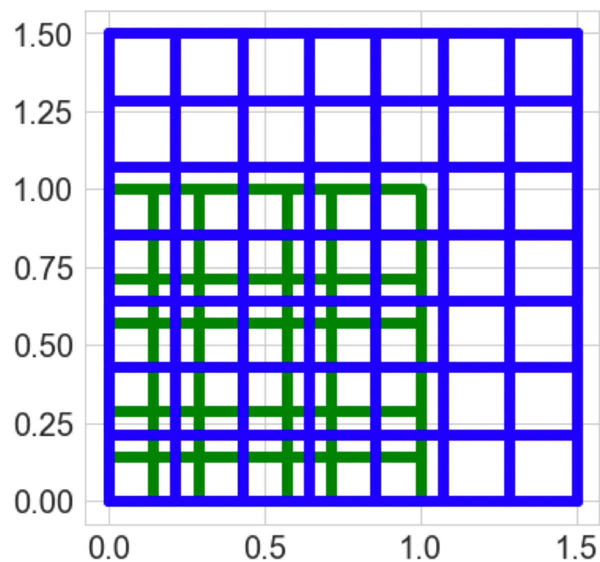
Scale operator

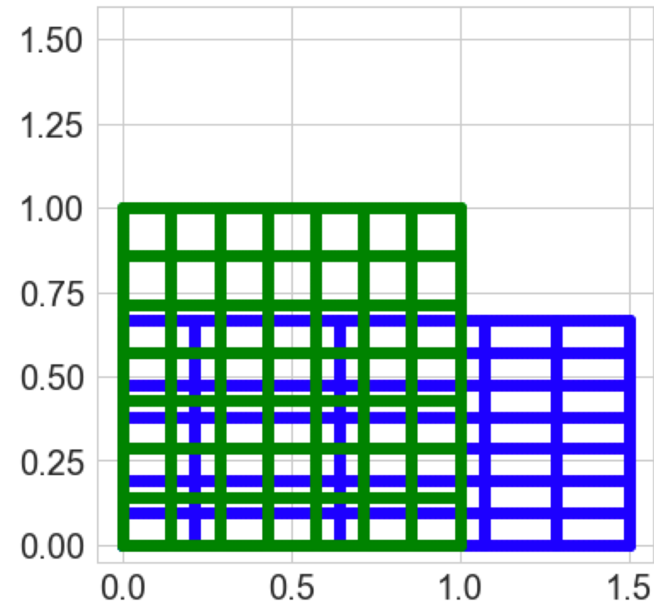
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$



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

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



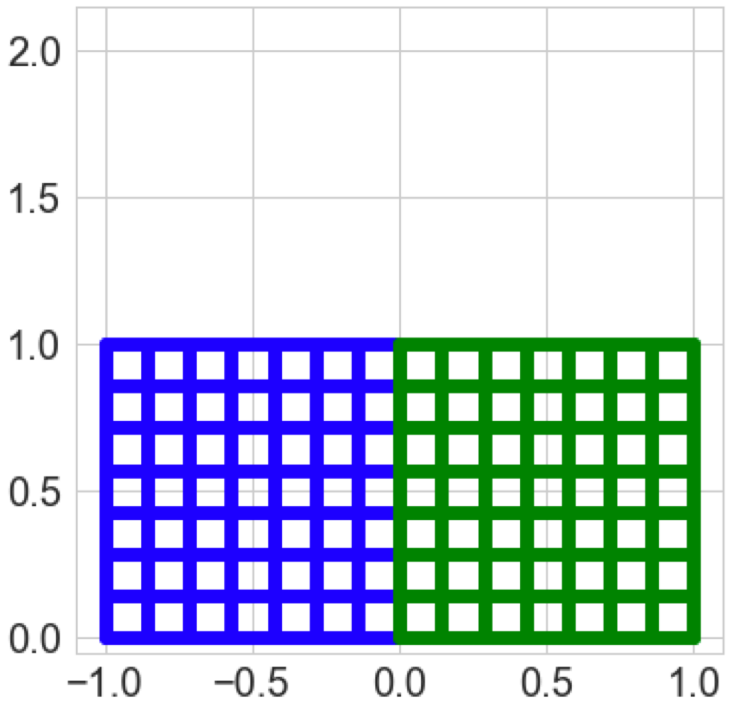


$$\begin{pmatrix} 3/2 & 0 \\ 0 & 2/3 \end{pmatrix}$$

Reflection operator

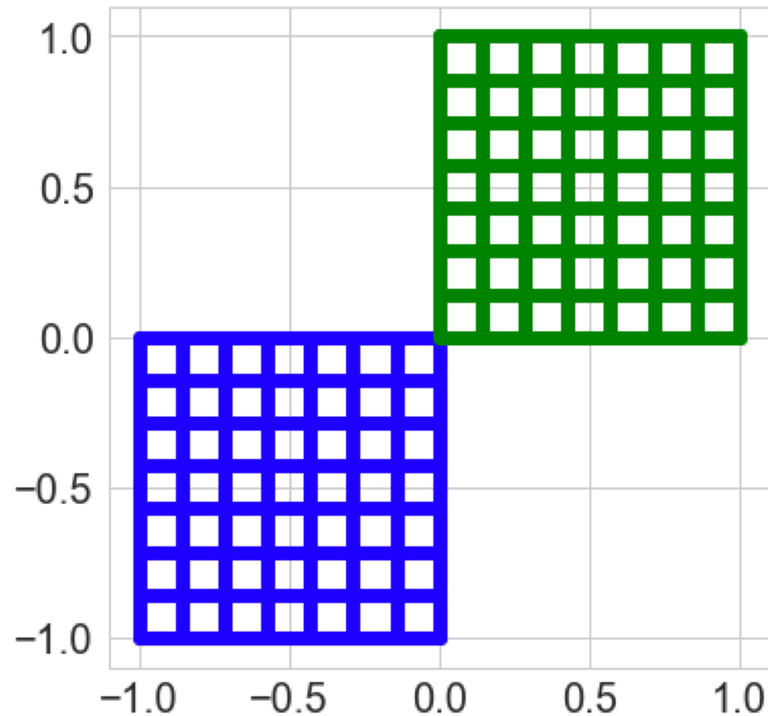
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

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



Reflect about y-axis

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

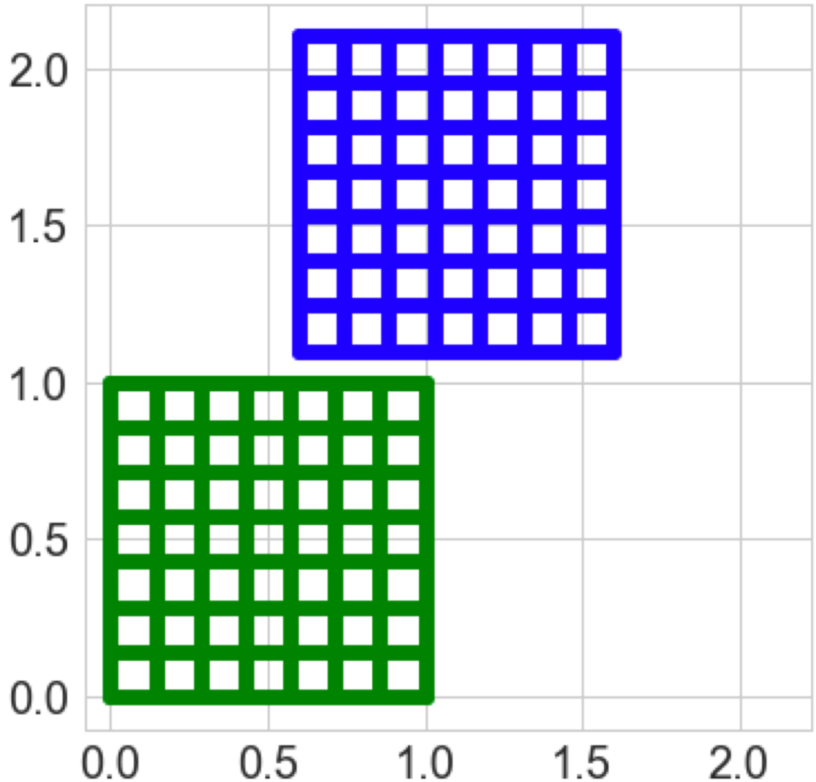


Reflect about x and y-axis

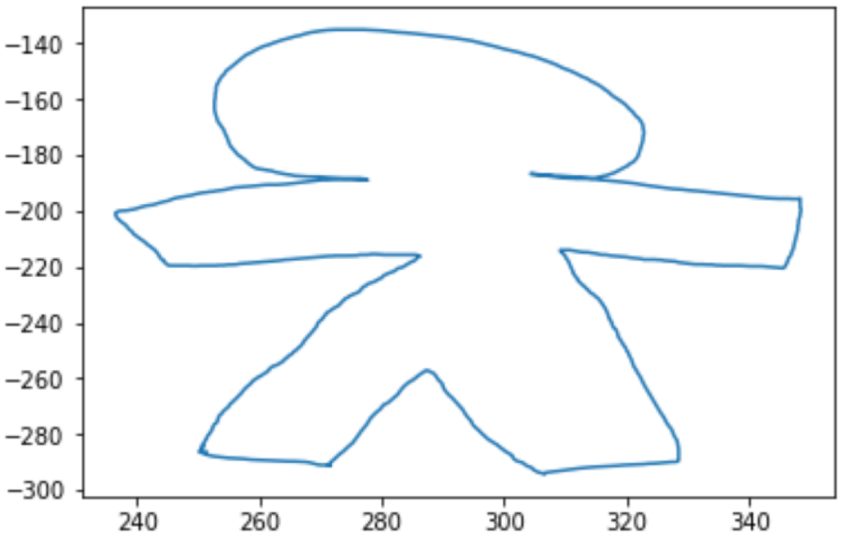
Translation (shift)

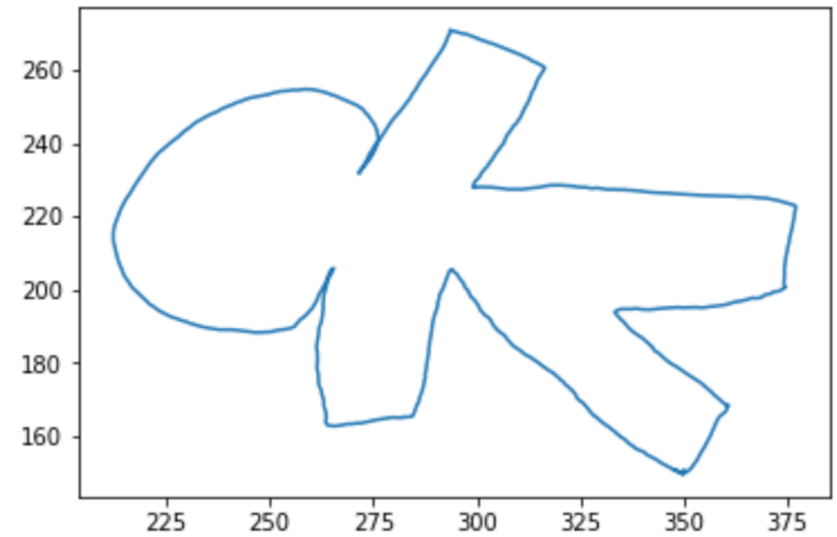
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$$

$a = 0.6; b = 1.1$



Matrices operating on data





Data set: *A*

Data set: *B*

Rotation

Demo “Matrices for geometry transformation”

Notation and special matrices

- Square matrix: $m = n$

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

- 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

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

- 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.

Norms

What's a norm?

- A generalization of 'absolute value' to vectors.
- $f(x) : \mathbb{R}^n \to \mathbb{R}_0^+$, 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 $\|\mathbf{x} + \mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$

Example of Norms

What are some examples of norms?

The so-called *p-norms*:

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

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

$p=1 \Rightarrow \|\underset{\sim}{x}\|_1 = |x_1| + |x_2| + \dots |x_n|$

$p=2 \Rightarrow \|\underset{\sim}{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$

$p=\infty \Rightarrow \|\underset{\sim}{x}\|_\infty = \max_i |x_i|$

Unit Ball:

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

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!

How? Attempt 1:

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

Attempt 2:

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

Absolute and Relative Errors

What are the absolute and relative errors in approximating the location of Siebel center $(40.114, -88.224)$ as $(40, -88)$ using the 2-norm?

$$\underset{\sim}{x}_{exact} = (40.114, -88.224) \qquad \underset{\sim}{x}_{approx} = (40, -88)$$

$$\underset{\sim}{error}_a = \underset{\sim}{x}_{exact} - \underset{\sim}{x}_{approx} \rightarrow \text{this is a vector!}$$

$$\|\underset{\sim}{e}_a\|_2 = \|(0.114, -0.224)\| = \sqrt{0.114^2 + 0.224^2} = 0.2513$$

$$\|\underset{\sim}{e}_r\|_2 = \frac{\|x_{exact} - x_{approx}\|_2}{\|x_{exact}\|_2} = \frac{0.2513}{\sqrt{40.114^2 + 88.224^2}} = 2.593 \times 10^{-3}$$

Matrix Norms

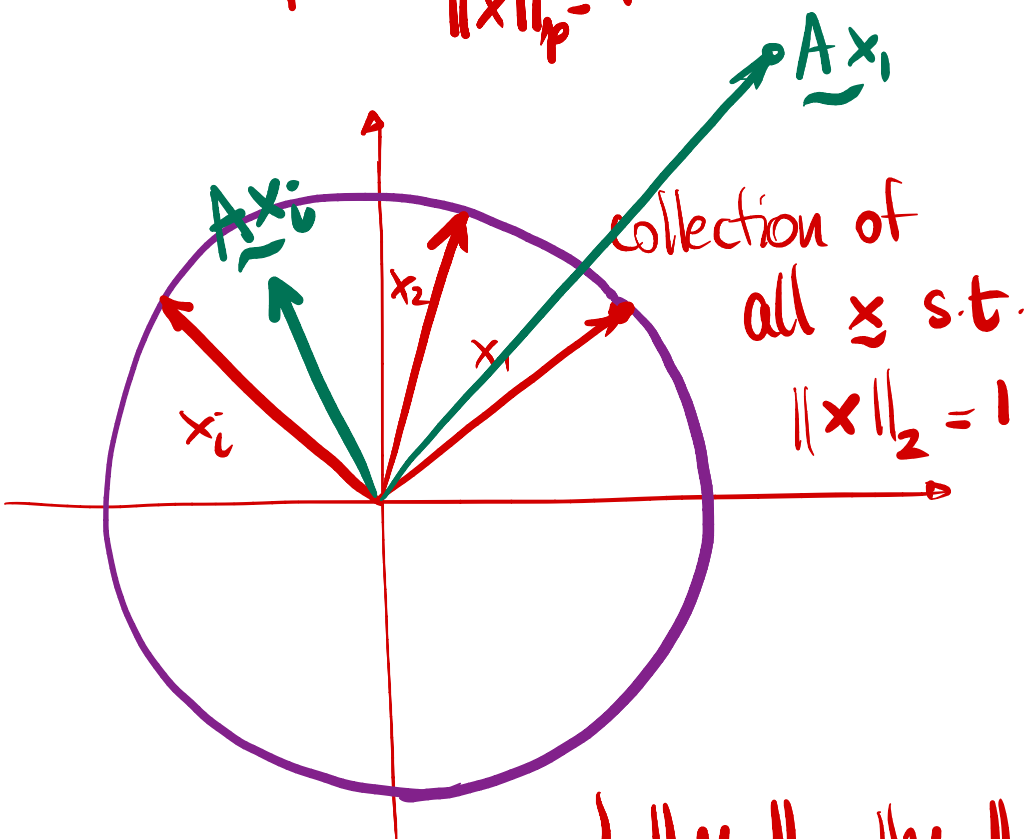
What norms would we apply to matrices?

- 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 Matrix Norms

$$\|A\|_p = \max_{\|x\|_p=1} \|Ax\|_p = \max_{\|x\|\neq 0} \frac{\|Ax\|_p}{\|x\|_p}$$



$p = 2$

$y_1 = \underset{\sim}{A}\,\underset{\sim}{x}_1 \rightarrow \|y_1\|_2 =$ —

$y_2 = \underset{\sim}{A}\,\underset{\sim}{x}_2 \rightarrow \|y_2\|_2 =$ —

$\vdots$

$y_i = \underset{\sim}{A}\,\underset{\sim}{x}_i \rightarrow \|y_i\|_2 =$ —

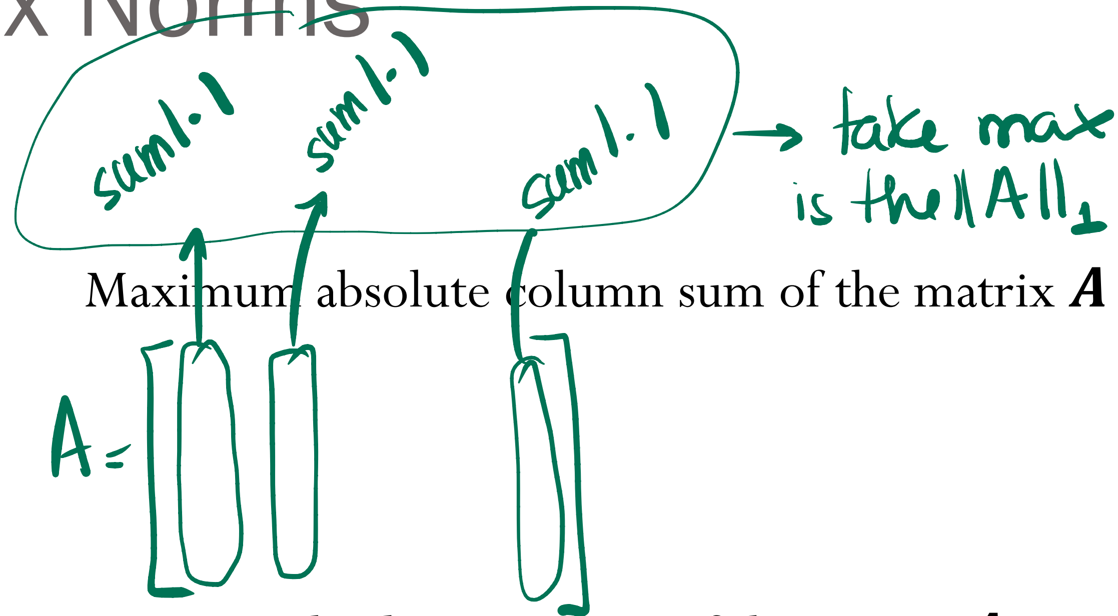
$$\{\|y_1\|_2, \|y_2\|_2, \ldots, \|y_i\|_2, \ldots\}$$

$$\Longrightarrow \|A\|_2 = \max \|y_i\|_2$$

Induced Matrix Norms

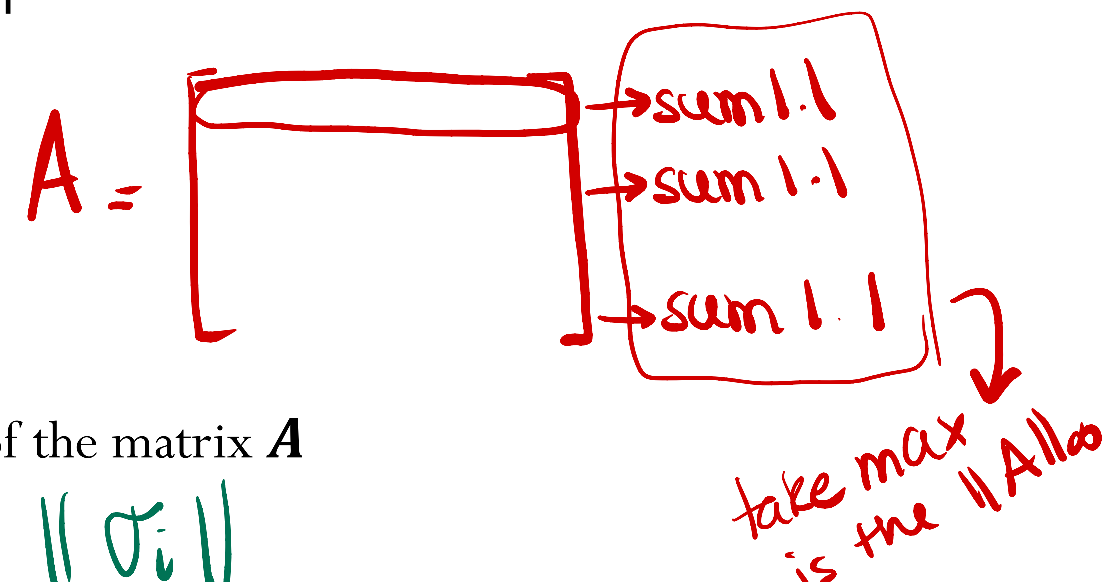
$$\|\boldsymbol{A}\|_1 = \max_j \sum_{i=1}^{n} |A_{ij}|$$

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



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

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



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

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

$\max \|\sigma_i\|$

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

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

a) 3

b) 4

c) 5

d) 6

e) 7

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

$$\max\left\{\frac{||Ax||}{||x||}, \frac{||Ay||}{||y||}, \frac{||Az||}{||z||}\right\}$$

$$= \max\left\{\frac{20}{2}, \frac{5}{1}, \frac{90}{3}\right\} = \max\{10, 5, 30\}$$

Induced Matrix Norm of a Diagonal Matrix

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

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

singular values of diagonal matrix are the diagonal entries.

singular values of $A = \{100, 13, 0.5\}$

$\|A\|_2 = 100$

we will show this later when introduce SVD.

Induced Matrix Norm of an Inverted Diagonal Matrix

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

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

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

$$\text{sing}(A^{-1}) = \left\{ \frac{1}{100}, \frac{1}{13}, \frac{1}{0.5} \right\}$$

$$\| A^{-1} \| = 2$$