

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

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

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

Clicker question

1) Is

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

a linear function?

A) YES

B) NO

2) Is

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

a linear function?

A) YES

B) NO

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 $f(\mathbf{x})$ can be represented by a Matrix-Vector multiplication.
- Think of a matrix \mathbf{A} as a linear function that takes vectors \mathbf{x} and transforms them into vectors \mathbf{y}

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

- Hence we have:

$$\mathbf{A} (\mathbf{u} + \mathbf{v}) = \mathbf{A} \mathbf{u} + \mathbf{A} \mathbf{v}$$

$$\mathbf{A} (\alpha \mathbf{u}) = \alpha \mathbf{A} \mathbf{u}$$

Matrix-Vector multiplication

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

Dot product of \mathbf{x} with
rows of \mathbf{A}

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

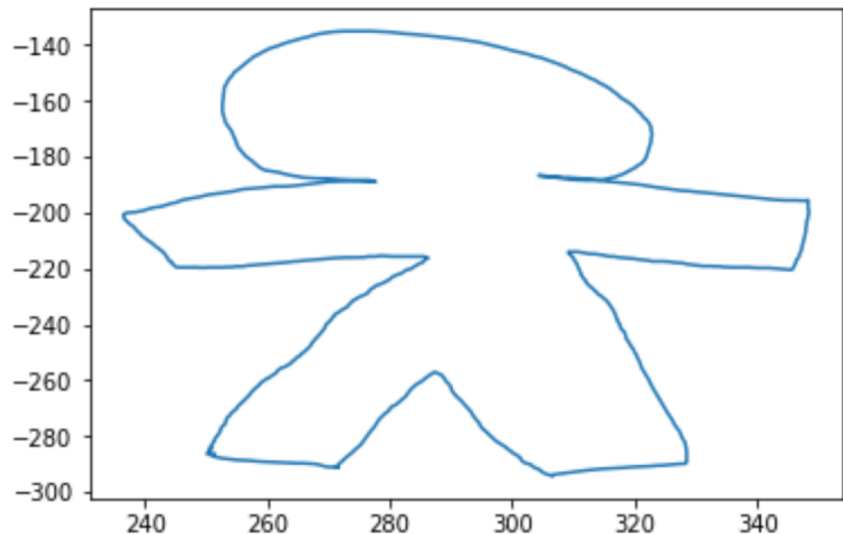
Matrix-Vector multiplication

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

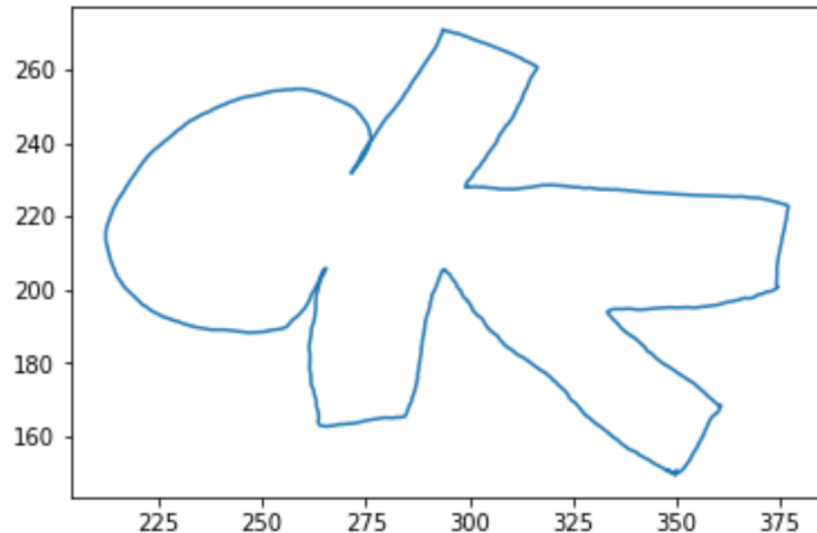
Linear combination of
column vectors of \mathbf{A}

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

Matrices operating on data



Data set: x



Data set: y

Rotation

$$y = f(x)$$

or

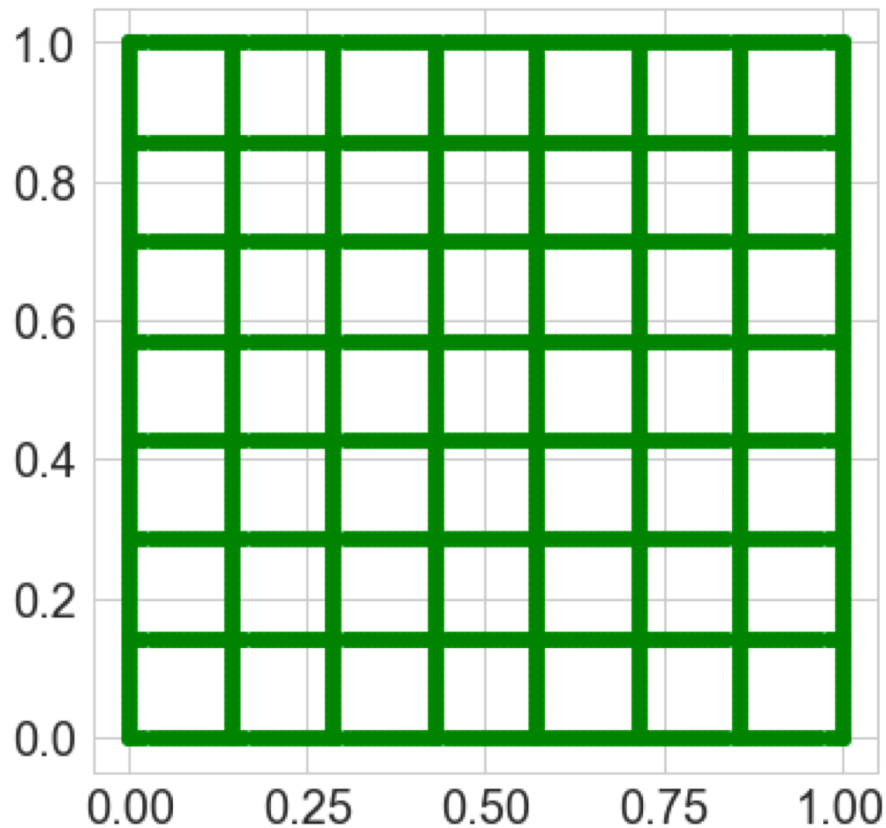
$$y = A x$$

Example: Shear operator

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

Matrices as operators

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



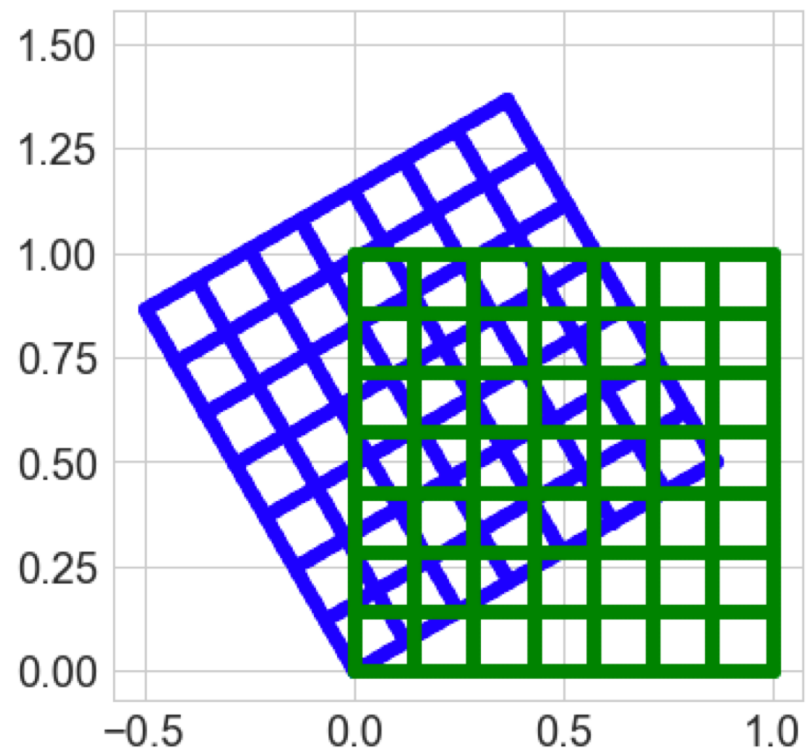
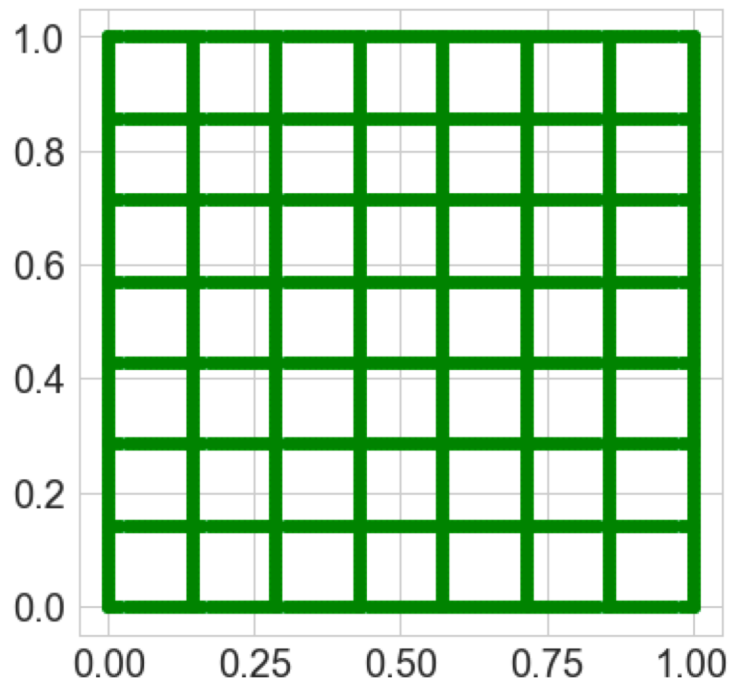
What can matrices do?

1. Shear
2. Rotate
3. Scale
4. Reflect
5. Can they translate?

Rotation operator

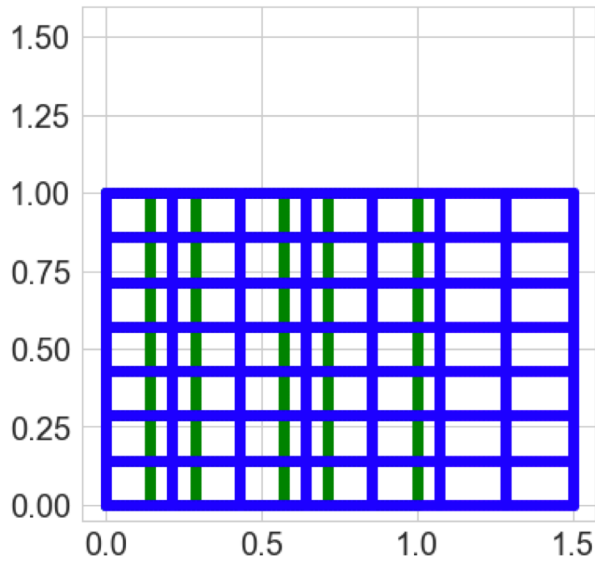
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\theta = \pi/6$$



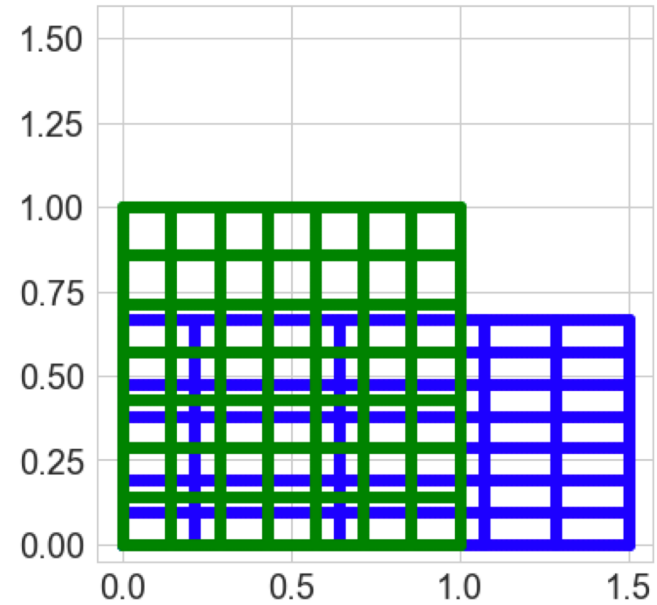
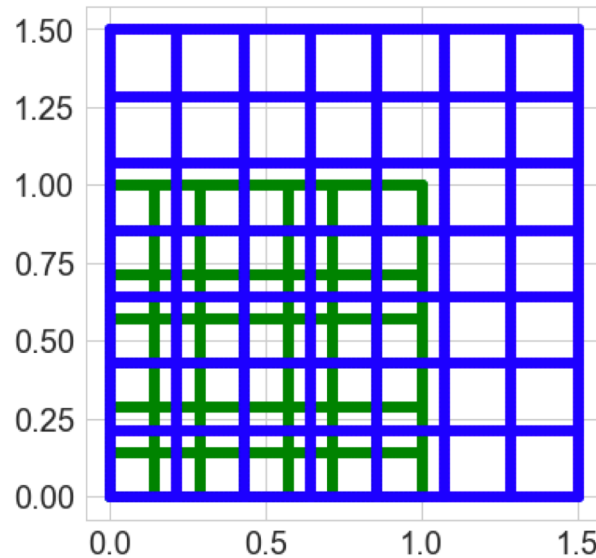
Scale operator

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



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

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

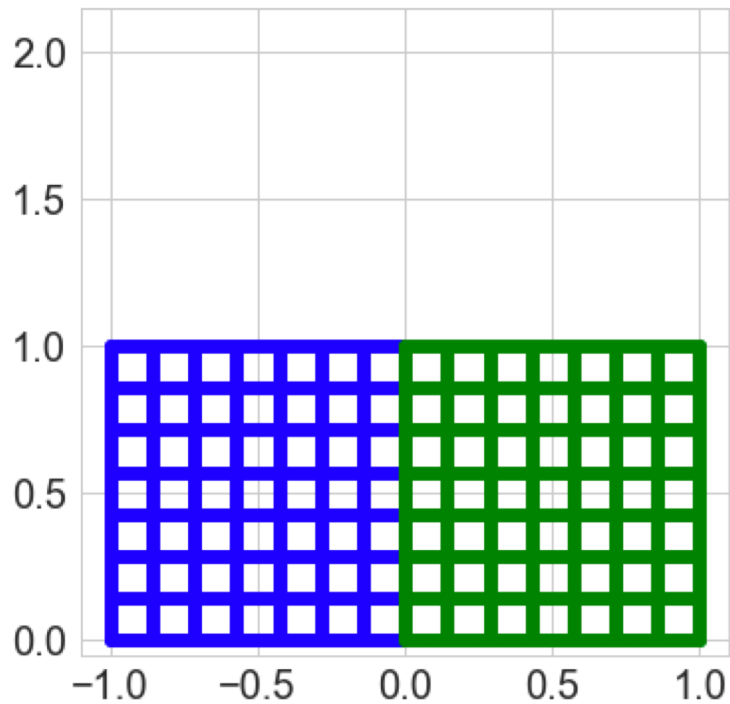


$$\begin{pmatrix} 3/2 & 0 \\ 0 & 2/3 \end{pmatrix}$$

Reflection operator

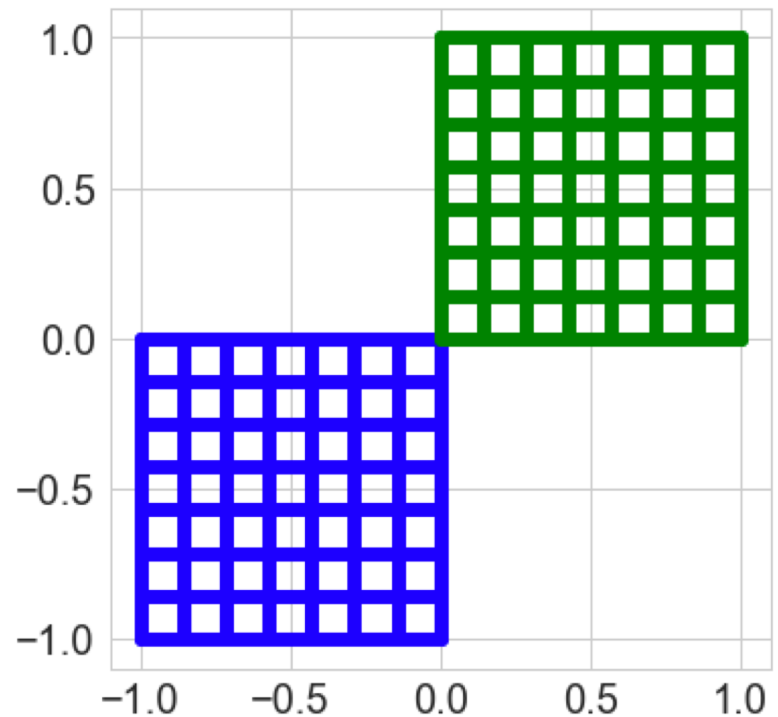
$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -a & 0 \\ 0 & -b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

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



Reflect about y-axis

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

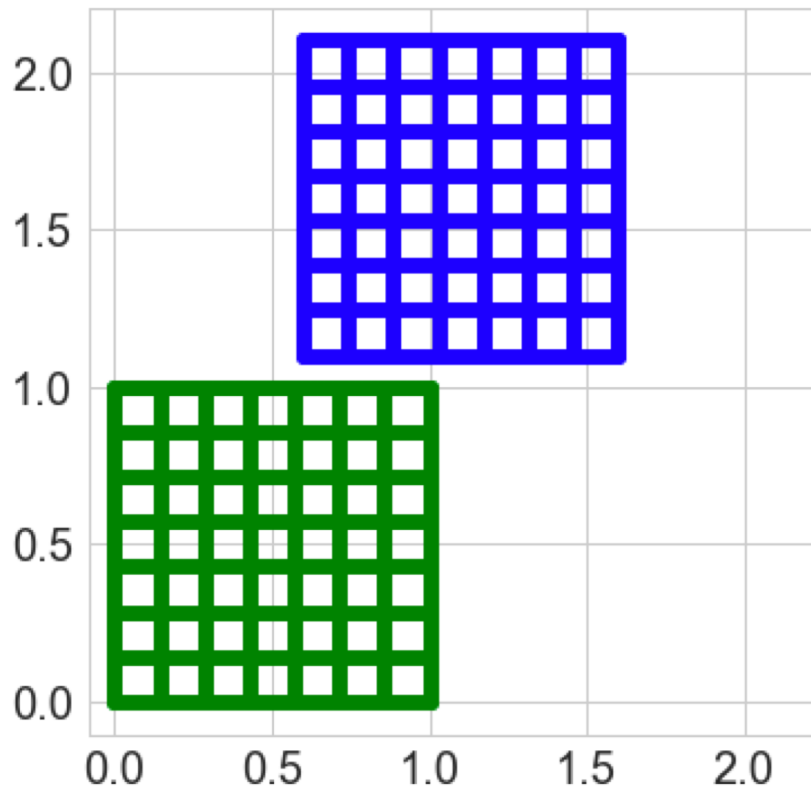


Reflect about x and y-axis

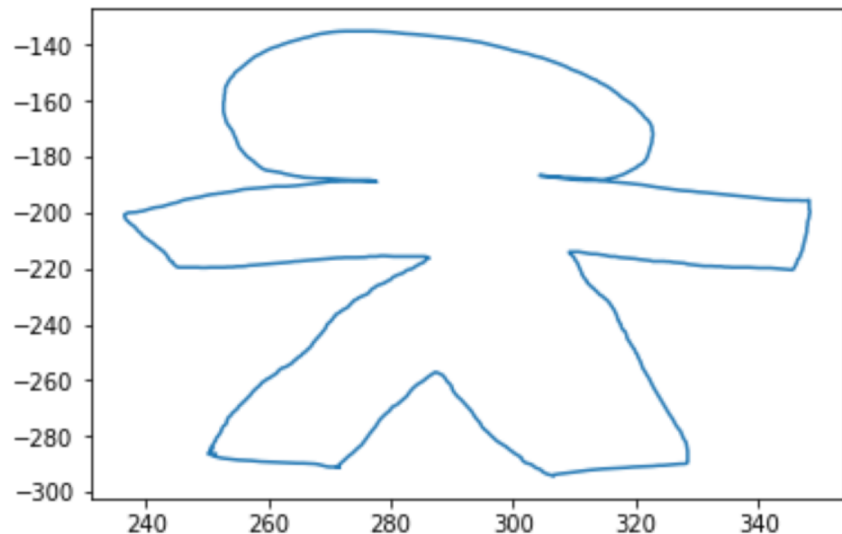
Translation (shift)

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

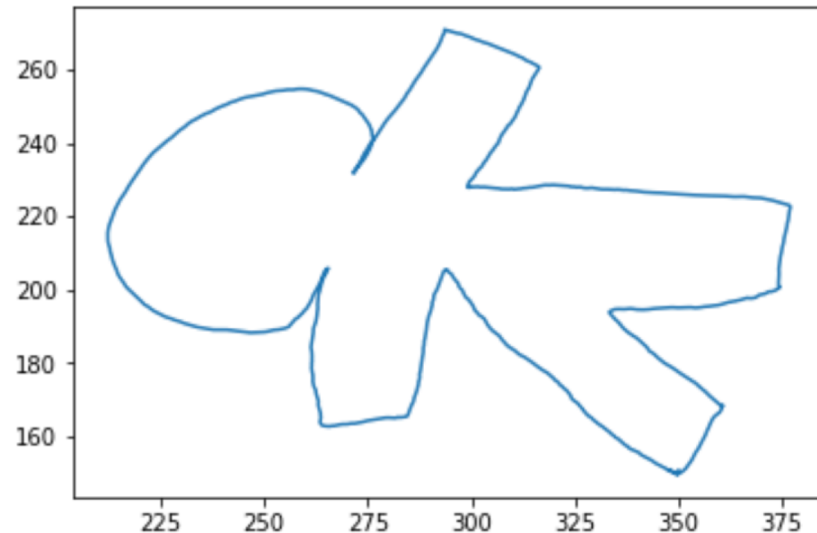
$$a = 0.6; b = 1.1$$



Matrices operating on data



Data set: *A*



Data set: *B*

Rotation

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 $[\mathbf{I}] = [\delta_{ij}]$

- Symmetric matrix: $A_{ij} = A_{ji}$ $[\mathbf{A}] = [\mathbf{A}]^T$

- Permutation matrix:

- Permutation of the identity matrix

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

- Permutes (swaps) rows

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

- Triangular matrix:

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

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

More about matrices

- Rank: the rank of a matrix \mathbf{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 \mathbf{A} has shape $m \times n$:
 - $\text{rank}(\mathbf{A}) \leq \min(m, n)$
 - Matrix \mathbf{A} is **full rank**: $\text{rank}(\mathbf{A}) = \min(m, n)$. Otherwise, matrix \mathbf{A} is **rank deficient**.
- Singular matrix: a square matrix \mathbf{A} is invertible if there exists a square matrix \mathbf{B} such that $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$. If the matrix is not invertible, it is called singular.

Norms

What's a norm?

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

Define **norm**.

A function $\|\mathbf{x}\| : \mathbb{R}^n \rightarrow \mathbb{R}_0^+$ 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 $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

Example of Norms

What are some examples of norms?

The so-called p -norms:

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

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

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

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:

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

Attempt 2:

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?

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

Induced Matrix Norms

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

Maximum absolute column sum of the matrix \mathbf{A}

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

Maximum absolute row sum of the matrix \mathbf{A}

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

σ_k are the singular value of the matrix \mathbf{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 $\|A + B\| \leq \|A\| + \|B\|$

But also some more properties that stem from our definition:

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

Both of these are called **submultiplicativity** of the matrix norm.

Iclicker question

Determine the norm of the following matrices:

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

a) 3

b) 4

c) 5

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

d) 6

e) 7

Clicker 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

Induced Matrix Norm of a Diagonal Matrix

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

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

Induced Matrix Norm of an Inverted Diagonal Matrix

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

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