Introduction to Linear Algebra

Kwanho Park and Kishan Sawant

Table of Contents

Table of Contents

Brief intro to MRC

1	frames
2	Matrices, vector norms, eigenvectors and eigenvalues
3	Matrix Decomposition (SVD)
4	Singular Value Decomposition Applications
5	Jacobians
6	Ordinary Differential Equations I
7	Ordinary Differential Equations II
8th	Midterm (NO lecture)
9	Basics of probability
10	Bayesian reasoning
11	Random variables
12	Joint random variables + covariance
13	TBD
14	recap

Linear Algebra

Bayesian Statistics

Markov Process

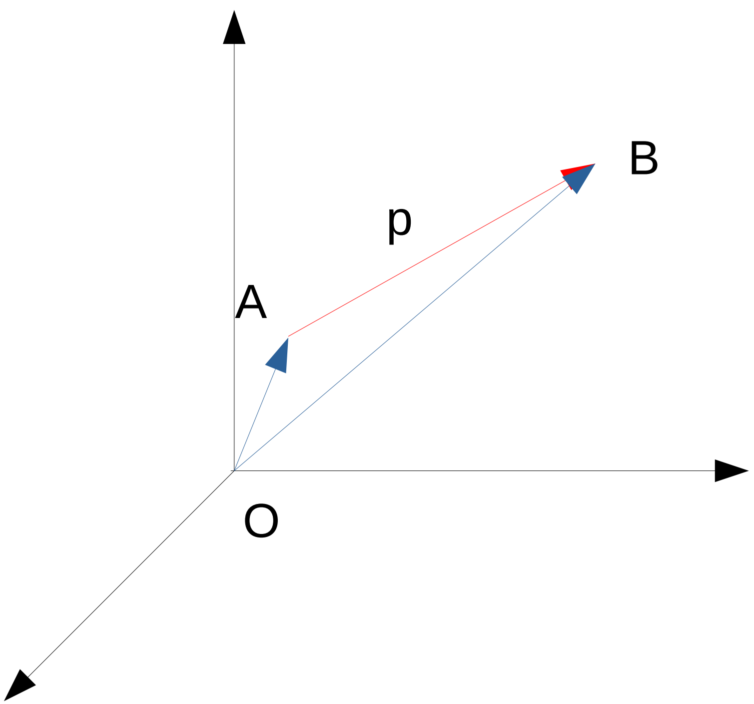
Mathematical description

- It starts from definitions and is ended by definitions

- $\mathbf{R}$: set of real number
- $\mathbf{R}^n$: set of n dimensional real number
- $X : A \rightarrow B$: mapping (morphism) X from A to B

Vectors

- n tuple of point = n dim'l vector p ∈ $\mathbf{R}^n$



$$\begin{aligned} p &= OB - OA \\ &= AB \\ &= (x, y, z) \end{aligned}$$

Vectors

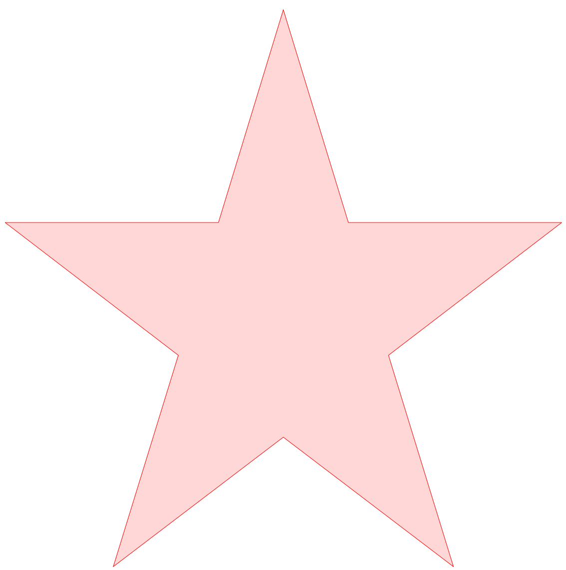
- Then, why not mathematical object with magnitude and direction?

Vectors

- Then, why not mathematical object with magnitude and direction?
- What is “*magnitude*”?

Norm & Inner product

- What is “*magnitude*” of following star?



Area?

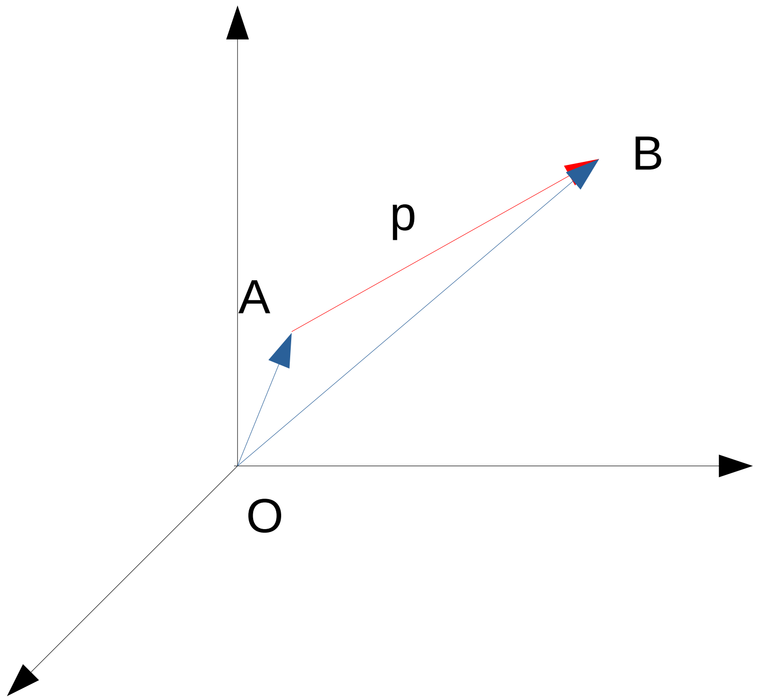
The number of vertices?

The number of edges?

Degree of concavity?

Norm & Products

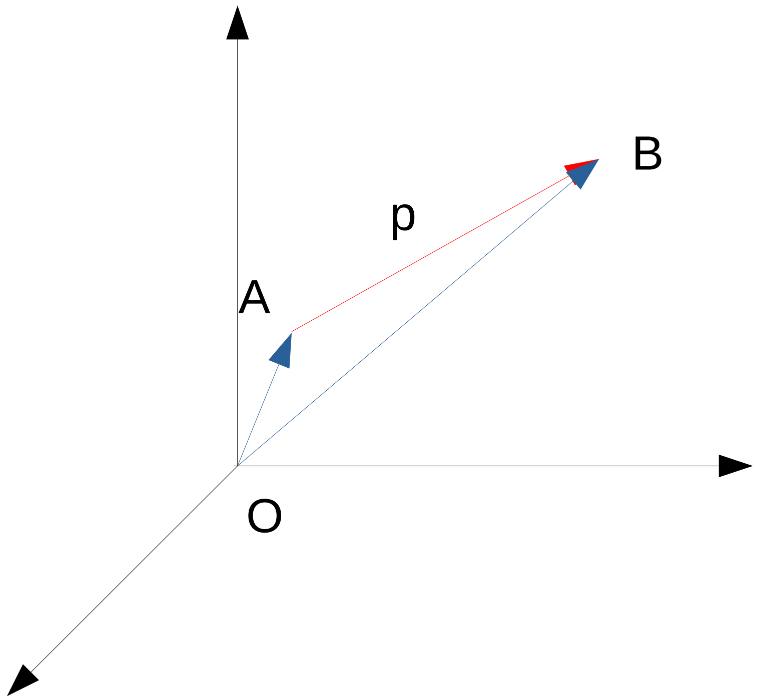
- Norm defines "*magnitude*" of vector
- Typical definition is Euclidean length (L2 norm)



Just measure length of arrow p
to estimate L2 norm, ||p|| = 4

Norm & Products

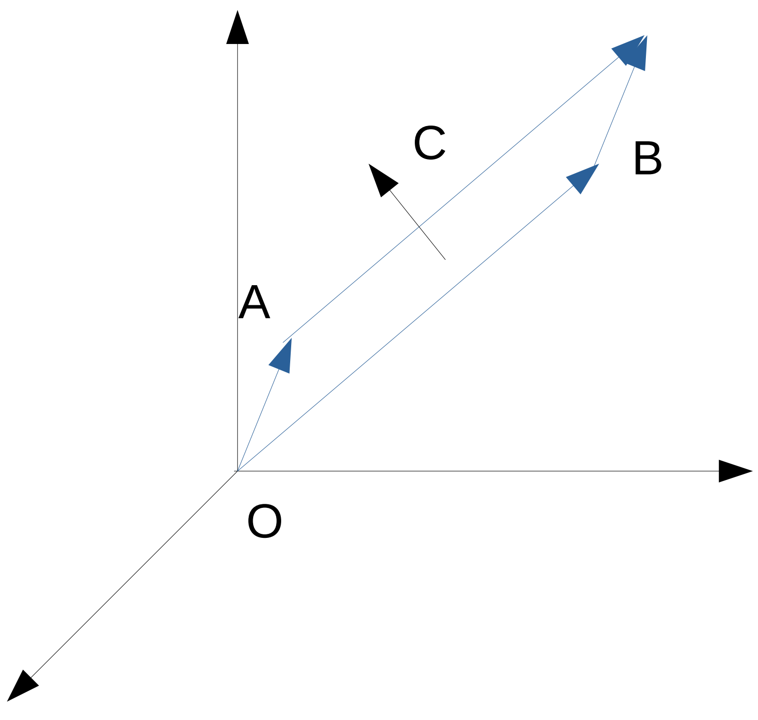
- Inner product <,> : $\mathbf{R}^n \times \mathbf{R}^n \rightarrow R$
- <A,B> : project A into B or vice versa



Norm and inner product can be related by ||p|| = <p, p>

Norm & Products

- Cross product x : $\mathbf{R}^3 \times \mathbf{R}^3 \rightarrow \mathbf{R}^3$
- B x A : generate new vector C

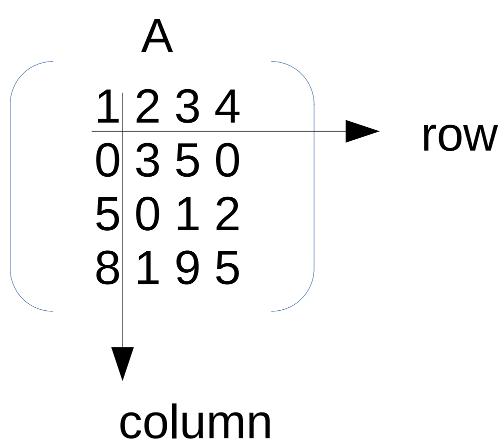


C is normal to both A and B
Size of C is square area
Only for 3d vector

Matrix & Operators

- A matrix is a table of numbers, it can represent abstract idea (isomorphism).

- ***Linear transform***, groups, tensors, etc



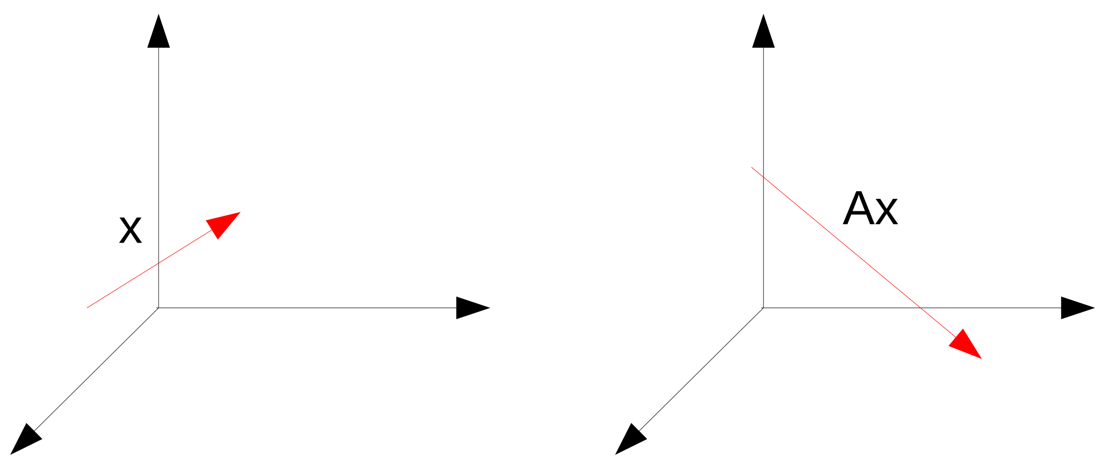
Matrix & Operators

- A matrix is a table of numbers, it can represent abstract idea (isomorphism).

- ***Linear transform***, groups, tensors, etc

A

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 5 & 0 \\ 5 & 0 & 1 & 2 \\ 8 & 1 & 9 & 5 \end{pmatrix}$$



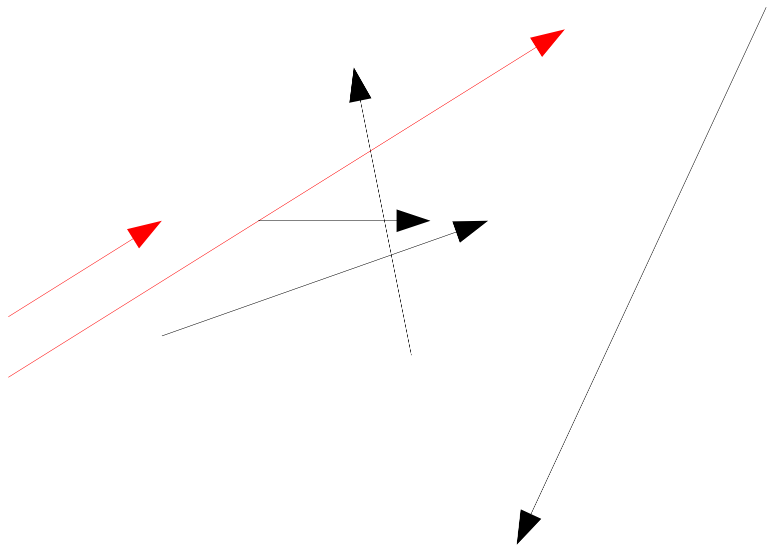
Matrix & Operators

- Commuteness : $[A,B] = AB - BA \neq 0$, In general, matrix multiplication is not commute
- Transpose : A^T, Flip the matrix over the diagonal axis.
- Trace : $Tr(A)$, Sum of diagonal elements
- Invertible : Let $Ax = b$, then $x = A^{-1}b$

Not every matrices are invertible! → only square and non singular matrices are invertible.

Vector space

- The world of vectors



Vector space

- Vector space is closed under the *vector addition* and *scalar multiplication.*
- The vector space over field F where F is set of scalars satisfying axioms for field defined by (X, Y, Z are vectors and a, b are scalars)

Commuteness : X + Y = Y + X

Associativity : (X + Y) + Z = X + (Y + Z), a(bX) = (ab)X
Distributivity : a(X + Y) = aX + aY, (a + b)X = aX + bX

Identity : 0 + X = X, aX = X when a = 1
Inverse : X + (-X) = 0

Linear span & Null space

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 5 & 0 \\ 5 & 0 & 1 & 2 \\ 8 & 1 & 9 & 5 \end{pmatrix}$$

row

column

Column space : vector space where it is linearly combined (or spanned) by column vectors.

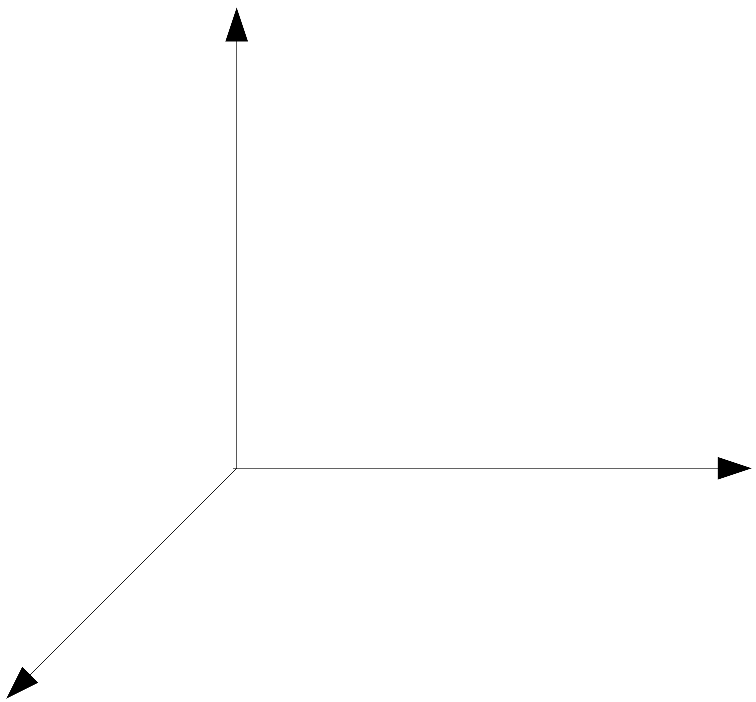
Row space : vector space where it is linearly combined (or spanned) by row vectors.

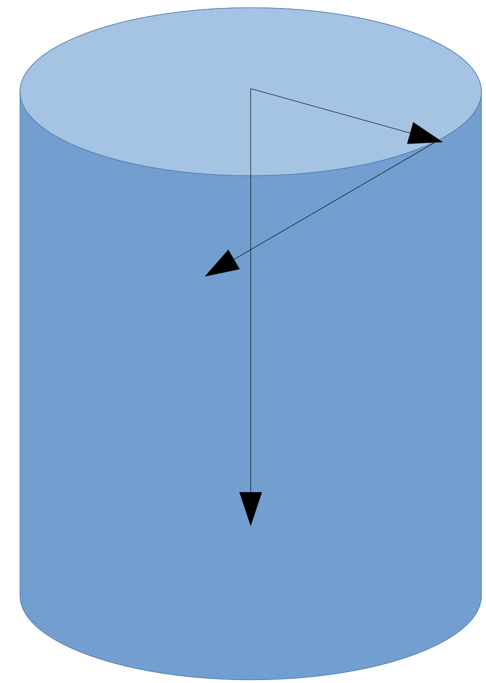
Null space : subspace of $\mathbf{R}^n$ which satisfies $Ax = 0$
rank(T) + null(T) = dim(T)

Orthonormal bases

- Orthogonality : linearly independent → <p, q> = 0
- Normal : <p, p> = 1
- Basis : set of vectors which can describes every element in vector space V

Orthonormal bases





Linear systems

$$\mathbf{A}x = b$$

- ‘b’ is linear combination of columns of **A** or independent co-efficients of equation of lines
- Linear dependency:
  - set of n-vectors $\{a_1, \ldots, a_k\}$ (with $k \geq 1$) are linearly dependent if,
    - $\beta_1 a_1 + \cdots + \beta_k a_k = 0$, with at least one non-zero β_k
- Note: If columns of A are linearly independent then $Ax = 0$ implies $x = 0$
- **Rank** of a matrix: maximum number of linearly independent rows or columns

Linear systems

$$\mathbf{A}x = b$$

- Over-constrained system:
  - Rank of **A** > dimension of b
  - No solution! But ‘x’ can be found which minimizes the norm, $||\mathbf{A}x-b||$, using *pseudo inverse, i.e.*, $(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$
- Under-constrained system:
  - Rank of of **A** < dimension of b
  - It allows infinitely many solutions

Linear systems

$$\mathbf{A}x = b$$

- Invertible matrix:
  - If **A** is a square matrix os size nxn and of rank ‘n’ (full-rank) then it is invertible
  - It is also called as non-singular matrix, i.e., a matrix with non-zero **determinant**

- How to find rank of a matrix ??
  - Perform **Gaussian elimination** to reduce the matrix to **row-echelon form**
  - Number of non-zero rows → Rank

Determinants

- Determinants are defined for square matrices and det(**A**) is a real number
- If matrix **A** is reduced to **B** after Gaussian elimination, then the determinant of **A** can be determined by,

$$\det(A) = (-1)^r \frac{(\text{product of the diagonal entries of } B)}{(\text{product of the scaling factors})}$$

- Determinants for 2D and 3D matrices
- A square matrix is invertible if it has a non-zero determinant

Determinants

- Applications:
  - The absolute value of det(**A**) is the volume or area spanned by the columns **A**
  - Calculate Eigenvalues by solving det(**A**-λ**I**) = 0
- Note: In matrix-vector multiplication, **A**x streches the volume of imaginary sphere with radius vector ‘x’ by a factor of |det(A)|
  - The directions of the strech is determined by its **eigenvalues** and **eigenvectors**

Eigenvalues & Eigenvectors

- An **eigenvector** of **A** is a nonzero vector ‘x’ such that $\mathbf{A}x = \lambda x$, for some scalar real number ‘λ’, which is termed as an **eigenvalue**
- Intuition: An eigenvector in the vector space of **A**, only gets transformed by magnitude
- Note: Eigenvalue can be zero, but not the eigenvector

Eigenvalues & Eigenvectors

- How to find **eigenvalue** and **eigenvector**?
  - Solve det(**A** - λ**I**) = 0 to find all eigenvalues (λ)
  - Substitute them in **A**x = λx to find the respective eigenvectors

- Note: Single **eigenvalue** can have multiple **eigenvectors** associated with it !

Matrix decomposition

- Why do we need matrix decompositions?
  - Computational efficiency
  - To extract underlying pattern in the matrix
- Different types:
  - LU decomposition
  - QR decomposition
  - Singular Value Decomposition (SVD)
  - Cholesky decomposition
  - Eigendecomposition
  - and many more ...

Diagonalization (eigendecomposition)

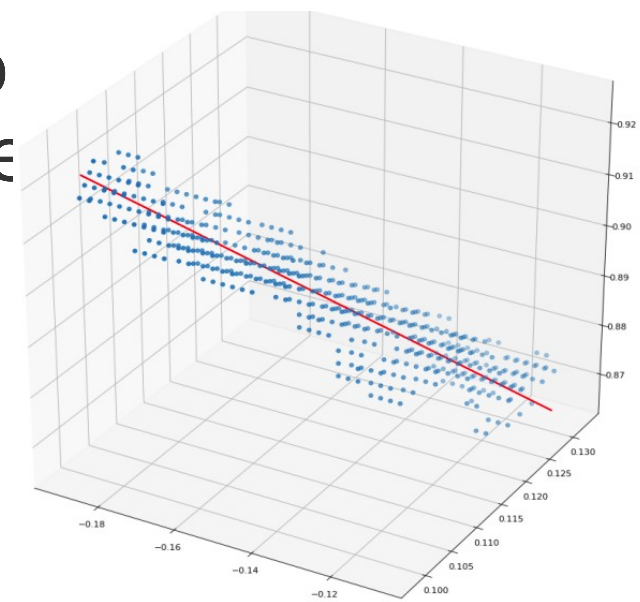
- An n × n matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.
- $\mathbf{A} = \mathbf{C}\,\mathbf{B}\,\mathbf{C}^{-1}$,
  - where **B** is a diagonal matrix with **eigenvalues** as its diagonal elements and **C** is a matrix having corresponding **eigenvectors** as its columns
- Note: If **A** is a symmetric matrix then the **eigenvectors** are orthogonal to each other

Diagonalization

- Significance?
  - Many matrix operations requires finding powers of a matrix
  - $\mathbf{A}^n = \mathbf{C}\ \mathbf{B}^n\ \mathbf{C}^{-1}$

Applications

- **PCA**:
  - Useful for pattern recognition and dimensionality reduction
  - Eg: Given a set of data po
    of a matrix **A**, find the dire
    spread of data points
  - Considering the ‘n’ points
    are given in 3D space, **A**
    will be of size ‘3xn’



Applications

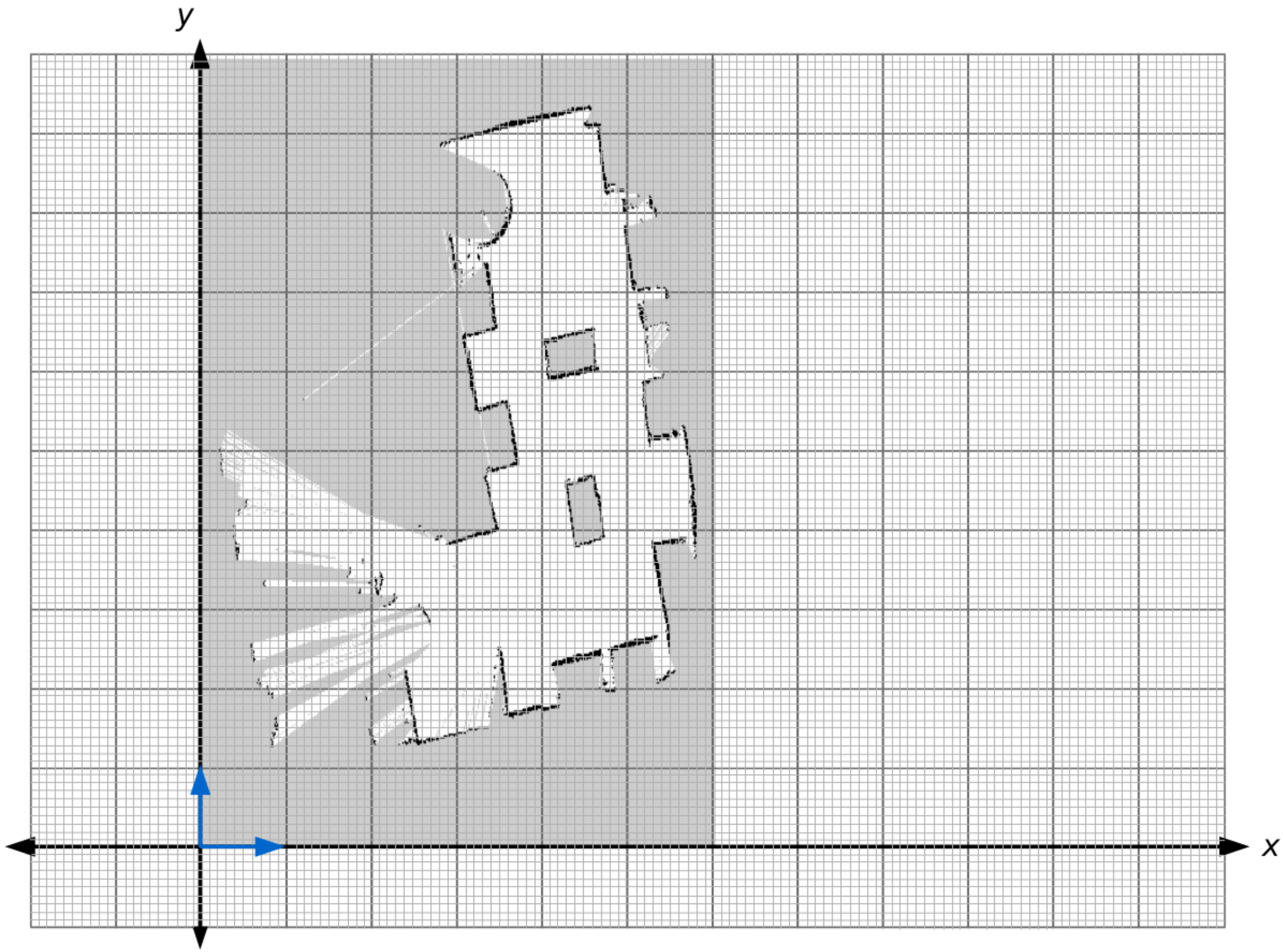
- **PCA:**
- Covariance matrix:
  - Formula: $$\begin{bmatrix} var(x) & cov(x,y) & cov(x,z) \\ cov(x,y) & var(y) & cov(y,z) \\ cov(x,z) & cov(y,z) & var(z) \end{bmatrix}$$
  - Sample variance: $\mathrm{var(x)} = \frac{\sum_1^n (x_i - \bar{x})^2}{n-1}$
  - Sample covariance: $\mathrm{cov(x, y)} = \frac{\sum_1^n (x_i - \bar{x})(y_i - \bar{y})}{n-1}$

Applications

- **PCA:**
- Step 1: Subtract mean of all rows from every row of the matrix
- Step 2: find covariance matrix
  - For given sample points in $\mathbf{A}$, this can be found by, cov(A) = ($\mathbf{AA}^T$)/(n-1)
- Step 3: Find the eigenvector corresponding to the highest eigenvalue, which represents the direction of maximum spread of the data
- Why is it so? What does other eigenvectors signify?

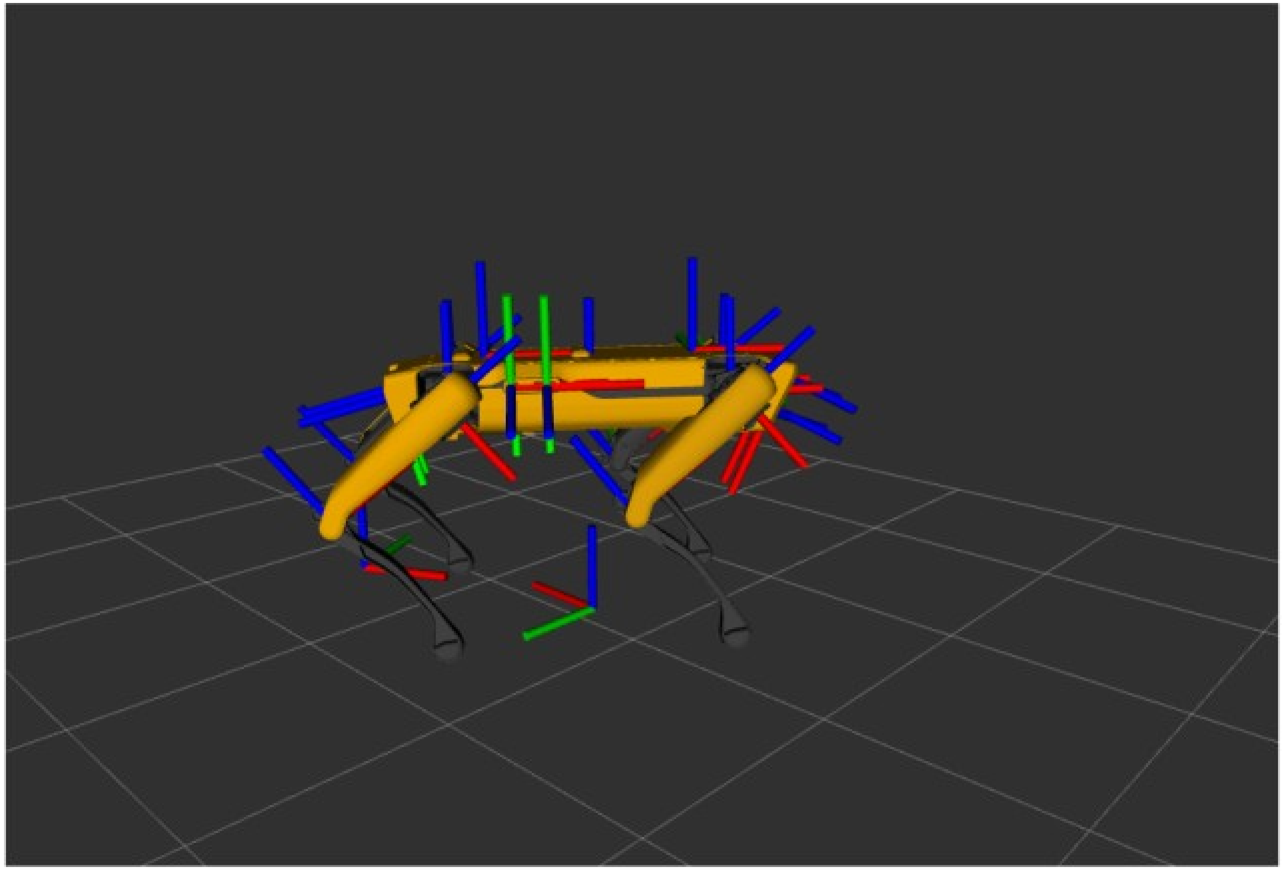
Applications

- Navigation



Applications

- Robot control



Applications

- Computer vision
  - Image processing
  - PCA (dimensionality reduction and pose estimation)
  - Object recognition
  - Image transformation
  - Camera calibration
  - many more ..

References:

- https://course-repos.math.gatech.edu/math1553/slides/allslides-web.pdf
- http://ais.informatik.uni-freiburg.de/teaching/ss11/robotics/slides/02-linear-algebra.ppt.pdf
- https://www.cuemath.com/algebra/covariance-matrix/

Suggested Readings

- (Intro) Contemporary Linear Algebra written by Howard Anton & Robert C. Busby
- (Regular) Linear Algebra written by Hoffman & Kunz
- Introduction to Linear Algebra written by Gilbert Strang
- Lecture series on Linear Algebra by Professor Gilbert Strang (available on youtube)