Maths Preliminaries

September 9, 201

Introduction

- This review focuses on Linear Algebra, in the context of COMP6714.
- Key take-away points
  - Matrices as Linear mappings/functions

Note

- You've probability learned Linear Algebra from matrix/system of linear equations, etc. We will review key concepts in LA from the perspective of linear transformations (think of it as *functions* for now). This perspective provides semantics and
-
- models/operations in this perspective e!

A Common Trick in Maths I

Question

Calculate 2^{10}, 2^{-1}, $2^{\ln 5}$ and 2^{4-3i}?

-
  - a a a
  - $f(u) * f(v) = f(u+v)$.
  - $f(x) = y \Leftrightarrow \ln(y) = x \ln(a) \Leftrightarrow f$
  - $e^{ix} = cos(x) + i \cdot sin(x)$.
- The trick:
- Same in Linear algebra

Objects and Their Representations

Goal

- We need to study the objects
- On one side:
  - A good representation helps (a lot)!
-
  -

Basic Concepts I

Algebra

- a set of objects
- two operations and their identity objects (aka. *identify element*):
  -
  -
-
  - Closed for both operations
  - Some nice properties of these operation

    - Commutative: $\mathbf{a} + \mathbf{b} = \mathbf{b}$
    - Associative: $(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a}$
    - Distributive: $\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a}$

Basic Concepts II

Think: *What about substraction and division?*

Tips

Always use analogy from algebra on integers ($\mathbb{Z}$) *and* algebra on Poly

Why t

Basic Concepts III

Representation matters?

Consider even geometric vectors: $\mathbf{c} = \mathbf{a} + \mathbf{b}$

What if we represent vectors by a column of their coordinates?

What if by their polar coordinates?

Notes

- Informally, the objects we are concerned w are (column) vectors.
- The set of all n-dimensional real vectors is called $\mathbb{R}^n$.

(Column) Vector

Vector

- A n-dimensional vector, $\mathbf{v}$, is a $n \times 1$ matrix. We can emphasize its shape by calling it a *column vector*.
  - A *row vector* is a transposed column vector: $\mathbf{v}^\top$.

Oper

-
-

Linearity I

Linear Combination: Generalization of Univariate Linear Functions

- Let $\lambda_i \in \mathbb{R}$, given a set of k vectors $\mathbf{v}_i$ ($i \in [k]$), a linear combination of them is
  $\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \ldots + \lambda_k \mathbf{v}_k = \sum_{i \in [k]} \lambda_i \mathbf{v}_i$
  -

$$\mathbf{V} = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \ldots & v_k \\ | & | & & | \end{bmatrix} \qquad \lambda =$$

- Span: All linear combination of a set of vector of them.
- Basis: The minimal set of vectors whose span is exactly the whole $\mathbb{R}^n$.

Linearity II

- Benefit: every vector has a unique decomposition into basis.

  Think: *Why uniqueness is desirable?*

Examples

-
- *dant*

  Think: *Who?*

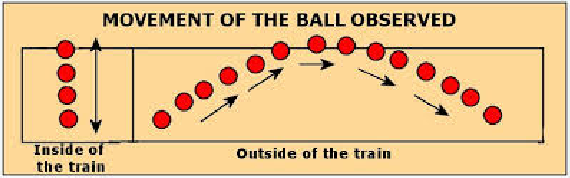
  - Decompose $\begin{bmatrix} 4 \\ 6 \end{bmatrix}$

Linearity III

Exercises

- What are the (natural) basis of all (univariate) Polynomials of degrees up to d?
- Decompose $3x^2 + 4x \quad 8$ into *the* linear combination of 2,

).
- The "same" polynomial is mapped to two di
  under two different bases. Think



Matrix I

Linear Transformation

- is a "nice" linear function that maps a vector in $\mathbb{R}^n$ to another vector in $\mathbb{R}^m$.

$$y_1$$

- The general form:

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \xrightarrow{f} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \implies y_3 = M_{31}x_1 + M_{32}x_2$$

Matrix II

Nonexample

$$\underset{x}{x_1} \xrightarrow{f} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \implies \begin{array}{l} y_1 = \alpha x_1^2 + \beta x_2 \\ y_2 = \gamma x_1^2 + \theta x_1 + \tau x_2 \end{array}$$

Matrix III

Why Only Linear Transformation?

- Simple and nice properties:
  - $(f_1 + f_2)(x) = f_1(x) + f_2(x)$
  - $(\lambda f)(x) = \lambda \cdot f(x)$
  - What about $f(g(x))$?
- Useful

Matrix I

Definition

- A $m \times n$ matrix *corresponds to* a *linear transformation* from $\mathbb{R}^n$ to $\mathbb{R}^m$
  - $f(\mathbf{x}) = \mathbf{y} \implies \mathbf{M}\mathbf{x} = \mathbf{y}$, where matrix-vector
  -
  - g

    mapping; the latter is more or less the understanding of a *function*. The *morphism* in category theory.

Semantic Interpretation

Matrix II

- Linear combination of columns of **M**:

$$\begin{bmatrix} | & | & & | \\ M_1 & M_2 & \dots & M_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} | \\ x \\ | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ M_1 & M_2 & \dots & M_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$

Matrix III

- Example:

$$\begin{bmatrix} 1 & 2 \\ -4 & 9 \\ 25 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 10 \end{bmatrix} = 1 \begin{bmatrix} 1 \\ -4 \\ 25 \end{bmatrix} + 10 \begin{bmatrix} 2 \\ 9 \\ 1 \end{bmatrix} = \begin{bmatrix} 21 \\ 86 \\ 35 \end{bmatrix}$$

Think: *What does* $\mathbf{M}$ *do for the last e*

- Rotation and scaling

- When $\mathbf{x}$ is also a matrix,

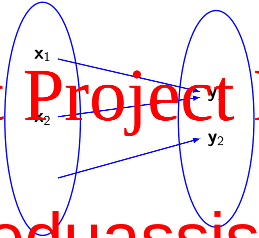
$$\begin{bmatrix} 1 & 2 \\ -4 & 9 \\ 25 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 10 & 20 \end{bmatrix} = \begin{bmatrix} 21 & 42 \\ 86 & 172 \\ 35 & 70 \end{bmatrix}$$

System of Linear Equations I

$$
\begin{aligned}
y_1 &= M_{11}x_1 + M_{12}x_2 \\
y_2 &= M_{21}x_1 + M_{22}x_2 \\
y &= M \quad x + M \quad x
\end{aligned}
\implies
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} =
\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}
$$
$$M \qquad M$$

- Interpretation: find a vector in $\mathbb{R}^2$ … $\mathbf{M}\mathbf{x}$ is exactly the given vector $\mathbf{y}$
- How to solve it?

System of Linear Equations II



The above transformation is *injec*

A Matrix Also Specifies a (Generalized) Coordinate System I

Yet another interpretation

- $\mathbf{y} = \mathbf{M}\mathbf{x} \implies \mathbf{I}\mathbf{y} = \mathbf{M}\mathbf{x}$.
- e

A Matrix Also Specifies a (Generalized) Coordinate System II

Example for polynomials

for 1 1 0 0 for 3 3 1 −4

$\mathbf{I}$: fo

fo

Let $\mathbf{x} = \quad -2 \quad 3 \quad \Longrightarrow \quad \mathbf{Mx} = \mathbf{I} \quad 13 \quad 6$

Exercise I

- What if **y** is given in the above example?
- What does the following mean?

3 1 4 1 0 0 3 1 4

- Think about representing polynomials using the basis: $(x-1)^2$, x^2-1, x^2+1.

Inner Product

THE binary operator – some kind of "similarity"

- Type signature: vector $\times$ vector $\to$ scalar: $\langle \mathbf{x}, \mathbf{y} \rangle$.
  - In $\mathbb{R}^n$, usually called *dot product*: $\mathbf{x} \cdot \mathbf{y} \stackrel{\text{def}}{=} \mathbf{x}^\top \mathbf{y} = \sum_i x_i y_i$.
  - For certain functions, $\langle f, g \rangle = \int_a^b f(t)g(t)\,\mathrm{d}t$. leads to the
-
  -
  - linearity in the first argument: $\langle a\mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = a\langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$
  - positive definitiveness: $\langle \mathbf{x}, \mathbf{x} \rangle \geq$
- Generalizes many geometric concepts to v (orthogonal), projection, norm
  - $\langle \sin nt, \sin mt \rangle = 0$ within $[-\pi, \pi]$ $(m \neq n) \Rightarrow$ they are orthogonal to each other.
- $\mathbf{C} = \mathbf{A}^\top \mathbf{B}$: $C_{ij} = \langle A_i, B_j \rangle$
  - Special case: $\mathbf{A}^\top \mathbf{A}$.

Eigenvalues/vectors and Eigen Decomposition

"Eigen" means "characteristic of" (German)

- A (right) eigenvector of a square matrix $\mathbf{A}$ is $\mathbf{u}$ such that $\mathbf{A}\mathbf{u} = \lambda\mathbf{u}$.

- Not all matrices have eigenvalues. Here, we only consider

- $\mathbf{U}$ (as 1

  columns). Then $\mathbf{AU} = \mathbf{U}\Lambda$, or equiv

  This is the Eigen-Decomposition.

  - We can interpret $\mathbf{U}$ as a transform coordinate systems. Note that vectors i y orthogonal.
  - Λ as the scaling on each of the directions in the "new" coordinate system.

Applications

Compute $\mathbf{A}^n$

- Apply Eigen Decomposition. $\mathbf{A} = \mathbf{U}\Lambda\mathbf{U}^{-1}$.
- Then

Exercises I

- Rewrite $\sum_{i=1}^{n} a_i b_i$ in vector/matrix operations.

- turns

$\begin{bmatrix} x_1 \\ 2x_2 \\ 3x_3 \end{bmatrix}$.

Exercises II

- Suppose we want to apply the linear mapping $\mathbf{W}$ to more than one $\mathbf{x}$ vectors. Draw a schematic diagram to show how this can be done using matrix operations (rather than a loop).

- In machine learning, we usually store training data as a data matrix; if it is $n \times m$, then it has
  and each sample is characterized by its
  can I apply the above linear mapping to all the tr
  samples.

References and Further Reading I

- Gaussian Quadrature:
  https://www.youtube.com/watch?v=k-yUdqRXijo
- Linear Algebra Review and Reference.
  http://cs229.stanford.edu/section/cs229-linalg.pdf
- y/

- We Recommend a Singular Value Decomposition.
  http://www.ams.org/samplings/ vd