21-670 Linear Algebra For Data Science

- All important info on Canvas:
  canvas.cmu.edu/courses/24962

Highlights

me: Garrett Ervin
e-mail: gervin@andrew.cmu.edu
office: Wean 6124

OH: TBA

Books: none required (will post lecture notes)
will draw from Strang + Trefethen and Bau + Golub.

Grading Four to Five HWs — 75% (submit on Gradescope)
Final exam — 25%

Cutoffs: not worse than standard:

90 - 100 = A
80 - 89 = B
etc...

Course Overview

aldobar...

* Linear algebra fundamentals important for data science *

- will assume only basic background, e.g. that you know about:
  - what vectors/matrices are
  - basic vector/matrix operations (addition, multiplication, transposes, determinants)
  - basic vector/matrix operation properties (associativity, distributivity, $(AB)^T = B^T A^T$ etc...)

stop me if ever confused by some algebra!

↳ hope to post some slides on Canvas w/ review material.

- class focuses on matrix calculations, may sometimes invoke basic theoretical facts from linear algebra (e.g. all bases of a vector space have same size)

Major Topics

- matrix multiplication
- column/row/null spaces
- eigenstuff
- symmetric and positive definite matrices
- SVD
- tensors

A theme: matrix factorizations!

- Some factorizations we'll cover:
  1. $A = CR$ (explains column = row rank)
  2. $A = LU$ (matrix form of Gaussian elim)
  3. $A = XDX^{-1}$ (diagonalization)
  4. $S = QDQ^T$ (S symmetric, D diagonal, Q orthonormal)
  5. $A = U\Sigma V^T$ (SVD)
  6. $A = QR$ (Gram-Schmidt)

- Factorizations not our purpose per se, but each encapsulates important ideas.

Definitions + Conventions

- A usually refers to $m \times n$ matrix (or $n \times n$ if square)

$$A = \begin{array}{cc} & j \\ i & \underbrace{\left[\begin{array}{ccc} & \vdots & \\ \cdots & a_{ij} & \cdots \\ & \vdots & \end{array}\right]}_{n} \Big\}\, m \end{array}$$

- write a_{ij} (or A_{ij}) for ij-th entry
- Scalars and matrix entries from $\mathbb{R}$ or $\mathbb{C}$ (usually $\mathbb{R}$) not an arbitrary field.

- vectors, written $\vec{u}, \vec{v}, \ldots,$ are column vectors, i.e. $n \times 1$ matrices:

$$\vec{v} = \underbrace{[\;]}_{}\Big\}n$$

- write $\vec{v}^T$ (or sometimes $\vec{v}^*$) for "row version"

$$\vec{v}^T = \underbrace{[\qquad]}_{n}$$

- $\vec{0}$ = 0 vector = $\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

- a <u>linear combination</u> of vectors $\vec{v}_1, \ldots, \vec{v}_k$ is a vector of form $\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k$ where c_i scalars (usually in $\mathbb{R}$)

- these vectors are <u>independent</u> if only way $c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = \vec{0}$ is if $c_i = 0$ $\forall i \leq k$.

e.g. $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ independent (why?)

but $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 6 \end{bmatrix}$ not since:

$$(-1)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + 2\begin{bmatrix} 3 \\ 4 \end{bmatrix} + (-1)\begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

in general: checking independence = solving system of eq'ns

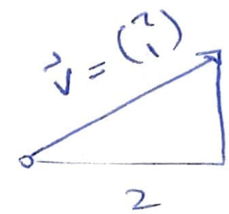
- the <u>span</u> of $\vec{v}_1, \ldots, \vec{v}_k$ is the set of all vectors $\vec{v}$ that are linear combos of $\vec{v}_1, \ldots, \vec{v}_k$

- for vectors $\vec{u}, \vec{v}$ over $\mathbb{R}$, inner product is $\vec{u}^T \vec{v} = [\quad][\,] = \_$

$\underset{1\times n}{\vec{u}^T} \ \underset{n \times 1}{\vec{v}} = \text{scalar}$

(← over $\mathbb{C}$ it's $\vec{u}^* \vec{v}$ ← transpose of complex conjugate of $\vec{u}$)

- $\|\vec{v}\|$ = length of $\vec{v} = \sqrt{\vec{v}^T \vec{v}}$



$\|v\| = \sqrt{(2\ 1)\binom{2}{1}}$
$= \sqrt{2^2 + 1^2}$
$= \sqrt{5}$

- All vector spaces for us of form $\mathbb{R}^n$, $\mathbb{C}^n$ or subspaces of these.
- a <u>subspace</u> V of e.g. $\mathbb{R}^n$ is a subset $V \subseteq \mathbb{R}^n$ closed under linear combos: if $\vec{v}_1, \dots, \vec{v}_k \in V$ then $c_1\vec{v}_1 + \dots + c_k\vec{v}_k \in V$ for any $c_1, \dots, c_k \in \mathbb{R}$
- span$(\{\vec{v}_1, \dots, \vec{v}_k\})$ always a subspace (why?)
- a <u>basis</u> for V is an independent set $\{\vec{v}_1, \dots, \vec{v}_k\} \subseteq V$ s.t. for any other $\vec{v} \in V$, $\{\vec{v}_1, \dots, \vec{v}_k, \vec{v}\}$ is dependent.
 ← i.e. $\vec{v}$ is linear combo of $\vec{v}_1, \dots, \vec{v}_k$

<u>Fact</u>: all bases for a vector space have same cardinality.

(5)

- $\dim V$ = dimension of V = size of any basis
- e.g. $\dim \mathbb{R}^n = n$, $\dim \{\vec{0}\} = 0$
  ↖ legit subspace

- if $\{\vec{v}_1, \dots, \vec{v}_k\}$ independent, form a basis for $V = \text{span}(\{\vec{v}_1, \dots, \vec{v}_k\})$ which has dim k.
- standard basis for $\mathbb{R}^n$ is $\{\vec{e}_1, \dots, \vec{e}_n\}$ where $\vec{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ ← ith place

---

- a function $T: \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if $T(c\vec{u} + d\vec{v}) = cT(\vec{u}) + dT(\vec{v})$ $\forall \vec{u}, \vec{v} \in \mathbb{R}^n$ $\forall c, d \in \mathbb{R}$.
- Given $m \times n$ A, if we define a map $\mathbb{R}^n \to \mathbb{R}^m$ by rule $\vec{v} \to A\vec{v}$, then this map is a linear transformation since:

$$A(c\vec{u} + d\vec{v}) = c(A\vec{u}) + d(A\vec{v})$$

(using properties of matrix multiplication)

Fact: every linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ can be represented as a matrix transformation for some $m \times n$ A.

(why?)