ECS130

Introduction

Monday, January 7, 2019

About

Course: ECS130 Scientific Computing

Professor: Zhaojun Bai

Webpage: http://web.cs.ucdavis.edu/~bai/ECS130/

Today's Agenda

Mathematics Review: Linear Algebra

Vector spaces over $\mathbb{R}$

Denote a (abstract) vector by $\vec{v}$. A vector space

$$\mathcal{V} = \{\text{a collection of vectors } \vec{v}\}$$

which satisfies

- All $\vec{v}, \vec{w} \in \mathcal{V}$ can be *added* and *multiplied by* $a \in \mathbb{R}$:

$$\vec{v} + \vec{w} \in \mathcal{V}, \quad a \cdot \vec{v} \in \mathcal{V}$$

- The operations '$+, \cdot$' must satisfy the *axioms*:

For arbitrary $\vec{u}, \vec{v}, \vec{w} \in \mathcal{V}$,

1. '$+$' commutativity and associativity: $\vec{v} + \vec{w} = \vec{w} + \vec{v}$, $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$.
2. Distributivity: $a(\vec{v} + \vec{w}) = a\vec{v} + a\vec{w}$, $(a + b)\vec{v} = a\vec{v} + b\vec{v}$, for all $a, b \in \mathbb{R}$.
3. '$+$' identity: there exists $\vec{0} \in \mathcal{V}$ with $\vec{0} + \vec{v} = \vec{v}$.
4. '$+$' inverse: for any $\vec{v} \in \mathcal{V}$, there exists $\vec{w} \in \mathcal{V}$ with $\vec{v} + \vec{w} = 0$.
5. '$\cdot$' identity: $1 \cdot \vec{v} = \vec{v}$.
6. '$\cdot$' compatibility: for all $a, b \in \mathbb{R}$, $(ab) \cdot \vec{v} = a \cdot (b \cdot \vec{v})$.

Example

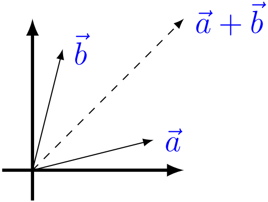
- Euclidean space:

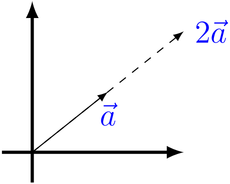
$$\mathbb{R}^n = \left\{ \vec{a} \equiv (a_1, a_2, \ldots, a_n) \colon a_i \in \mathbb{R} \right\}.$$

  - Addition:
    $(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$
  - Multiplication:

$$c \cdot (a_1, \ldots, a_n) = (ca_1, \ldots, ca_n)$$

- Illustration in $\mathbb{R}^2$:





Example

- Polynomials:

$$\mathbb{R}[x] = \left\{ p(x) = \sum_i a_i x^i \colon a_i \in \mathbb{R} \right\}.$$

- Addition and multiplication in the usual way, e.g. $p(x) = a_0 + a_1 x + a_2 x^2$, $q(x) = b_1 x$:
  - Addition:

$$p(x) + q(x) = a_0 + (a_1 + b_1)x + a_2 x^2.$$

  - Multiplication:

$$2p(x) = 2a_0 + 2a_1 x + 2a_2 x^2.$$

Span of vectors

- Start with $\vec{v}_1, \ldots, \vec{v}_n \in \mathcal{V}$, and $a_i \in \mathbb{R}$, we can define

$$\vec{v} \equiv \sum_{i=1}^{n} a_i \vec{v}_i = a_1 \vec{v}_1 + a_2 \vec{v}_2 + \cdots + a_n \vec{v}_n,$$

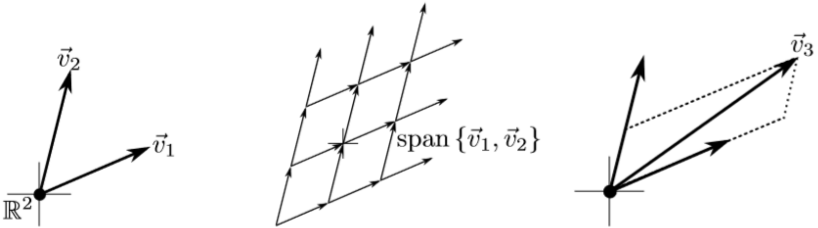
  Such a $\vec{v}$ is called a *linear combination* of $\vec{v}_1, \ldots, \vec{v}_n$.
- For a set of vectors

$$S = \{\vec{v}_i \colon i \in \mathcal{I}\},$$

  all its linear combinations define

$$\operatorname{span} S \equiv \left\{ \sum_i a_i \vec{v}_i \colon \vec{v}_i \in S \text{ and } a_i \in \mathbb{R} \right\}$$

Example in $\mathbb{R}^2$



(a) $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^2$ (b) span $\{\vec{v}_1, \vec{v}_2\}$ (c) span $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$

- Observation from (c): adding a new vector does not always increase the span.

Linear dependence

- A set S of vectors is *linearly dependent* if it contains a vector

  $$\vec{v} = \sum_{i=1}^{k} c_i \vec{v}_i, \quad \text{for some } v_i \in S \backslash \{\vec{v}\} \text{ and nonzero } c_i \in \mathbb{R}.$$

- Otherwise, S is called *linearly independent*.
- Two other equivalent defs. of linear dependence:
  - There exists $\{\vec{v}_1, \dots, \vec{v}_k\} \subset S \backslash \{\vec{0}\}$ such that

    $$\sum_{i=1}^{k} c_i \vec{v}_i = 0 \quad \text{where } c_i \neq 0 \text{ for all } i.$$

  - There exists $\vec{v} \in S$ such that

    $$\operatorname{span} S = \operatorname{span}(S \backslash \{\vec{v}\}).$$

Dimension and basis

- Given a vector space $\mathcal{V}$, it is natural to build a finite set of linearly independent vectors:

$$\{\vec{v}_1, \ldots, \vec{v}_n\} \subset \mathcal{V}.$$

- The max number n of such vectors defines the *dimension* of $\mathcal{V}$.
- Any set S of such vectors is a basis of $\mathcal{V}$, and satisfies

$$\operatorname{span} S = \mathcal{V}.$$

Examples

- The standard basis for $\mathbb{R}^n$ is given by the n vectors

$$\vec{e}_i = (\underbrace{0, \dots, 0}_{i-1}, 1, \underbrace{0, \dots, 0}_{n-i}) \quad \text{for } i = 1, \dots, n$$

  Since
  - $\vec{e}_i$ is not linear combination of the rest of vectors.
  - For all $\vec{c} \in \mathbb{R}^n$, we have $\vec{c} = \sum_{i=1}^{n} c_i \vec{e}_i$.

  Hence, the dimension of $\mathbb{R}^n$ is n.

- A basis of polynomials $\mathbb{R}[x]$ is given by monomials

$$\{1, x, x^2, \dots\}.$$

  The dimension of $\mathbb{R}[x]$ is ∞.

More about $\mathbb{R}^n$

- Dot product: for $\vec{a} = (a_1, \ldots, a_n), \vec{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$

$$\vec{a} \cdot \vec{b} = \sum_{i=1}^{n} a_i b_i.$$

- Length of a vector

$$\|a\|_2 = \sqrt{a_1^2 + \cdots + a_n^2} = \sqrt{\vec{a} \cdot \vec{a}}.$$

- Angle between two vectors

$$\theta = \arccos \frac{\vec{a} \cdot \vec{b}}{\|a\|_2 \|b\|_2}.$$

(*Motivating trigonometric in $\mathbb{R}^3$: $\vec{a} \cdot \vec{b} = \|a\|_2 \|b\|_2 \cos\theta$.)

- Vectors $\vec{a}, \vec{b}$ are *orthogonal* if $\vec{a} \cdot \vec{b} = 0 = \cos 90°$.

Linear function

- Given two vector spaces $\mathcal{V}, \mathcal{V}'$, a function

$$\mathcal{L} \colon \mathcal{V} \to \mathcal{V}'$$

is *linear*, if it preserves *linearity*.

- Namely, for all $\vec{v}_1, \vec{v}_2 \in \mathcal{V}$ and $c \in \mathbb{R}$,
  - $\mathcal{L}[\vec{v}_1 + \vec{v}_2] = \mathcal{L}[\vec{v}_1] + \mathcal{L}[\vec{v}_2]$.
  - $\mathcal{L}[c\vec{v}_1] = c\mathcal{L}[\vec{v}_1]$.
- $\mathcal{L}$ is completely defined by its action on a basis of $\mathcal{V}$:

$$\mathcal{L}[\vec{v}] = \sum_i c_i \mathcal{L}[\vec{v}_i],$$

where $\vec{v} = \sum_i c_i \vec{v}_i$ and $\{\vec{v}_1, \vec{v}_2, \ldots\}$ is a basis of $\mathcal{V}$.

Examples

- Linear map in $\mathbb{R}^n$:

$$\mathcal{L}\colon \mathbb{R}^2 \to \mathbb{R}^3$$

  defined by

$$\mathcal{L}[(x,y)] = (3x, 2x+y, -y).$$

- Integration operator: linear map

$$\mathcal{L}\colon \mathbb{R}[x] \to \mathbb{R}[x]$$

  defined by

$$\mathcal{L}[p(x)] = \int_0^1 p(x)dx.$$

Matrix

- Write vectors in $\mathbb{R}^m$ in '*column forms*', e.g.,

$$\vec{v}_1 = \begin{bmatrix} v_{11} \\ \vdots \\ v_{m1} \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} v_{12} \\ \vdots \\ v_{m2} \end{bmatrix}, \ldots, \vec{v}_n = \begin{bmatrix} v_{1n} \\ \vdots \\ v_{mn} \end{bmatrix}.$$

- Put n columns together we obtain an $m \times n$ *matrix*

$$V \equiv \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \ldots & v_{1n} \\ v_{21} & v_{22} & \ldots & v_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ v_{m1} & v_{m2} & \ldots & v_{mn} \end{bmatrix}$$

- The space of all such matrices is denoted by $\mathbb{R}^{m \times n}$.

Unified notation: Scalars, Vectors, and Matrices

- A scalar $c \in \mathbb{R}$ is viewed as a 1×1 matrix

$$c \in \mathbb{R}^{1 \times 1}.$$

- A column vector $\vec{v} \in \mathbb{R}^n$ is viewed as an $n \times 1$ matrix

$$\vec{v} \in \mathbb{R}^{n \times 1}.$$

Matrix vector multiplication

- A matrix $V \in \mathbb{R}^{m \times n}$ can be multiplied by a vector $\vec{c} \in \mathbb{R}^n$:

$$\begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n.$$

- Elementwisely, we have

$$\begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ v_{m1} & v_{m2} & \dots & v_{mn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} c_1 v_{11} + c_2 v_{12} + \cdots + c_n v_{1n} \\ c_1 v_{21} + c_2 v_{22} + \cdots + c_n v_{2n} \\ \vdots \\ c_1 v_{m1} + c_2 v_{m2} + \cdots + c_n v_{mn} \end{bmatrix}.$$

Using matrix notation

- Matrix vector multiplication can be denoted by

$$\underbrace{A}_{\mathbb{R}^{m\times n}}\underbrace{\vec{x}}_{\mathbb{R}^{n}} = \underbrace{\vec{b}}_{\mathbb{R}^{m}}.$$

- $M \in \mathbb{R}^{m\times n}$ multiplied by another matrix in $\mathbb{R}^{n\times k}$ can be defined as

$$M[\vec{c}_1,\ldots,\vec{c}_k] \equiv [M\vec{c}_1,\ldots,M\vec{c}_k].$$

Example

- Identity matrix

$$I_n \equiv \begin{bmatrix} | & | & & | \\ \vec{e}_1 & \vec{e}_2 & \dots & \vec{e}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}.$$

It holds

$$I_n \vec{c} = \vec{c} \text{ for all } \vec{c} \in \mathbb{R}^n.$$

Example

- Linear map $\mathcal{L}[(x,y)] = (3x, 2x+y, -y)$ satisfies

$$\mathcal{L}[(x,y)] = \underbrace{\begin{bmatrix} 3 & 0 \\ 2 & 1 \\ 0 & -1 \end{bmatrix}}_{\mathbb{R}^{3\times 2}} \cdot \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\mathbb{R}^2} = \underbrace{\begin{bmatrix} 3x \\ 2x+y \\ -y \end{bmatrix}}_{\mathbb{R}^3}.$$

- All linear maps $\mathcal{L} \colon \mathbb{R}^n \to \mathbb{R}^m$ can be expressed as

$$\mathcal{L}[\vec{x}] = A\vec{x},$$

for some matrix $A \in \mathbb{R}^{m\times n}$.

Matrix transpose

- Use A_{ij} to denote the element of A at row i column j.
- The transpose of $A \in \mathbb{R}^{m \times n}$ is defined as $A^T \in \mathbb{R}^{n \times m}$

$$(A^T)_{ij} = A_{ji}.$$

Example:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \qquad \Rightarrow \qquad A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

- Basic identities:

$$(A^T)^T = A, \quad (A+B)^T = A^T + B^T, \quad (AB)^T = B^T A^T.$$

Examples: Matrix operations with transpose

- Dot product of $\vec{a}, \vec{b} \in \mathbb{R}^n$:

$$\vec{a} \cdot \vec{b} = \sum_{i=1}^{n} a_i b_i = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \vec{a}^T \vec{b}.$$

- Residual norms of $\vec{r} = A\vec{x} - \vec{b}$:

$$\begin{aligned} \|A\vec{x} - \vec{b}\|_2^2 &= (A\vec{x} - \vec{b})^T (A\vec{x} - \vec{b}) \\ &= (\vec{x}^T A^T - \vec{b}^T)(A\vec{x} - \vec{b}) \\ &= \vec{b}^T \vec{b} - \vec{b}^T A\vec{x} - \vec{x}^T A^T \vec{b} + \vec{x}^T A^T A\vec{x} \\ (\text{by } \vec{b}^T A\vec{x} = \vec{x}^T A^T \vec{b}) \quad &= \|\vec{b}\|_2^2 - 2\vec{b}^T A\vec{x} + \|A\vec{x}\|_2^2. \end{aligned}$$

Computation aspects

- Storage of matrices in memory:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \Rightarrow \begin{cases} \text{Row-major:} & \boxed{1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6} \\ \text{Column-major:} & \boxed{1 \mid 3 \mid 5 \mid 2 \mid 4 \mid 6} \end{cases}$$

- Multiplication $\vec{b} = A\vec{x}$ for $A \in \mathbb{R}^{m \times n}$ and $\vec{x} \in \mathbb{R}^n$:

Access A row-by-row:

```
b⃗ = 0
for i = 1, ..., m do
    for j = 1, ..., n do
        b_i = b_i + A_ij x_j
    end for
end for
```

Access column-by-column:

```
b⃗ = 0
for j = 1, ..., n do
    for i = 1, ..., m do
        b_i = b_i + A_ij x_j
    end for
end for
```

Linear systems of equations in matrix form

- **Example**: find (x, y, z) satisfying

$$\begin{aligned} 3x + 2y + 5z &= 0 \\ -4x + 9y - 3z &= -7 \\ 2x - 3y - 3z &= 1. \end{aligned} \quad \Rightarrow \quad \begin{bmatrix} 3 & 2 & 5 \\ -4 & 9 & -3 \\ 2 & -3 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \\ 1 \end{bmatrix}$$

- Given $A = [\vec{a}_1, \ldots, \vec{a}_n] \in \mathbb{R}^{m \times n}$, $\vec{b} \in \mathbb{R}^m$, find $\vec{x} \in \mathbb{R}^n$:

$$A\vec{x} = \vec{b}.$$

- Solution exists if $\vec{b}$ is in *column space* of A:

$$\vec{b} \in \operatorname{col} A \equiv \{A\vec{x} \colon \vec{x} \in \mathbb{R}^n\} = \left\{ \sum_{i=1}^{n} x_i \vec{a}_i \colon x_i \in \mathbb{R} \right\}.$$

The dimension of $\operatorname{col} A$ is defined as the *rank* of A.

The square case

- Let $A \in \mathbb{R}^{n \times n}$ be a square matrix, and suppose $A\vec{x} = \vec{b}$ has solution for all $\vec{b} \in \mathbb{R}^n$. We can solve

$$A\vec{x}_i = \vec{e}_i, \quad \text{for } i = 1, \ldots, n.$$

$$\Updownarrow$$

$$A \underbrace{\begin{bmatrix} \vec{x}_1 & \vec{x}_2 & \ldots & \vec{x}_n \end{bmatrix}}_{A^{-1}} = I_n$$

- The *inverse* satisfies (*why?*)

$$AA^{-1} = A^{-1}A = I_n \quad \text{and} \quad (A^{-1})^{-1} = A.$$

- Hence, for any $\vec{b}$, we can express the solution as

$$\vec{x} = A^{-1}A\vec{x} = A^{-1}\vec{b}.$$