Lecture 3: Linear Algebra Review

Baboo J. Cui, Yangang Cao

July 8, 2019

Contents

1 Vector Spaces

A (real) vector space V is a set with **two** operations(also known as close to linear operation):

- vector sum: $V + V \to V$
- scalar multiplication: $\mathbb{R} \times V \to V$

that has the following (7) properties:

1. commutativity: $x + y = y + x, \forall x, y \in V$
2. associativity: $(x + y) + z = x + (y + z), \forall x, y, z \in V$
3. zero element: $\exists! 0 \in V$ such that $0 + x = x, \forall x \in V$, $\exists!$ means **only exist**
4. inverse: $\forall x \in V, \exists (-x) \in V$ such that $x + (-x) = 0$
5. associativity in scalar product: $(\alpha\beta)x = \alpha(\beta x), \forall \alpha, \beta \in \mathbb{R}, x \in V$
6. distributivity: $\alpha(x + y) = \alpha x + \alpha y, \forall a \in \mathbb{R}, x, y \in V$
7. distributivity: $(\alpha + \beta)x = \alpha x + \beta x, \forall \alpha, \beta \in \mathbb{R}, x \in V$

The properties implies that:

$$1 \cdot x = x \text{ and } 0 \cdot \vec{x} = \vec{0}, \forall x \in V$$

1.1 Example of Vector Spaces

- $\mathbb{R}^n$
- $\mathbb{R}^{m \times n}$
- P_n : the set of all polynomials in λ with degree up to n, note that DOF is $n + 1$
- $\mathcal{F}(\mathcal{I}; \mathbb{R}^n)$: Set of all mappings from an index set $\mathcal{I}$ to $\mathbb{R}^n$
- set of all differentiable function $f : \mathbb{R}_+ \to \mathbb{R}$
- set of all square integrable function $f : \mathbb{R}_+ \to \mathbb{R}$
- set of all solutions to an autonomous LTI system

1.2 Subspaces and Product Spaces

- **subspace**: W is a subspace of vector space V if $W \subset V$ and W itself is a vector space under the same vector sum and scalar multiplication operations
- **product space**: given two vector spaces V_1 and V_2, their **direct product** is the vector space $V_1 \times V_2 := \{(v_1, v_2) \,| v_1 \in V_1, v_2 \in V_2\}$, essentially, link two vectors together

1.3 Bases and Dimension of Vector Spaces

$\{v_1, \dots, v_k\}$ in vector space V are linearly independent if for $\alpha_1, \dots, \alpha_k \in \mathbb{R}$,

$$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0 \quad \Rightarrow \quad \alpha_1 = \cdots = \alpha_k = 0$$

A set of vectors $\{v_1, \dots, v_k\}$ is a **basis** of the vector space V if

- $v_1, \dots, v_k$ are linearly independent in V
- $V = \operatorname{span}\{v_1, \dots, v_k\}$

Or equivalently,

- each $v \in V$ has a **unique** expression $v = \alpha_1 v_1 + \cdots + \alpha_k v_k$
- $(\alpha_1, \dots, \alpha_k)$ is the coordinate of v in this basis

The **dimension** of a vector space V is the number of vectors in any of its basis, and is denoted $\dim V$.

2 Linear Maps

A map $f : V \to W$ between two vector spaces V and W is linear if

$$f(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 f(v_1) + \alpha_2 f(v_2)$$

- A linear map $f : V \to W$ must map $0 \in V$ to $0 \in W$, also known as ZIZO
- The composition of two linear maps $f : V \to W$ and $g : W \to U$ is also linear: $g \circ f : v \in V \mapsto g(f(v)) \in U$

2.1 Null Spaces and Images of Linear Maps

- **null space**: the null space of a linear map $f : V \to W$ is $\mathcal{N}(f) := \{v \in V | f(v) = 0\}$, note that $\mathcal{N}(f)$ is a subspace of V
- **image(range)**: the image (or range) of a linear map $f : V \to W$ is

  $$\mathcal{R}(f) := \{w \in W | w = f(v) \text{ for some(some is enough) } v \in V\}$$

  $\mathcal{R}(f)$ is a subspace of W.

2.2 Injective (One-To-One) Linear Maps

A linear map $f : V \to W$ is **injective** (one-to-one) if for all $v_1, v_2 \in V$,

$$f(v_1) = f(v_2) \Rightarrow v_1 = v_2$$

Equivalent definitions:

- f maps different vectors to different vectors
- f maps linearly independent vectors to linearly independent vectors
- $\mathcal{N}(f) = \{0\}$

Matrix $A \in \mathbb{R}^{m \times n}$ considered as a linear map $\mathbb{R}^n$ to $\mathbb{R}^m$ has null space

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n | Ax = 0\}$$

$A \in \mathbb{R}^{m \times n}$ is **one-to-one** if and only if

- Columns of A are linearly independent
- Rows of A span $\mathbb{R}^n$
- A has rank n (full column rank)
- A has a left inverse: $\exists B \in \mathbb{R}^{n \times m}$ such that $BA = I_n$
- $\det(A^T A) \neq 0$

2.3 Surjective (Onto) Linear Maps

A linear map $f : V \to W$ is surjective (onto) if $\mathcal{R}(f) = W$, or equivalently, if for any $w \in W, w = f(v)$ for some $v \in V$. Matrix $A \in \mathbb{R}^{m \times n}$ considered as a linear map $\mathbb{R}^n$ to $\mathbb{R}^m$ has range space $\mathcal{R}(A) = \{Ax \in \mathbb{R}^m | x \in \mathbb{R}^n\}$, $A \in \mathbb{R}^{m \times n}$ is **onto** if and only if

- rows of A are linearly independent
- columns of A span $\mathbb{R}^m$
- rank of A is m (full row rank)
- A has a right inverse $\exists B \in \mathbb{R}^{n \times m}$ such that $AB = I_m$

2.4 Bijective (Invertible) Linear Maps

A linear map $f : V \to W$ is bijective (invertible) if it is both one-to-one and onto. Its inverse is the unique map $f^{-1} : W \to V$ such that $f \circ f^{-1} = id_W$ and $f^{-1} \circ f = id_V$, V and W must have the same dimension, a matrix $A \in \mathbb{R}^{n \times n}$ is invertible if $A : \mathbb{R}^n \to \mathbb{R}^n$ is bijective, it is equivalent to:

- columns (or rows) of A form a basis of $\mathbb{R}^n$
- A has inverse A^{-1} with $AA^{-1} = A^{-1}A = I_n$
- $\mathcal{N}(A) = \{0\}$
- $\mathcal{R}(A) = \mathbb{R}^n$
- $\det A \neq 0$

3 Matrix

3.1 Matrix Rank

The rank of a matrix $A \in \mathbb{R}^{m \times n}$ is its maximum number of linearly independent columns (or rows), or equivalently, $\dim \mathcal{R}(A)$, it has the following properties:

- $\text{Rank}(A) \leq \min(m, n)$

- $\text{Rank}(A) = \text{Rank}\left(A^T\right)$
- $\text{Rank}(A) + \dim \mathcal{N}(A) = n$ (conservation of dimension)

Full rank matrix $A \in \mathbb{R}^{m \times n}$: $\text{Rank}(A) = \min(m, n)$

- (for skinny matrices) independent column or injective maps
- (for fat matrices) independent rows or surjective maps
- (for square matrices) nonsingular or bijective maps

3.2 Matrix Transpose

When $A \in \mathbb{R}^{m \times n}$ is considered as a linearmap from $\mathbb{R}^n$ to $\mathbb{R}^m$, its transpose $A^T \in \mathbb{R}^{n \times m}$ is a linear map from $\mathbb{R}^m$ back to $\mathbb{R}^n$

The following statements are equivalent:

- A is one-to-one
- A^T is onto
- $\det A^T A \neq 0$
- $A^T A \in \mathbb{R}^{n \times n}$ is bijective

The following statements are equivalent:

- A is onto
- A^T is one-to-one
- $\det AA^T \neq 0$
- $AA^T \in \mathbb{R}^{m \times m}$ is bijective

More generally, for any $A \in \mathbb{R}^{m \times n}$

- $\mathcal{R}\left(A^T\right) = \mathcal{N}(A)^{\perp}$
- $\mathcal{N}\left(A^T\right) = \mathcal{R}(A)^{\perp}$

4 Inner Product on Euclidean Space

For $x, y \in \mathbb{R}^n$, their inner product is

$$\langle x, y \rangle := x^T y = y^T x = x_1 y_1 + \cdots + x_n y_n$$

For $x, y, z \in \mathbb{R}^n$:

- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$
- $\langle x + y, z \rangle = \langle x, z \rangle + \langle x, y \rangle$

- $\langle x, x\rangle = \|x\|^2 \geq 0$, where $||x||$ is the Euclidean norm of x:

$$\|x\| := \sqrt{x^T x} = \sqrt{x_1^2 + \cdots + x_n^2}$$

Cauchy-Schwartz Inequality:

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|, \quad \forall x, y \in \mathbb{R}^n$$

5 Finite Dimensional Vector Space vs $\mathbb{R}^n$

There is a **bijection** between any finite dimensional vector space V and $\mathbb{R}^n$ with

$$n = \dim V$$

Proof: coordinate of a vector with basis change.

6 More about Matrix

6.1 Matrix Representation of Linear Maps

Any linear map $f : V \to W$ between two finite dimensional vector spaces can be represented as a matrix $A \in \mathbb{R}^{m\times n}$ with $n = \dim V, m = \dim W$. Example:

- For $A \in \mathbb{R}^{m\times n}$, the map $x \in \mathbb{R}^n \mapsto Ax \in \mathbb{R}^m$ viewed in standard basis

$$\mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$$

- $\frac{d}{d\lambda} : p(\lambda) \in \mathcal{P}_n \mapsto \frac{dp(\lambda)}{d\lambda} \in \mathcal{P}_{n-1}$, the matrix is:

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & n-1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

6.2 Determinant and Inverse of Square Matrices

For any square matrix $A \in \mathbb{R}^{n\times n}$, its **determinant** is defined recursively as

$$\det A := \sum_{i=1}^{n} a_{ij} c_{ij}$$

- a_{ij} : entry of A on row i and column j
- $c_{ij} = (-1)^{i+j} \det M_{ij}$: cofactor corresponding to a_{ij}
- measures the volume amplification of linear map A

For nonsingular matrices ($\det A \neq 0$), the **inverse matrix** of $A \in \mathbb{R}^{n\times n}$ is the unique matrix $A^{-1} \in \mathbb{R}^{n\times n}$ satisfying $AA^{-1} = A^{-1}A = I_n$:

$$A^{-1} = \frac{\operatorname{Adj} A}{\det A} = \frac{[c_{ij}]^T}{\det A}$$

6.3 Spectrum of Square Matrices

The **characteristic polynomial** of a square matrix $A \in \mathbb{R}^{n\times n}$ is

$$\chi_A(\lambda) := \det(\lambda I_n - A) \in \mathcal{P}_n$$

The n roots (counting multiplicity, possibly complex) of $\chi_A(\lambda)$ are the **eigenvalues** of A. The **spectrum** of A is the set $\sigma(A)$ of all its eigenvalues. For each eigenvalue $\lambda_i \in \mathbb{C}$ of A,

- $v_i \in \mathbb{C}^n$ is called a (right) **eigenvector** if $Av_i = \lambda_i v_i$
- $w_i \in \mathbb{C}^n$ is called a **left eigenvector** if $w_i^T A = \lambda_i w_i^T$

7 Change of Basis in $\mathbb{R}^n$

A vector $x = [x_1 \cdots x_n]^T \in \mathbb{R}^n$ in standard basis has the coordinate in new basis $\{t_1, \cdots, t_n\}$:

$$\tilde{x} = T^{-1}x = \begin{bmatrix} t_1 & \cdots & t_n \end{bmatrix}^{-1} x \quad \text{this is } T^{-1}$$

$A \in \mathbb{R}^{n\times n}$ as a linear map in standard basis when viewed in a different basis $\{t_1, \ldots, t_n\}$ has matrix representation:

$$\tilde{A} = T^{-1}AT$$

$$\begin{array}{l} \mathbb{R}^n \xrightarrow{A} \mathbb{R}^n \\ \downarrow T^{-1} \quad\; \downarrow T^{-1} \\ \mathbb{R}^n \xrightarrow{\tilde{A}} \mathbb{R}^n \end{array}$$

Two matrices $A, \tilde{A} \in \mathbb{R}^{n\times n}$ are **similar** if there exists a nonsingular matrix $T \in \mathbb{R}^{n\times n}$ such that

$$\tilde{A} = T^{-1}AT$$

- representing the same linear map viewed in different bases
- determinant is invariant: $\det A = \det \tilde{A}$
- spectrum is invariant: $\sigma(A) = \sigma(\tilde{A})$

Matrix $A \in \mathbb{R}^{n\times n}$ is called **diagonalizable** if there exists a nonsingular matrix $T \in \mathbb{C}^{n\times n}$ such that $T^{-1}AT = \Lambda \in \mathbb{C}^{n\times n}$ is diagonal.

- Diagonal entries of $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ are the eigenvalues of A
- Column of $T = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$ are the right eigenvectors of A
- Rows of $T^{-1} = \begin{bmatrix} w_1 & \ldots & w_n \end{bmatrix}^T$ are (the transpose of) the left eigenvectors of A

Diagonalizable matrix $A \in \mathbb{R}^{n\times n}$ with $T^{-1}AT = \Lambda$ can be decomposed by Dyadic expansion as

$$A = \lambda_1 v_1 w_1^T + \lambda_2 v_2 w_2^T + \cdots \lambda_n v_n w_n^T$$

which is the sum of n rank-one matrices.

8 Jordan Canonical Form

For any $A \in \mathbb{R}^{n\times n}$, there exists a nonsingular $T \in \mathbb{C}^{n\times n}$ such that

$$T^{-1}AT = J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_q \end{bmatrix}, \quad J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix} \in \mathbb{C}^{n_i\times n_i}$$

- Unique up to permutation of Jordan blocks
- Diagonalizable matrices are special cases with all $n_i = 1$

The **algebraic multiplicity** of an eigenvalue λ_i is **the sum of the sizes** of all Jordan blocks corresponding to it, its **geometric multiplicity** is the number of all such Jordan blocks.

9 Geometric Characterization of Jordan Block Sizes

Given $A \in R^n$ with an eigenvalue λ, construct a cascade of subspaces:

$$\mathcal{N}\left(A-\lambda I_n\right) \subset \mathcal{N}\left[\left(A-\lambda I_n\right)^2\right] \subset \mathcal{N}\left[\left(A-\lambda I_n\right)^3\right] \subset \cdots$$

The geometric multiplicity of eigenvalue λ is

- $\dim \mathcal{N}\left(A-\lambda I_n\right)$
- The number of linearly independent eigenvectors corresponding to λ

In general, the number of Jordan blocks of λ with size at least k is

$$\dim \mathcal{N}\left[\left(A-\lambda I_n\right)^k\right] - \dim \mathcal{N}\left[\left(A-\lambda I_n\right)^{k-1}\right]$$

An example should be added here(coming soon)!

10 Generalized Eigenvectors

Vector v is a generalized eigenvector of A of grade d if

$$v \in \mathcal{N}(A-\lambda I_n)^d \quad \text{and} \quad v \notin \mathcal{N}(A-\lambda I_n)^{d-1}$$

When $d=1$, this reduces to the definition of eigenvectors, here is an example $(d=3)$: A chain of generalized eigenvectors of length 3:

$$\{v_3 := v, v_2 := \left(A-\lambda_n\right)v, v_1 := \left(A-\lambda l_n\right)^2 v\}$$

which satisfies

$$Av_1 = \lambda v_1, Av_2 = v_1 + \lambda v_2, Av_3 = v_2 + \lambda v_3$$

v_1, v_2, v_3 are the columns of T corresponding to a Jordan block of size 3.

11 Real Jordan Canonical Form

For any $A \in \mathbb{R}^{n \times n}$, there exists a nonsingular $T \in \mathbb{R}^{n \times n}$ such that:

$$T^{-1}AT = J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_q \end{bmatrix},$$

where $J_i = \begin{bmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_i \end{bmatrix}$ for real λ_i and $J_i = \begin{bmatrix} C_i & I_2 & & \\ & \ddots & \ddots & \\ & & \ddots & I_2 \\ & & & C_i \end{bmatrix}$ with $C_i = \begin{bmatrix} a_i & b_i \\ -b_i & a_i \end{bmatrix}$ for complex $\lambda_i = a_i + b_i\sqrt{-1}$.