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:

- Vector sum $+$: $V + V \to V$
- Scalar multiplication $\cdot$: $\mathbb{R} \times V \to V$

that has the following propertoes:

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 $0 + x = x, \forall x \in V$
4. Inverse: $\forall x \in V, \exists (-x) \in V$ such that $x + (-x) = 0$
5. $(\alpha\beta)x = \alpha(\beta x), \forall \alpha, \beta \in \mathbb{R}, x \in V$
6. $\alpha(x + y) = \alpha x + \alpha y, \forall a \in \mathbb{R}, x, y \in V$
7. $(\alpha + \beta)x = \alpha x + \beta x, \forall \alpha, \beta \in \mathbb{R}, x \in V$

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

2 Example of Vector Spaces

1. $\mathbb{R}^n$
2. $\mathbb{R}^{m \times n}$
3. P_n : the set of all polynomials in λ with degree up to n
4. $\mathcal{F}(\mathcal{I}; \mathbb{R}^n)$: Set of all mappings from an index set $\mathcal{I}$ to $\mathbb{R}^n$
5. Set of all differentiable function $f : \mathbb{R}_+ \to \mathbb{R}$
6. Set of all square integrable function $f : \mathbb{R}_+ \to \mathbb{R}$
7. Set of all solutions to an autonomous LTI system

3 Subspaces and Product Spaces

Definition (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 multiplicatien operations

Definition (Product space):
Given two vector spaces V_1 and V_2, their direct product is the vector space $V_1 \times V_2 := \{(v_1, v_2) \,|v_1 \in V_1, v_2 \in V_2\}$

4 Bases and Dimension of Vector Spaces

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

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

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

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

Or equivalently,

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

Definition (Dimension):
The dimension of a vector space V is the number of vectors in any of its basis, and is denoted $\dim V$.

5 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 $0 \in W$
- The composition of two linear maps $f : V \to W$ and $g : W \to U$ is also linear: $g \circ f : v \in V \mapsto g(f(v)) \in U$

6 Null Spaces and Images of Linear Maps

Definition (Null Space):
The null space of a linear map $f : V \to W$ is $\mathcal{N}(f) := \{v \in V | f(v) = 0\}$
• $\mathcal{N}(f)$ is a subspace of V

Definition (Image (Range)):
The image (or range) of a linear map $f : V \to W$ is

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

• $\mathcal{R}(f)$ is a subspace of W

7 Injective (One-To-One) Linear Maps

Definition (Injective Map):
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
- $\mathcal{N}(f) = \{0\}$

8 One-To-One Matrices

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 $BA = I_n$
- $\det(A^T A) \neq 0$

9 Surjective (Onto) Linear Maps

Definition (Surjective Map):
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 $v \in V$

10 Surjective Matrices

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 $AB = I_m$

11 Bijective (Invertible) Linear Maps

Definition (Bijective Map):
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

$\Leftrightarrow$ Columns (or rows) of A form a basis of $\mathbb{R}^n$

$\Leftrightarrow$ A has inverse A^{-1} with $AA^{-1} = A^{-1}A = I_n$

$\Leftrightarrow$ $\mathcal{N}(A) = \{0\}$

$\Leftrightarrow$ $\mathcal{R}(A) = \mathbb{R}^n$

$\Leftrightarrow$ $\det A \neq 0$

12 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, $\dim \mathcal{R}(A)$

- $\operatorname{Rank}(A) \leq \min(m, n)$
- $\operatorname{Rank}(A) = \operatorname{Rank}\left(A^T\right)$
- $\operatorname{Rank}(A) + \dim \mathcal{N}(A) = n$ (conservation of dimension)

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

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

13 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 are equivalent:

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

The following 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 $A \in \mathbb{R}^{m\times n}$

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

14 Inner Product on Euclidean Space

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

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

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:

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

Theorem (Cauchy-Schwartz Inequality):

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

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

Theorem:
There is a bijection between any finite dimensional vector space V and $\mathbb{R}^n$ with $n = \dim V$.

Proof:
$\rho_V : v \in V \mapsto$ its coordinate in a basis $\{v_1, \ldots, v_n\}$ of V.

16 Matrix Representation of Linear Maps

Theorem:
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 $n = \dim V, m = \dim W$.

17 Examples

- 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

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

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

18 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 $AA^{-1} = A^{-1}A = I_n$:

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

19 Spectrum of Square Matrices

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

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

Definition (Spectrum of A):
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 $w_i^T A = \lambda_i w_i^T$

Example:

$$\begin{bmatrix} 0.4 & 0.6 \\ 0.7 & 0.3 \end{bmatrix} \qquad \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

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

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

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

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

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

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

21 Similarity Transformations

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: $\sigma(A) = \sigma(\tilde{A})$

22 Diagonalizable Matrices

Definition (Diagonalizable Matrix):
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

23 Dyadic Expansion of Diagonalizable Matrices

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

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

• Sum of n rank-one matrices
Example:

$$\underbrace{\begin{bmatrix} \frac{7}{13} & \frac{6}{13} \\ 1 & -1 \end{bmatrix}}_{T^{-1}} \underbrace{\begin{bmatrix} 0.4 & 0.6 \\ 0.7 & 0.3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 & \frac{6}{13} \\ 1 & -\frac{7}{13} \end{bmatrix}}_{T} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0.3 \end{bmatrix}}_{\Lambda}$$

24 Jordan Canonical Form

Theorem (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 $n_i = 1$

Definition (Algebraic and Geometric Multiplicity): The **algebraic multiplicity** of an eigenvalue λ_i is the sum of the sizes of all Jordan blocks correpsonding to it; its **geometric multiplicity** is the number of all such Jordan blocks.

25 Geometric Characterization of Jordan Block Sizes

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

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

Fact: the geometric multiplicity (# of Jordan blocks) of eigenvalue λ is

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

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

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

26 Example

Suppose that λ_1 is an eigenvalue of the matrix $A \in \mathbb{R}^{10\times 10}$, and that the dimensions of the null space of $\left(A-\lambda_1 I\right)^k$ for $k = 1, 2, 3, 4$ and 5 are: $3, 6, 8, 9$ and 9, respectively.

- How many Jordan blocks are associated with λ_1?
- What are the sizes of these Jordan blocks?
- Does A have other eigenvalues? If yes, what are their multiplicities?

27 Generalized Eigenvectors

Definition (Generalized Eigenvector):
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

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

28 Real Jordan Canonical Form

Theorem (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 $\lambda_i = a_i + b_i\sqrt{-1}$.