Matrix Analysis: Review of linear algebra

Jean-Luc Bouchot

School of Mathematics and Statistics
Beijing Institute of Technology
jlbouchot@bit.edu.cn

2018/11/18

1. Review of linear algebra

Outline

Definition 2.1

A **group** *is a set G together with an operation $\odot$ such that*

- G *is close under* $\odot$*: for all* $a, b \in G, a \odot b \in G$,
- $\odot$ *is associative: for all* $a, b, c \in G$, $(a \odot b) \odot c = a \odot (b \odot c)$,
- G *contains an identity element* e *for* $\odot$*: for all* $a \in G, a \odot e = e \odot a = a$,
- G *is close by inversion: for all* $a \in G$*, there exists a* $b \in G$ *such that* $a \odot b = b \odot a = e$. *(usually written* $-a$ *or* a^{-1}*).*

If moreover $\odot$ *is commutative in* G*, i.e. for all* $a, b \in G$, $a \odot b = b \odot a$*, we say that* $(G, \odot)$ *is* **abelian group**.

Example 2.1

Show whether the following sets are groups or not. Are they abelian groups?

- $C(\mathbb{R}, \mathbb{R})$ the set of continuous functions on $\mathbb{R}$, together with the usual addition: $f + g$ is the function defined on $\mathbb{R}$ such that $(f + g)(x) = f(x) + g(x)$.
- It is also a *multiplicative* group?
- What if we use the composition?
- For a given $N \geq 2$, let $\mathcal{G}_N := \{\omega \in \mathbb{C} : \omega^N = 1\}$. Is it a multiplicative group with the usual scalar multiplication?

Definition 2.2

A **field** *is a set* G *with two operations* $\oplus$ *(usually called the addition) and* $\otimes$ *(the multiplication) such that*

- $(G, \oplus)$ *is an abelian group with (additive) identity* 0_G*,*
- $(G \backslash \{0_G\}, \otimes)$ *is an abelian group with (multiplicative) identity* 1_G*,*
- *the multiplication is distributive over the addition: for all* $a, b, c \in G$*,* $a \otimes (b \oplus c) = (a \oplus b) \otimes (a \oplus c)$.

Definition 2.3

A **vector space** *over a field* $\mathbb{F}$ *(with operations* $\oplus_F$ *and* $\otimes_F$ *and respective identitites* $0_F, 1_F$*) is a set of vectors* V *together with two operations* $\oplus_V$ *(vector addition) and* $\odot_S$ *(the scalar multiplication) such that*

1. $(V, \oplus_V)$ *is an abelian group, with the zero vector* 0_V*,*
2. *for all* $\mathbf{v} \in V$*,* $1_F \odot_S \mathbf{v} = \mathbf{v}$
3. *the scalar multiplication is distributive: for all* $\mathbf{u}, \mathbf{v} \in V$*, for all* $\alpha \in \mathbb{F}, \alpha \odot_S (\mathbf{u} \oplus_V \mathbf{v}) = \alpha \odot_S \mathbf{u} \oplus_V \alpha \odot_S \mathbf{v}$*,*
4. *the scalar multiplication is compatible: for all* $\alpha, \beta \in \mathbb{F}$*, for all* $\mathbf{v} \in V$*,* $\alpha \odot_S (\beta \odot_S \mathbf{v}) = (\alpha \otimes_F \beta) \odot_S \mathbf{v}$*,*
5. *Distributivity of scalar multiplication of the additive field: for all* $\alpha, \beta \in \mathbb{F}$*, and for all* $\mathbf{v} \in V$*,* $(\alpha \oplus_F \beta) \odot_S \mathbf{v} = \alpha \odot_S \mathbf{v} \oplus_V \beta \odot_S \mathbf{v}$*.*

Example 2.2

- Classical vectors $\mathbb{R}^n$, $\mathbb{C}^n$
- $\mathbb{R}_n[x] := \{f(x) = a_0 + a_1 x + \cdots + a_n x^n; (a_0, \cdots, a_n) \in \mathbb{R}^{n+1}\}$
- $\mathbb{R}[x]$?
- $\{(x, y, z)^T : ax + by + cz = 0\}$
- $\{(x, y, z)^T : ax + by + cz = 1\}$

Remark 2.1

It should be clear from the context whether the vector of scalar multiplication / addition is meant. We will therefore drop the subscripts to avoid overcomplicating the notation.
Moreover, the vector space $(V; \mathbb{F})$ will only be denoted V unless there are any ambiguities.

Definition 2.4

A subset $W \subseteq V$ *is a subspace of* V *if*

1. $0_V \in W$
2. *for all* $\mathbf{u}, \mathbf{v} \in W$, $\mathbf{u} + \mathbf{v} \in W$
3. *for all* $\mathbf{v} \in W$ *and* $\alpha \in \mathbb{F}$, $\alpha \mathbf{v} \in W$.

Exercise 2.1

Let U be a vector space and $V, W \subset U$ two subspaces. Are the following sets subspaces of U?

1. $V \cap W := \{\mathbf{u} : \mathbf{u} \in V \text{ and } \mathbf{u} \in W\}$
2. $V \cup W := \{\mathbf{u} : \mathbf{u} \in V \text{ or } \mathbf{u} \in W\}$
3. $V + W := \{\mathbf{u} : \exists \mathbf{v} \in V, \mathbf{w} \in W : \mathbf{u} = \mathbf{v} + \mathbf{w}\}$

Definition 2.5

Let $V \subset U$ *be a subset of* U *(not necessarily a subspace). We define its* **span** *has the intersection of all subsets of* U *which contain* V*. We write* $W = \mathrm{span}(V)$*.* W *is a subspace of* U *(verify this).*

Proposition 2.1

Let $V \subset U$. $\mathrm{span}(V) = \{\sum_{k=1}^{n} \alpha_k \mathbf{v}_k, k = 1, \cdots\}$.

Exercise 2.2

Let $\mathbf{u}$ and $\mathbf{v}$ be two linearly independent vectors. Show that $\text{span}\{\mathbf{u}, \mathbf{v}, \mathbf{u}+\mathbf{v}\} = \text{span}\{\mathbf{u}, \mathbf{v}\} = \text{span}\{\mathbf{u}, \mathbf{u}+\mathbf{v}\}$.

Definition 2.6

Let V be a vector space and $\mathcal{F} = (\mathbf{v}_1, \cdots, \mathbf{v}_n)$ be a family of n vectors in V. We say that the family $\mathcal{F}$ is a **linearly independent** *set of vectors if*

$$\sum_{i=1}^{n} \alpha_i \mathbf{v}_i = 0 \Leftrightarrow \alpha_1 = \cdots = \alpha_n = 0.$$

A family which is not linearly independent is said to be a linearly dependent.

Exercise 2.3

Write down the definition of what it means to be linearly dependent.

Example 2.3

- $((1,0),(0,1))$
- $((1,0),(1,1))$
- $((1,0),(0,1),(1,1))$
- $\big((x \mapsto \cos(x)),(x \mapsto \cos(2x)),(x \mapsto \cos^2(x))\big)$

Exercise 2.4

Consider $V = \mathbb{R}_n[x]$. Are the following families linearly dependent?

- $(1, x, \cdots, x^n)$
- $(1, 1+x, 1+x+x^2, \cdots, 1+x+\cdots+x^{n-1}+x^n)$
- $(1, 1+x, 1+x^2, \cdots, 1+x^n)$
- $(1+x, x+x^2, x^2+x^3, \cdots, x^{n-1}+x^n, x^n+1)$

Definition 2.7

A family $\mathcal{F} = (\mathbf{v}_1, \cdots, \mathbf{v}_n) \subset V$ *is a* **generating family** *or* **spanning set** *if for all* $\mathbf{v} \in V$*, there exists scalars* $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$ *such that* $\mathbf{v} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$*.*

Definition 2.8

A family $\mathcal{F}$ of vectors is a basis if it is a linearly independent spanning set.

Exercise 2.5

Are the following families generating? Linearly independent? Basis?

- $(1, x, \cdots, x^n)$
- $(1, 1 + x, 1 + x + x^2, \cdots, 1 + x + \cdots + x^{n-1} + x^n)$
- $(1, 1 + x, 1 + x^2, \cdots, 1 + x^n)$
- $(1 + x, x + x^2, x^2 + x^3, \cdots, x^{n-1} + x^n, x^n + 1)$

Theorem 2.1

Let V be a vector space and $\mathcal{F} = \{\mathbf{u}_1, \cdots, \mathbf{u}_n\}$ be a basis for V. Then for all $\mathbf{v} \in V$, there exists unique scalars $\alpha_1, \cdots, \alpha_n \in \mathbb{F}$ such that

$$\mathbf{v} = \sum_{i=1}^{n} \alpha_i \mathbf{v}_i.$$

This unique representation gives rise to the notion of **coordinates** *of a vector with respect to a certain basis.*

Theorem 2.2

Let V be a vector space and B and C two basis. Then B and C have the same number of vectors.

Definition 2.9

The **dimension** *of a vector space is the number of vectors in any of its basis. We write* $\dim(V) = n$. *A vector space can be*

- *Finite dimensional if* $\dim(V) < \infty$, *or*
- *Infinite dimensional if* $\dim(V) = \infty$.

Exercise 2.6

What is the dimension of the following vector spaces:

- $\mathbb{R}_n[x]$
- $\mathbb{R}[x]$
- $\mathbb{R}^n$
- $\mathbb{C}^n$

Theorem 2.3

Let V be a finite dimensional vector space with $\dim(V) = n < \infty$ and let $S = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$. The following statements are equivalent:

1. *S is a basis for V.*
2. *S is a spanning set.*
3. *S is linearly independent.*

Definition 2.10

Let U *and* V *be two vector spaces over the same field* $\mathbb{F}$*. A map* $f : U \to V$ *is said to be a* **linear map** *if*

- *for all* $\mathbf{u}, \mathbf{v} \in U$, $f(\mathbf{u} +_U \mathbf{v}) = f(\mathbf{u}) +_V f(\mathbf{v})$,
- *for all* $\alpha \in \mathbb{F}$ *and* $\mathbf{u} \in U$, $f(\alpha \mathbf{u}) = \alpha f(\mathbf{u})$.

Example 2.4

- $x \mapsto 2x, \alpha x$
- For a given vector $\mathbf{a} \in \mathbb{K}^n$, the map $T_{\mathbf{a}} : \mathbb{K}^n \to \mathbb{K}, \mathbf{x} \mapsto \mathbf{a}^T \mathbf{x} = \sum a_i x_i$ is linear.

Exercise 2.7

Let $C^1(\mathbb{R})$ be the set of continuously differentiable functions. Verify that $T : C^1 \to C^0, f \mapsto f'$ is a linear map.

Exercise 2.8

Prove that for any vector spaces V, W and any linear map $f : V \to W$, $f(0) = 0$.

Definition 2.11

A **matrix** *is a table of numbers. We denote the set of matrices of size m times n over the field $\mathbb{F}$ as $\mathbb{F}^{m\times n}$.*

Proposition 2.2

Let V and W be two finite dimensional vectors spaces with $\dim(U) = n$ and $\dim(V) = m$ and let $f : V \to W$ be a linear map. Let $S = (\mathbf{v}_1, \cdots, \mathbf{v}_n)$ be a basis for V. Then f is completely determined by the values of $f(\mathbf{v}_i)$.

Exercise 2.9

Let $f : U = \mathbb{R}_3[x] \to V = \mathbb{R}_3[x]$ be defined as the differentiation operator. Compute the matrices associated to f given the following basis

- $U = \operatorname{span}(1, x, x^2, x^3)$ and $V = \operatorname{span}(1, x, x^2, x^3)$.
- $U = \operatorname{span}(1, x, x^2, x^3)$ and $V = \operatorname{span}(1, 1 + x, 1 + x^2, 1 + x^3)$.
- $U = \operatorname{span}(1, 1 + x, 1 + x + x^2, 1 + x + x^2 + x^3)$ and $V = \operatorname{span}(1, 1 + x, 1 + x + x^2, 1 + x + x^2 + x^3)$.

Definition 2.12

Let V and W be two vector spaces and $\phi : V \to W$ a linear transformation. The **range** *or* **image** *of ϕ is the subspace*
$R(\phi) = Im(\phi) = \{\mathbf{w} \in W : \exists \mathbf{v} \in V \text{ with } \mathbf{w} = \phi(\mathbf{v})\} \subset W$.

Definition 2.13

Let V and W be two vector spaces and $\phi : V \to W$ a linear transformation. The **nullspace** *or* **kernel** *of ϕ is the subspace*
$N(\phi) = Ker(\phi) = \phi^{-1}(0) = \{\mathbf{v} \in V : \phi(\mathbf{v}) = 0\} \subset V$.

Exercise 2.10

Prove that the range and kernel of a linear mapping are indeed subspaces.

Exercise 2.11

Let $f : V \to W$, $S = (\mathbf{v}_1, \mathbf{v}_k)$ and $T = (f(\mathbf{v}_i))_i$. What can be said about T if

- S is a spanning set?
- S is linearly dependent?
- S is linearly independent?
- S is a basis?

Definition 2.14

The **rank** *of a linear application is the dimension of its range:*
$rk(f) = \dim(f(V))$.

Theorem 2.4 (Rank-nullity theorem)

Let V and W be two vector spaces with $\dim(V) = n < \infty$ and let $f : V \to W$ be a linear map. It holds

$$\dim(ker(f)) + rk(f) = \dim(V).$$

Definition 2.15

Let $A \in \mathbb{F}^{m\times m}$. Its **trace** *is defined as the sum of its diagonal entries:*

$$tr : \begin{array}{ccc} \mathbb{F}^{m\times m} & \rightarrow & \mathbb{F} \\ A & \mapsto & tr(A) = \sum_{i=1}^{m} a_{i,i} \end{array}$$

Exercise 2.12

Show that the trace is linear and prove the following identity:

$$tr(AB) = tr(BA), \text{ for any } A \in \mathbb{F}^{m\times n}, B \in \mathbb{F}^{n\times m}.$$

Definition 2.16

The **determinant** *of a matrix is defined in one of the following ways:*

1. *It is the only function* $f : \mathbb{F}^n \times \cdots \mathbb{F}^n \to \mathbb{F}$ *that is linear with respect to each column, alternating* $f(\cdots, \mathbf{u}, \cdots, \mathbf{v}, \cdots) = -f(\cdots, \mathbf{v}, \cdots, \mathbf{u}, \cdots)$ *and normalized such that* $f(I) = 1$*.*
2. $\det(A) = \sum_{\sigma \in P_n} \text{sign}(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$ *where* P_n *is the set of permutations of* $\{1, \cdots, n\}$ *and* $\text{sign}(\sigma) = (-1)^s$ *where* s *is the number of pairwise interchanges in* σ*.*
3. $\det(A) = \sum_{j=1}^n a_{i,j} \det(A_{i,j})$ *where* $A_{i,j}$ *is the matrix obtained from* A *by deleting the row* i *and column* j*.*

Exercise 2.13

Prove or compute the following results:

- $\det(AB) = \det(A)\det(B)$
- Computations for 2×2 matrices and Sarrus' rule for 3×3.
- $\det(A^T) =?$
- $A adj(A) = adj(A)A = \det(A)I$, where $adj(A)_{i,j} = (-1)^{i+j} A_{j,i}$ is the adjunct or adjugate matrix.

Definition 2.17

A matrix A is said to be **diagonal** *if $a_{i,j} = 0$ for $i \neq j$.*

Definition 2.18

A matrix A *is said to be* **upper triangular** *if* $a_{i,j} = 0$ *for* $i > j$*.*

Definition 2.19

A matrix A is said to be **lower triangular** *if $a_{i,j} = 0$ for $i < j$.*

Definition 2.20

A matrix A *is said to be* **symmetric** *if* $A^T = A$.

Definition 2.21

A matrix A is said to be **skew-symmetric** *if $A^T = -A$.*

Definition 2.22

A matrix A *is said to be* **Hermitian** *if* $A^* := \bar{A}^T = A$.

Definition 2.23

A matrix A is said to be **invertible** *if there exists a matrix B such that $AB = BA = I$. We write $B = A^{-1}$.*
If it is not invertible, it is said to be **singular**.

Exercise 2.14

Are all sets of these particular matrices subspaces of the vector space of matrices? In case of vector subspaces, what are their dimensions and give some basis.

Exercise 2.15

Which kind of structure does the set of symmetric matrices have?

Exercise 2.16

Prove that A is invertible if and only if $\det(A) \neq 0$ and give a formula for its inverse.

Exercise 2.17

Let T be an upper triangular matrix. Show that $\det(T) = \prod t_{ii}$.

Proposition 2.3

Given a square matrix A, the following statements are equivalent

1. A *is invertible.*
2. $ker(A) = \{0\}$.
3. $R(A) = \mathbb{K}^n$.

Definition 2.24

We say that a matrix A is **similar** *to a matrix B and write $A \sim B$ if there exists an invertible matrix P such that $A = PBP^{-1}$.*

Exercise 2.18

Let f be the differential operator on the set of degree 2 polynomials. Let $S = (1, x, x^2)$ and $T = (1, 1 + x, 1 + x + x^2)$. Furthermore, let A be the representation of f in the basis S and B the matrix representing f in T. Show that $A \sim B$. What does P represent?

Definition 2.25

$V = S \oplus T$ *is the* **direct sum** *of the subspaces* S *and* T *if*

1. $S \cap T = \{0\}$ *and*
2. $V = S + T$.

Exercise 2.19

Let S be the set of symmetric matrices and T the set of skew-symmetric matrices. Show that $\mathbb{K}^{n\times n} = S \oplus T$.

Definition 2.26

Given a square matrix $A \in \mathbb{K}^{n\times n}$. A pair of vector and scalar $(\mathbf{x}, \lambda) \in \mathbb{K} \times \mathbb{K}^n$ is called an **eigenpair** *if*

- $\mathbf{x} \neq 0$,
- $A\mathbf{x} = \lambda\mathbf{x}$.

$\mathbf{x}$ *is called an* **eigenvector** *with* **eigenvalue** λ.
The set of all eigenvectors corresponding to an eigenvalue λ is called the **eigenspace** *corresponding to λ*

Exercise 2.20

Verify that the eigenspaces are indeed vector spaces.

Proposition 2.4

$\lambda \in \mathbb{K}$ *is an eigenvalue for* $A \in \mathbb{K}^{n\times n}$ *if and only if*

$$\det(A - \lambda I) = 0.$$

Definition 2.27

For a given square matrix $A \in \mathbb{K}^{n \times n}$, its characteristic polynomial $p_A(x)$ is defined as

$$p_A(x) = \det(A - xI).$$

Hence, the zeros of the characteristic polynomial corresponds to the eigenvalues!

Definition 2.28

The set $\sigma(A) = \{x \in \mathbb{K} : p_A(x) = 0\}$ *is called the* **spectrum** *of* A.

Exercise 2.21

Show that

$$p_A(x) = (-1)^n x^n + (-1)^{n-1} tr(A) x^{n-1} + \cdots + \det(A)$$

and show that

$$tr(A) = \sum_{i=1}^{n} \lambda_i \quad \det(A) = \prod_{i=1}^{n} \lambda_i$$

where the λ_i are the n (possibly complex and repeated eigenvalues of A).
Conclude that A is invertible $\Leftrightarrow 0 \notin \sigma(A)$.

Exercise 2.22

Let A and B be two square matrices such that $A \sim B$. It holds

$$tr(A) = tr(B)$$
$$\det(A) = \det(B).$$

Theorem 2.5 (Invertible matrix theorem)

Let $A \in \mathbb{K}^{n \times n}$. The following statements are equivalent

1. A *is non-singular*
2. A^{-1} *exists*
3. $rk(A) = n$
4. *the columns of* A *are linearly independent*
5. *the rows of* A *are linearly independent*
6. $\det(A) \neq 0$
7. *the dimension of the range of* A *is* n
8. *the nullity of* A *is* 0
9. $A\mathbf{x} = \mathbf{y}$ *is consistent (= admits at least one solution) for each* $\mathbf{y} \in \mathbb{K}^n$
10. *if* $A\mathbf{x} = \mathbf{y}$ *is consistent then the solution is unique*
11. $A\mathbf{x} = \mathbf{y}$ *has a unique solution for each* $\mathbf{y} \in \mathbb{K}^n$
12. *the only solution to* $A\mathbf{x} = 0$ *is* $\mathbf{x} = 0$
13. 0 *is not an eigenvalue of* A

Proposition 2.5

Let $\mathbf{u}$ *and* $\mathbf{v}$ *be two eigenvectors associated to the two different eigenvalues* $\lambda \neq 0$ *and* $\mu \neq 0$ *respectively. Then* $\mathbf{u}$ *and* $\mathbf{v}$ *are linearly independent.*

Exercise 2.23

Show the following: there exists a non-singular matrix V and a diagonal matrix D such that $A = VDV^{-1}$ if and only if there exists n linearly independent eigenvectors $\mathbf{v}_i$ with respective eigenvalues λ_i.

Definition 2.29

We say that a matrix A *is* **diagonalizable** *if there exists a non-singular matrix* P *and a diagonal matrix* D *such that*

$$A = PDP^{-1}.$$

Definition 2.30

Let $p_A(x) = (-1)^n (x - \lambda_1)^{p_1} \cdots (x - \lambda_r)^{p_r}$ *with* $\sum p_i = n$, *be written in its (complex) factorized form. Then*

- p_i *is the* **algebraic multiplicity** *of the eigenvalue* λ_i
- $\dim(ker(A - \lambda_i I)) = n - rk(A - \lambda_i I) =: q_i$ *is the* **geometric multiplicity** *of the eigenvalue* λ_i.

Exercise 2.24

Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Find the eigenvalues, their algebraic and geometric multiplicities of A and B.

Proposition 2.6

Let $A \in \mathbb{K}^{n\times n}$ *and let* $\lambda_1, \cdots, \lambda_r$ *be* r *distinct eigenvalues with respective geometric multiplicities* $q_1, \cdots, q_r$. *Let furthermore* $\mathbf{v}_i^j$ *be the* j^{th} *eigenvector with eigenvalue* λ_i, $1 \leq i \leq r$, $1 \leq j \leq q_i$. *Then the family* $\{\mathbf{v}_i^j\}_{i,j}$ *is a linearly independent family of vectors.*

Theorem 2.6

Let $A \in \mathbb{K}^{n \times n}$. A is diagonalizable if and only if $q_i = p_i$ for all r distinct eigenvalues.

Corollary 2.1

If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable.

Example 2.5

The process of diagonalizing a matrix is always the same:

1. Compute the characteristic polynomial
2. Find the eigenvalues and their respective algebraic multiplicities
3. For each eigenvalue, find a basis of the eigenspaces
4. Side: if you find less eigenvectors than the total dimension, the matrix is not diagonalizable
5. Define the matrix $V = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ containing all the eigenvectors
6. Define the matrix $D = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$
7. You obtain the diagonalization $A = VDV^{-1}$.

Apply this to

$$A = \begin{bmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{bmatrix}.$$