MATRIX ANALYSIS

A Quick Guide

Huan Bui

Colby College
Physics & Statistics
Class of 2021

February 10, 2019

Preface

Greetings,

Matrix Analysis: A Quick Guide to is compiled based on my MA353: Matrix Analysis notes with professor Leo Livshits. The sections are based on a number of resources: *Linear Algebra Done Right* by Axler, *A Second Course in Linear Algebra* by Horn and Garcia, *Matrices and Linear Transformations* by Cullen, *Matrices: Methods and Applications* by Barnett, *Problems and Theorems in Linear Algebra* by Prasolov, *Matrix Operations* by Richard Bronson, and professor Leo Livshits' own textbook (in the making).

The development of this text will come in layers. The first layer, one that I am working on during the course of S'19 MA353, will be an overview of the key topics listed in the table of contents. As the semester progresses, I will be constantly updating the existing notes, as well as adding prof. Livshits' problems and my solutions to the problems. The second layer will come after the course is over, when concepts will have hopefully "come together."

I will decide how much narrative I should put into the text as text is developed over the semester. I'm thinking that I will only add detailed explanations wherever I find fit or necessary for my own studies. I will most likely keep the text as condensed as I can.

Technical note: for the conjugate transpose operation, I will use "†" and "*" interchangeably. I'm more biased towards † as a physics student, but to appeal to my mathematics professor, I should also learn the mathematicians' notations.

Enjoy!

Contents

1 List of Special Matrices & Their Properties

1. **Hermitian/Self-adjoint**: $H = H^\dagger$. A Hermitian matrix is matrix that is equal to its own conjugate transpose:

$$H \text{ is Hermitian } \iff H_{ij} = \bar{H}_{ji}$$

Properties 1.1.

(a) H is Hermitian $\iff \langle w, Hv \rangle = \langle Hw, v \rangle$, where $\langle , \rangle$ denotes the inner product.

(b) H is Hermitian $\iff \langle v, Hv \rangle \in \mathbb{R}$.

(c) H is Hermitian $\iff$ it is *unitarily diagonalizable* with *real eigenvalues.*

Unitary: $U^*U = UU^* = I = U^\dagger U = UU^\dagger$. The real analogue of a unitary matrix is an orthogonal matrix. The following list contain the properties of U:

(a) U preserves the inner product:

$$\langle Ux, Uy \rangle = \langle x, y \rangle.$$

(b) U is normal: it commutes with $U^* = U^\dagger$.

(c) U is diagonalizable:

$$U = VDV^*,$$

where D is diagonal and unitary, and V is unitary.

(d) $|\det(U)| = 1$ (hence the real analogue to U is an orthogonal matrix)

(e) Its eigenspaces are orthogonal.

(f) U can be written as

$$U = e^{iH},$$

where H is a Hermitian matrix.

(g) Any square matrix with unit Euclidean norm is the average of two unitary matrices.

2. **Cofactor**: ?

2 List of Operations

1. **Conjugate transpose** is what its name suggests.

2. **Classical adjoint/Adjugate/adjunct** of a square matrix is the transpose of its cofactor matrix.

3 List of Algorithms

4 Complex Numbers

4.1 A different point of view

We often think of complex numbers as

$$a + ib$$

where $a, b \in \mathbb{R}$ and $i = \sqrt{-1}$. While there is nothing "bad" about this way of thinking - in fact thinking of complex numbers as $a+ib$ allows us to very quickly and intuitively do arithmetics operations on them - a "matrix representation" of complex numbers can give us some insights on "what we actually do" when we perform complex arithmetics.

Let us think of

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

as a different representation of the same object - the same complex number "$a + ib$." Note that it does not make sense to say the matrix representation **equals** the complex number itself. But we shall see that a lot of the properties of complex numbers are carried into this matrix representation under interesting matricial properties.

$$\boxed{\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \sim a + ib}$$

First, let us break the the matrix down:

$$a + ib = a \times 1 + i \times b \sim \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = aI + b\mathcal{I}.$$

Right away, we can make some "mental connections" between the representations:

$$I \sim 1$$
$$\mathcal{I} \sim i.$$

Now, we know that complex number multiplications commute:

$$(a + ib)(c + id) = (c + id)(a + ib).$$

Matrix multiplications are not commutative. So, we might wonder whether commutativity holds under the this new representation of complex numbers. Well, the answer is yes. We can readily verify that

$$(aI + b\mathcal{I})(cI + b\mathcal{I}) = (cI + b\mathcal{I})(aI + b\mathcal{I}).$$

How about additions? Let's check:

$$(a+ib)+(c+id)=(a+c)+i(b+d)\sim\begin{pmatrix}a+c & -(b+d)\\(b+d) & a+c\end{pmatrix}=\begin{pmatrix}a & -b\\b & a\end{pmatrix}+\begin{pmatrix}c & -d\\d & c\end{pmatrix}.$$

Ah! Additions work. So, the new representation of complex numbers seems to be working flawlessly. However, we have yet to gain any interesting insights into the connections between the representations. To do that, we have to look into changing the form of the matrix. First, let's see what conjugation does:

$$(a+ib)^*=a-ib\sim\begin{pmatrix}a & b\\-b & a\end{pmatrix}=\begin{pmatrix}a & -b\\b & a\end{pmatrix}^\top$$

Ah, so conjugation to a complex number in the traditional representation is the same as transposition in the matrix representations. What about the amplitude square? Let us call

$$M=\begin{pmatrix}a & -b\\b & a\end{pmatrix}.$$

We have

$$(a+ib)(a-ib)\sim\begin{pmatrix}a & -b\\b & a\end{pmatrix}\begin{pmatrix}a & b\\-b & a\end{pmatrix}=MM^\top=(a^2+b^2)I=\det(M)I$$

Interesting. But observe that if $\det(M)\neq 0$

$$\frac{1}{\det(M)}MM^\top=I.$$

This tells us that

$$M^\top=M^{-1},$$

where M^{-1} is the inverse of M, and, not surprisingly, it corresponds to the reciprocal to the complex number $a+ib$. We can readily show that

$$M^{-1}\sim(a+ib)^{-1}=\frac{1}{a^2+b^2}(a-ib).$$

Remember that we can also think of a complex number as a column vector:

$$c+id\sim\begin{pmatrix}c\\d\end{pmatrix}.$$

Let us look back at complex number multiplication under matrix representation:

$$(a+ib)(c+id)=(ac-bd)+i(bc+ad)\sim\begin{pmatrix}a & -b\\b & a\end{pmatrix}\begin{pmatrix}c\\d\end{pmatrix}=\begin{pmatrix}ac-bd\\bc+ad\end{pmatrix}.$$

Multiplication actually works in this "mixed" way of representing complex numbers as well. Now, observe that what we just did was performing a linear transformation on a vector in $\mathbb{R}^2$. It is always interesting to look at the geometrical interpretation of this transformation. To do this, let us call N the "normalized" version of M:

$$N = \frac{1}{\sqrt{a^2+b^2}} \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

We immediately recognize that N is an orthogonal matrix. This means N is an orthogonal transformation (length preserving). Now, it is reasonable to define

$$\cos\theta = \frac{a}{\sqrt{a^2+b^2}}$$
$$\sin\theta = \frac{b}{\sqrt{a^2+b^2}}.$$

We can write N as

$$N = \begin{pmatrix} \cos\theta & -\sin\theta \sin\theta & \cos\theta \end{pmatrix},$$

which is a physicists' favorite matrix: the rotation by θ. So, let us write M in terms of N:

$$M = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \sqrt{a^2+b^2} N = \sqrt{a^2+b^2} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

We can interpret M as a rotation by θ, followed by a scaling by $\sqrt{(a^2+b^2)}$. But what $\sqrt{a^2+b^2}$ exactly is just the "length" or the "amplitude" of the complex number $a+ib$, if we think of it as an arrow in a plane.

4.2 Relevant properties and definitions

1. The *modulus* of $z = a+ib$ is the "amplitude" of z, denoted by $|z| = \sqrt{a^2+b^2} = z\bar{z}$.
2. The modulus is *multiplicative*, i.e.

$$|wz| = |w||z|.$$

3. Triangle inequality:

$$|z+w| \leq |z| + |w|.$$

   We can readily show this geometrically, or algebraically.

4. The *argument* of $z = a+ib$ is θ, where

$$\theta = \begin{cases} \tan^{-1}\left(\frac{b}{a}\right), \text{ if } a > 0 \\ \frac{\pi}{2} + k2\pi, k \in \mathbb{R} \text{ if } a = 0, b > 0 \\ -\frac{\pi}{2} + k2\pi, k \in \mathbb{R} \text{ if } a = 0, b < 0 \\ \text{Undefined if } a = b = 0. \end{cases}$$

5. The *conjugate* of $a+ib$ is $a-ib$. Conjugation is *additive* and *multiplicative*, i.e.

$$z \bar{+} w = \bar{z} + \bar{w}$$
$$\bar{w}z = \bar{w}\bar{z}.$$

Note that we can also show the multiplicative property with the matrix representation as well:

$$\bar{w}z \sim (WZ)^\top = Z^\top W^\top \sim \bar{z}\bar{w} = \bar{w}\bar{z}.$$

6. Euler's identity, generalized to de Moivre's formula:

$$z^n = r^n e^{in\theta}.$$

5 Vector Spaces & Linear Functions

5.1 Rank-Nullity Theorem

5.2 Isomorphisms

5.3 Coordinatization & matricial representation of linear functions

6 Products of vector spaces & Sums of subspaces

6.1 Direct Sums

Definition 6.1. Let U_j, $j = 1, 2, \dots m$ are subspaces of V. $\sum_1^m U_j$ is a *direct sum* if each $u \in \sum U_j$ can be written in only one way as $u = \sum_1^m u_j$. The direct sum $\sum_i^m U_j$ is denoted as $U_1 \oplus \cdots \oplus U_m$.

Properties 6.1.

1. Condition for direct sum: If all U_j are subspaces of V, then $\sum_1^m U_j$ is a direct sum $\iff$ the only way to write 0 as $\sum_1^m u_j$, where $u_j \in U_j$ is to take $u_j = 0$ for all j.

2. If U, W are subspaces of V and $U \bigcap W = \{0\}$ then $U + W$ is a direct sum.

6.2 Products and Quotients of Vector Spaces

6.3 Products of Vector Spaces

Definition 6.2. Product of vectors spaces

$$V_1 \times \cdots \times V_m = \{(v_1, \dots, v_m) : v_j \in V_j, j = 1, 2, \dots, m\}.$$

Definition 6.3. Addition on $V_1 \times \cdots \times V_m$:

$$(u_1, \dots, u_m) + (v_1, \dots, v_m) = (u_1 + v_1, \dots, v_m + u_m).$$

Definition 6.4. Scalar multiplication on $V_1 \times \cdots \times V_m$:

$$\lambda(v_1, \dots, v_m) = (\lambda v_1, \dots, \lambda v_m).$$

Properties 6.2.

1. Product of vectors spaces is a vector space.

   V_j are vectors spaces over $\mathbb{F} \implies V_1 \times \cdots \times V_m$ is a vector space over $\mathbb{F}$.

2. Dimension of a product is the sum of dimensions:

$$\dim(V_1 \times \cdots \times V_m) = \sum_1^m \dim(V_j)$$

6.4 Products & Direct Sums

Properties 6.3.

1. Let $U_1, \dots, U_m$ be subspaces of V. Define a linear map $\Gamma : U_1 \times \cdots \times U_m \to U_1 + \cdots + U_m$ by:

$$\Gamma(u_1, \dots, u_m) = \sum_1^m u_j.$$

   $U_1 + \cdots + U_m$ is a direct sum $\iff \Gamma$ is injective.

2. Let U_j be finite-dimensional and are subspaces of V.

$$U_1 \oplus \cdots \oplus U_m \iff \dim(U_1 + \cdots + U_m) = \sum_1^m \dim(U_j)$$

6.5 Quotients of Vector Spaces

6.6 Nullspaces

6.7 Ranges of Operator Powers

7 Idempotents & Resolutions of Identity

8 Block-representations of operators

8.1 Direct sums of operators

9 Invariant subspaces

9.1 Reducing subspaces

10 Polynomials applied to operators

10.1 Minimal polynomials of block-Δ^r operators

10.2 Minimal polynomials at a vector

11 Eigentheory

11.1 Spectral Mapping Theorem

12 Triangularization

12.1 Compression to invariant subspaces

12.2 Simultaneously Δ-ity of commuting families

13 Diagonalization

13.1 Spectral resolutions

13.2 Compressions to reducing subspaces

13.3 Simultaneous diagonalizability for commuting families

14 Primary decomposition over $\mathbb{C}$ and generalized eigenspaces

15 Cyclic decomposition and Jordan form

15.1 Square roots of operators

15.2 Similarity of a matrix and its transpose

15.3 Similarity of a matrix and its conjugate

15.4 Jordan forms of $\mathcal{AB}$ and $\mathcal{BA}$

15.5 Power-convergent operators

15.6 Power-bounded operators

15.7 Row-stochastic matrices

16 Determinant & Trace

16.1 Classical adjoints

16.2 Cayley-Hamilton theorem

17 Inner products and norms

17.1 Riesz representation theorem

17.2 Adjoints

17.3 Grammians

17.4 Orthogonal complements and orthogonal decompositions

17.5 Ortho-projections

17.6 Closest point solutions

17.7 Gram-Schmidt and orthonormal bases

18 Isometries and unitary operators

19 Ortho-triangularization

20 Spectral resolutions

21 Ortho-diagonalization; self-adjoint and normal operators; spectral theorems

22 Positive (semi-)definite operators

22.1 Classification of inner products

22.2 Positive square roots

23 Polar decomposition

24 Single value decomposition

24.1 Spectral/operator norm

24.2 Singular values and approximation

24.3 Singular values and eigenvalues