# Matrix Analysis Notes

Jasper Ty

September 29, 2024

## What is this?

These are typed up notes for my Matrix Analysis class (MATH-504 at Drexel University). The class is based on the textbook "Matrix Analysis" by Horn and Johnson. I am using the second edition of the book [HoJo13]

The sections match up with the book, but theorem and definition numbers do not.

## Contents



# 0 Review

Omitted.

# 1 Eigenvectors, eigenvectors, and similarity

## 1.0 Introduction

**Proposition 1.0.1.** If $T$ is a linear transformation on a vector space $V$, then if we fix a particular basis $\mathscr{B}$ of $V$ and put $A = [T]_{\mathscr{B}}$, then the set of *all possible basis representations of* $T$ is the set of all *matrices similar to* $A$.

**Definition 1.0.2.** Let $T$ be a linear transformation on a vector space $V$ over $\mathbb{F}$. An *eigenvector* of $T$ is a vector $\mathbf{v} \in V$ such that

$$T\mathbf{v} = \lambda\mathbf{v}$$

for some $\lambda \in \mathbb{F}$.

An *eigenvalue* of $T$ is a scalar $\lambda \in \mathbb{F}$ such that *there exists* an eigenvector for which the above equation holds.

Eigenvectors arise, for example, from the problem of *maximizing a real symmetric quadratic form on the boundary of the unit disk*, namely the task of finding

$$\max\{\mathbf{x}^\top A\mathbf{x} : \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{x}^\top\mathbf{x} = 1\}.$$

where $A \in M_n(\mathbb{R})$ is a real symmetric matrix.

This can be solved with Lagrange multipliers, and the appropriate Lagrangian function is

$$\mathcal{L}(\mathbf{x}, \lambda) = \mathbf{x}^\top A\mathbf{x} - \lambda\mathbf{x}^\top\mathbf{x}.$$

To compute its stationary points: we have the following fact: $\nabla(\mathbf{x}^\top A\mathbf{x}) = 2A\mathbf{x}$, which tells us that

$$\nabla\mathcal{L} = 2(A\mathbf{x} - \lambda\mathbf{x}).$$

Hence it must be that $A\mathbf{x} = \lambda\mathbf{x}$ is a necessary condition to find a maximizing or minimizing vector for our problem.

By Weierstrass's theorem, which is an avatar of the general topological theorem that *continuous images of compact spaces are compact*, we *know* that $\mathbf{x}^\top A\mathbf{x}$ attains its minimum and minimum for some $\mathbf{x}_{\text{min}}, \mathbf{x}_{\text{max}}$.

**Theorem 1.0.3** (Weierstrass)**.** Continuous images of compact Fix a vector space $V$ with a norm $\|\cdot\|$. If $S \subseteq V$ is compact, and $f : S \to \mathbb{R}$ is continuous, then $\sup f(S) \in f(S)$ and $\inf f(S) \in f(S)$ respectively.

Then, since $A\mathbf{x} = \lambda\mathbf{x}$ was a *necessary condition* for such extrema to exist, it must follow that $A\mathbf{x}_{\text{max}} = \lambda_1\mathbf{x}_{\text{max}}$ and $A\mathbf{x}_{\text{min}} = \lambda_2\mathbf{x}_{\text{min}}$ for some $\lambda_1, \lambda_2 \in \mathbb{R}$.

What we have just proved is the following:

**Theorem 1.0.4.** Every real symmetric matrix has at least one real eigenvalue.

## 1.1 The eigenvalue-eigenvector equation.

We review basic terminology about eigenvalues and eigenvectors.

**Definition 1.1.1.** An *eigenpair* for a matrix $A \in M_n$ is a pair $(\mathbf{x}, \lambda)$ where $\mathbf{x} \in \mathbb{C}^n \setminus \{0\}$ and $\lambda \in \mathbb{C}$ such that

$$A\mathbf{x} = \lambda\mathbf{x}.$$

We call $\mathbf{x}$ an *eigenvector* of $A$, and we call $\lambda$ an *eigenvalue* of $A$.

This definition can readily be manipulated to show the following:

**Proposition 1.1.2.** Eigenvectors are the nontrivial solutions to the matrix-vector equation

$$(\lambda I - A)\mathbf{x} = 0.$$

**Definition 1.1.3.** Let $A \in M_n$. The *spectrum,* $\sigma(A)$ of $A$ is the set of all eigenvalues of $A$.

We note that *evaluating polynomials at matrices* is entirely well defined, as summing and taking powers of matrices is well defined.

We have the following theorem

**Theorem 1.1.4.** Let $p$ be a polynomial. If $\lambda, x$ is an eigenvalue-eigenvector pair of $A \in M_n$, then $p(\lambda), x$ is an eigenvalue-eigenvector pair of $p(A)$.

Conversely, if $k \geq 1$ and $\mu$ is an eigenvalue of $p(A)$, then there is some eigenvalue $\lambda$ of $A$ such that $\mu = p(\lambda)$.

## 1.2 The characteristic polynomial and algebraic multiplicity

**Definition 1.2.1.**

# 2 Unitary similarity and unitary equivalence

## 2.1 Unitary matrices and the $QR$ factorization

**Definition 2.1.1.** A list of vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k \in \mathbb{C}^n$ is said to be *orthogonal* if $\mathbf{x}_i^* \mathbf{x}_j = 0$ for all $1 \le i < j \le k$. If, in particular, $\mathbf{x}_i^* \mathbf{x}_i$ for all $1 \le i \le k$, then we say that the list is *orthonormal*.

**Theorem 2.1.2.** Every orthonormal list of vectors in $\mathbb{C}^n$ is linearly independent.

*Proof.* Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ be an orthonormal set, and suppose $\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k = 0$. Then

$$\begin{aligned} 0 &= 0^* 0 \\ &= (\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k)^* (\alpha_1 \mathbf{x}_1 + \cdots + \alpha_k \mathbf{x}_k). \\ &= \sum_{i,j} \overline{\alpha_i} \alpha_j \mathbf{x}_i^* \mathbf{x}_j \\ &= \sum_i |\alpha_i|^2. \end{aligned}$$

Hence each $\alpha_i$ must be zero. □

**Definition 2.1.3.** A matrix $U \in M_n$ is unitary if $U^* U = I$. In particular, a real matrix $U \in M_n(\mathbb{R})$ is *orthogonal* if $U^T U = I$.

**Theorem 2.1.4.** Let $U \in M_n$. Then the following are equivalent

(a) $U$ is unitary.

(b) $U$ is nonsingular and $U^* = U^{-1}$.

(c) $UU^* = I$.

(d) $U^*$ is unitary.

(e) The columns of $U$ are orthonormal.

(f) The rows of $U$ are orthonormal.

(g) For all $\mathbf{x} \in \mathbb{C}^n$, $\|\mathbf{x}\|_2 = \|U\mathbf{x}\|_2$.

*Proof.* (a) implies (b) follows from the defining equation

$$U^*U = I.$$

(b) implies (c) since matrix inverses are both left and right inverses.

(d) is verified directly from (c)

$$(U^*)^*(U^*) = UU^* = I.$$

TODO: Finish proof of characterizations of unitarity □

**Definition 2.1.5.** A linear transformation $T : \mathbb{C}^n \to \mathbb{C}^n$ is called a *Euclidean isometry* if $\|\mathbf{x}\|_2 = \|T\mathbf{x}\|_2$ for all $x \in \mathbb{C}^n$.

The previous theorem tells us that all such transformations are in fact represented by *square* unitary matricess.

**Remark 2.1.6.** If $U, V \in M_n$ are unitary, then so is $UV$.

*Proof.*

$$(UV)^*(UV) = V^*U^*UV = V^*V = I.$$

□

**Definition 2.1.7.** The subset of $\mathrm{GL}(n, \mathbb{C})$ consisting of all $n \times n$ unitary matrices is a subgroup called the unitary group $\mathrm{U}(n)$.

**Theorem 2.1.8** (QR factorization)**.** Let $A \in M_{n,m}$. If $n \geq m$, then there *always exist* matrices $Q_{n,m}$ and $R \in M_m$ such that

$$A = QR.$$

where $Q$ is an isometry, and $R$ is diagonal.

*Proof.* The proof in the book uses the machinery of Householder matrices. The professor gave us a proof via keeping track of intermediate results in applying the Gram-Schmidt algorithm. □

# References

[HoJ013] Roger A. Horn and Charles R. Johnson, *Matrix Analysis, Second Edition*, Cambridge University Press 2013.