

MATH 20700 (Honors Analysis in \mathbb{R}^n I) Notes

Steven Labalme

October 22, 2021

Contents

I	Linear Algebra	1
1	Basic Notions	2
2	Systems of Linear Equations	4
3	Determinants	6
4	Introduction to Spectral Theory	8
5	Inner Product Spaces	10
6	Structure of Operators on Inner Product Spaces	14
7	Bilinear and Quadratic Forms	19
8	Dual Spaces and Tensors	22
9	Advanced Spectral Theory	23
	References	24

List of Figures

3.1	Visualizing properties of determinants.	6
5.1	The unit ball of norms corresponding to $p = 1, 2, \infty$	11

Part I

Linear Algebra

Chapter 1

Basic Notions

- 9/27:
- **Vector space:** Basically, a set for which you have an addition and multiplication.
 - \mathbb{F}^d is used for \mathbb{R}^d or \mathbb{C}^d in Treil (2017).
 - \mathbb{P}_n is the vector space of polynomials up to degree n .
 - $C([0, 1])$ is the set of continuous functions defined on $[0, 1]$, an infinite-dimensional vector space.
 - **Generating set:** A subset of a vector space, all linear combinations of which generate the vector space. *Also known as spanning set.*
 - Any element of VS is a linear comb. of elements of the generating set.
 - **Linearly independent (list):** A list of vectors $\mathbf{v}_1, \dots, \mathbf{v}_k \in V$ such that $\sum_{i=1}^k \alpha_i \mathbf{v}_i = 0$ implies $\alpha_i = 0$ for all i .
 - **Base:** A generating set consisting of linearly independent vectors.
 - Any element of a VS can be written as a *unique* linear combination of the vectors in a base.
 - If $\mathbf{x} = \sum_{i=1}^k \alpha_i \mathbf{v}_i = \sum_{i=1}^k \beta_i \mathbf{v}_i$, then $\alpha_i = \beta_i$ for all i .
 - **Linear transformation:** A function $T : X \rightarrow Y$, where X, Y are VSs, such that

$$T(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha T\mathbf{x} + \beta T\mathbf{y}$$

for all $\mathbf{x} \in X, \mathbf{y} \in Y$.

- Examples of linear transformations:
 - Consider \mathbb{P}_n . Let $Tp_n = p'_n$. This T is linear.
 - Rotation in \mathbb{R}^d .
 - Think graphically about two vectors \mathbf{x}, \mathbf{y} .
 - Rotating and summing them is the same as summing and rotating. Same for scaling.
 - Thus, rotation is actually linear!
 - Reflection as well.
- Consider $T : \mathbb{R} \rightarrow \mathbb{R}$.
 - Any linear map on the line is a line.
 - We must have $Tx = \alpha x$: $Tx = T(1x) = xT(1) = x\alpha$.
- Consider $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear.

- Any linear map between \mathbb{R}^n and \mathbb{R}^m is linear.
- Thus, $T(\mathbf{x}) = A\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^n$, where A is an $m \times n$ matrix.
- To find A , do the same calculation as for $T\mathbf{x} = \alpha\mathbf{x}$ but more carefully:
 - Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be a basis.
 - So $\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{e}_i$.
 - Thus, $T\mathbf{x} = \sum_{i=1}^n \alpha_i T(\mathbf{e}_i)$.
 - Each $T(\mathbf{e}_i)$ is part of the matrix that we multiply by the column vector representing \mathbf{x} .
- Multiplication of matrices is equivalent to composition of linear maps.
- Consider $T_1 : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $T_2 : \mathbb{R}^m \rightarrow \mathbb{R}^r$.
 - $T_2 \circ T_1$ is equivalent to BA , if A represents T_1 and B represents T_2 . In other words, $(T_2 \circ T_1)(\mathbf{x}) = BA\mathbf{x}$ for all \mathbf{x} .
- Recall that if $A = (\alpha_{ij})$ and $B = (\beta_{ij})$, then $(BA)_{ij} = (\sum \beta_{ik} \alpha_{kj})$.
- Properties of multiplication:

$$(AB)C = A(BC)$$

$$A(B + C) = AB + AC$$

$$(A + B)C = AC + BC$$

- However, it is not true in general that $AB = BA$.
- **Trace** (of an $n \times n$ matrix A): The sum of the diagonal entries of A . Denoted by $\mathbf{tr}(A)$. Given by

$$\mathbf{tr}(A) = \sum \alpha_{ii}$$

- It is true that $\mathbf{tr}(AB) = \mathbf{tr}(BA)$.
 - Indeed, on the diagonals, multiplication is commutative; it's the other terms that mess you up in general.
- Invertibility of matrices.
 - In general, matrices are not invertible: Not every system of equations is solveable; $Ax = b$ does not always have a solution $x = A^{-1}b$.
- C is the inverse from the left: $CA = I$. B is the inverse from the right: $AB = I$. A matrix can have a left and a right inverse and still not be invertible. A matrix is invertible iff $C = B$.
- Any time we write “inverse,” we do so under the assumption that it exists.
- $(AB)^{-1} = B^{-1}A^{-1}$ — easy proof by multiplication.
- If $A = (a_{ij})$, $A^T = (a_{ji})$.
 - $(A^{-1})^T = (A^T)^{-1}$.
 - $(AB)^T = B^T A^T$.
- Let X, Y VS.
 - $X \cong Y^{[1]}$ if there exists a linear $T : X \rightarrow Y$ that is one-to-one and onto.
 - Check: $A(\text{basis of } X) = \text{basis of } Y$. Prove by definition and expression of elements as linear combinations.
- **Subspace**: A subset of a vector space which happens to be a vector space, itself.

¹“ X is isomorphic to Y .”

Chapter 2

Systems of Linear Equations

9/29:

- Row elimination:

- Let

$$A = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 3 & 1 & 2 & 7 \\ 2 & 1 & 2 & 1 \end{pmatrix}$$

- Then the **echelon form** matrix

$$A_e = \begin{pmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 2 & -4 \end{pmatrix}$$

- Lastly, the **reduced echelon form** matrix

$$A_{re} = \begin{pmatrix} 1 & 0 & 0 & 7 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{pmatrix}$$

- **Echelon form:**

- All zero rows are below nonzero rows.
 - For any nonzero row, its leading element is strictly to the left of the nonzero entry of the next row.

- **Reduced echelon form:**

- All pivots are 1.
 - Used to solve systems of the form $Ax = b$.

- **Inconsistent** (system of equations): A system with no solution.

- If the last row is of the form $(0, \dots, 0, b)$ where $b \neq 0$, then there is no solution.

- Unique solution if A_e has a pivot in every column.

- There exists a solution for every b if there is a pivot in every row?

- Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a matrix. Then $\ker A = \{x \in \mathbb{R}^n : Ax = 0\}$ (subspace of \mathbb{R}^n) and $\text{range } A = \{Ax : x \in \mathbb{R}^n\}$ (subspace of \mathbb{R}^m).

- Also consider $\ker(A^T)$ and $\text{range}(A^T)$, the basis of the kernel and range, and dimension.

- Finite-dimensional vector spaces:

- A basis is a generating set (so every element of V can be written uniquely as a linear combination of the basis) the length of which is equal to the dimension of V .
- All bases of finite-dimensional vector spaces have the same number of elements.
 - Let v_1, v_2, v_3 and w_1, w_2 be two generating sets of V .
 - Then

$$v_1 = \lambda_{11}w_1 + \lambda_{12}w_2$$

$$v_2 = \lambda_{21}w_1 + \lambda_{22}w_2$$

$$v_3 = \lambda_{31}w_1 + \lambda_{32}w_2$$

- Suppose the only solution to $\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0$ is $\alpha_1 = \alpha_2 = \alpha_3 = 0$.
 - But this is not true, as we can find another one in terms of the λ s.
- If you have a list of linearly independent vectors, you can complete it into a basis.
 - If there exists a vector that can't be written as a linear combination of the list, add it to the list.
- If you find any particular solution to a system $Ax = b$, and you add to it any element of $\ker A$, you will obtain another solution.
 - $Ax_1 = b$ and $Ax_h = 0$ implies that $A(x_1 + x_h) = b$.
 - $Ax_1 = b$ and $Ax_2 = b$ imply that $A(x_1 - x_2) = 0$, i.e., that $x_1 - x_2 \in \ker A$.
- If $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\dim \text{range } A = m$, then $Ax = b$ is solvable for all $b \in \mathbb{R}^m$.
- Let $\text{rank } A = \dim \text{range } A$.
- Rank theorem:
 - $\text{rank } A = \text{rank } A^T$.
 - Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$. We know that $\dim \ker A + \dim \text{range } A = n$.
 - $\dim \ker A^T + \text{rank } A^T = m$.
 - This theorem survives linear algebra and enters functional analysis under the name **Fredholm's alternative**.

- **Fredholm's alternative:** $Ax = b$ has a solution for all $b \in \mathbb{R}^n$ iff $\dim \ker A^T = 0$.
 - $\dim \ker A^T = 0$ implies $\text{rank } A^T = m$ implies $\text{rank } A = m$ implies $\dim \text{range } A = m$, as desired.
- **Pivot column** (of A): A column of A where A_e has pivots.
- The **pivot columns** of A give a basis for $\text{range } A$.
- The pivot rows of A_e give a basis for $\text{range } A^T$.
- A basis for the kernel is enough to solve $Ax = 0$.
- If you take these three things as givens, you can prove the rank theorem.

Chapter 3

Determinants

- 9/29:
- The determinant, geometrically, is the volume of the object (in \mathbb{R}^3) you get when you take linear combinations of the vectors.
 - In 2D:
 - Let v_1, v_2 be two vectors. Put tail to tail and forming a parallelogram, the determinant of the matrix (v_1, v_2) is the area of said parallelogram.
 - Linearity 1: $D(av_1, v_2, \dots, v_n) = aD(v_1, \dots, v_n)$ is the same as saying that if you stretch one vector by a , you scale up the area by that much, too.
 - Linearity 2: $D(v_1, \dots, v_{k+} + v_{k-}, \dots, v_n) = D(-) + D(+)$.
 - Antisymmetry: $D(v_1, \dots, v_k, \dots, v_j, \dots, v_n) = -D(v_1, \dots, v_j, \dots, v_k, \dots, v_n)$. Interchanging columns flips the sign of the determinant.
 - Basis: $D(e_1, \dots, e_n) = 1$.
 - Determinant: Denoted by $D(v_1, \dots, v_n)$, where (v_1, \dots, v_n) is an $n \times n$ matrix.
- 10/1:
- Consider an $n \times n$ matrix A consisting of n columns containing vectors $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$.
 - $D(A)$ is the volume of the solid $V = \sum_{i=1}^n \alpha_i v_i$.
 - $D(\mathbf{e}_1, \dots, \mathbf{e}_n) = 1$.

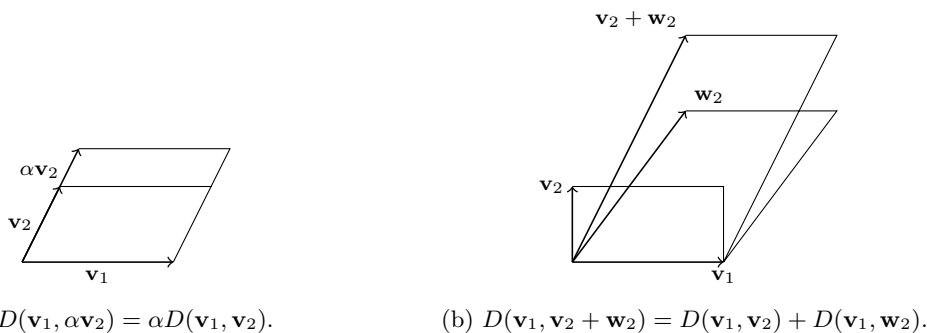


Figure 3.1: Visualizing properties of determinants.

- Basic properties of the determinant.
 - If A has a zero column, then $\det A = 0$: Scalar property.
 - If A has two equal columns, then $\det A = 0$: Multiply one by minus and add.

- If A has a column which is a multiple of another, then $\det A = 0$: Pull out the multiple and then you have the previous one.
- If columns are linearly dependent, then $\det A = 0$: Decompose it into sums, split, add back up with previous properties.
- The determinant is preserved under column reduction.
- $\det A^T = \det A$: Put everything in rref.
- If A is not invertible, then $\det A = 0$ (not invertible implies linearly dependent columns, implies $\det A = 0$).
- $\det(AB) = \det A \det B$.
- Determinant of...
 - A diagonal matrix: The product of the diagonal entries (pull out the terms, and then note that the remaining identity matrix has determinant 1).
 - An upper triangular matrix: The product of the diagonal entries (column reduction to make it into a diagonal matrix, and then the property above).

Chapter 4

Introduction to Spectral Theory

10/1: • **Difference equation:** Like a differential equation, but instead of writing a differentials, you write differences.

• Suppose we want to solve $x_{n+1} = Ax_n$ with x_0 given.

– You will find that $x_n = A^n x_0$.

– This gets hard to compute, so we want to find a way to simplify the computation.

• Thus, we want to diagonalize the matrix, and this concept is inherently linked to eigenvalues and eigenvectors.

– If you can decompose the x_0 into a linear combination of eigenvectors, then you can simplify the computation a lot:

$$x_n = \sum \alpha_i A^n v_i = \sum \alpha_i \lambda_i^n v_i$$

– An $n \times n$ matrix will have n eigenvalues. You want n linearly independent eigenvectors, creating an eigenbasis.

• To find eigenvalues and eigenvectors, we need to solve $Ax = \lambda x$, i.e., $(A - \lambda I)x = 0$. Thus, $\ker(A - \lambda I) \neq \{0\}$, so $\det(A - \lambda I) = 0$.

• The eigenvalues of A are independent of the choice of basis of the domain of A or the range.

10/4: • We need to know everything in Treil (2017).

– We don't need to know the applications sections, but you should be interested.

• **Spectral theory:** Decomposing a linear operator.

• Let $A : V \rightarrow V$ be a linear operator. $\lambda \in \mathbb{C}$ is an eigenvalue if there exists $x \in V$ nonzero such that $Ax = \lambda x$.

– Let A be an $n \times n$ matrix over \mathbb{C} or \mathbb{R} .

– The eigenvalues are the roots of the polynomial $\det(A - \lambda I) = 0$ in λ .

• Things we want to do:

– Given A , find the eigenvalues and eigenvectors (solve $(A - \lambda I)x = 0$).

– In order to simplify A , make it a diagonal matrix:

$$A = S \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} S^{-1}$$

- Eigenvalues are independent of the choice of basis.

- From the book, we have that

$$[A]_{\mathcal{A}\mathcal{A}} = [S]_{\mathcal{A}\mathcal{B}}[B]_{\mathcal{B}\mathcal{B}}[S]_{\mathcal{A}\mathcal{B}}^{-1}$$

- It follows that

$$A - \lambda I = [S]_{\mathcal{A}\mathcal{B}}(B - \lambda I)[S]_{\mathcal{A}\mathcal{B}}^{-1}$$

so

$$\det(A - \lambda I) = \det([S]_{\mathcal{A}\mathcal{B}}(B - \lambda I)[S]_{\mathcal{A}\mathcal{B}}^{-1}) = \det([S]_{\mathcal{A}\mathcal{B}}[S]_{\mathcal{A}\mathcal{B}}^{-1}(B - \lambda I)) = \det(B - \lambda I)$$

- If $p(z) = (z - \lambda)^k q(z)$, then k is the **algebraic multiplicity** of λ . The **geometric multiplicity** of λ is $\dim \ker(A - \lambda I)$.

- These terms are not always the same, but they are related.

- Diagonalization:

- Given A that corresponds to $T : V \rightarrow V$, can we find a basis of V in which the operator is a diagonal matrix?

- $A = SDS^{-1}$ iff there exists a basis of V consisting of the eigenvectors of A .

- Proves $A^N = SD^N S^{-1}$ via $A^2 = SDS^{-1}SDS^{-1} = SDIDS^{-1} = SD^2 S^{-1}$.

- Let A be an $n \times n$ matrix over \mathbb{F} . If $\lambda_1, \dots, \lambda_r$ are distinct eigenvalues, then their eigenvectors are linearly independent.

- Prove with induction contradiction argument. Assume true for \mathbf{v}_{r-1} . Then

$$0 = (A - \lambda_r I)[\mathbf{v}_1 + \dots + \mathbf{v}_r] = (\lambda_1 - \lambda_r)\mathbf{v}_1 + \dots + (\lambda_{r-1} - \lambda_r)\mathbf{v}_{r-1}$$

- Implies $\lambda_r = \lambda_i$ for all $i \in [r-1]$, a contradiction.

- If A has n distinct eigenvalues, then A is diagonalizable.

- If $A : V \rightarrow V$ has n complex eigenvalues, then A is diagonalizable iff the algebraic multiplicity equals the geometric multiplicity for each eigenvalue.

- Goes through a sample diagonalization with $\begin{pmatrix} 1 & 2 \\ 8 & 1 \end{pmatrix}$.

- We have

$$A - \lambda I = \begin{pmatrix} 1 - \lambda & 2 \\ 8 & 1 - \lambda \end{pmatrix}$$

so

$$0 = \det(A - \lambda I) = (1 - \lambda)^2 - 16$$

- It follows that $\lambda = 5, -3$.

- This yields

$$\begin{pmatrix} 1 & 2 \\ 8 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 2 & -2 \end{pmatrix}^{-1}$$

by inspection.

- As another example, consider $\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$.

- Here, we have $\lambda = 1 \pm 2i$.

Chapter 5

Inner Product Spaces

10/6: • We define

$$\ell^2(\mathbb{R}) = \left\{ \{a_n\}_{n \geq 1} \subset \mathbb{R} : \sum_1^\infty |a_n|^2 < \infty \right\}$$

• **Inner product:** A map $V \times V \rightarrow \mathbb{F}$ that takes $(\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x} \cdot \mathbf{y}$. Denoted by $\cdot, (\cdot, \cdot), \langle \cdot, \cdot \rangle$.

• Properties of the inner product:

- $(\mathbf{x}, \mathbf{y}) = \overline{(\mathbf{y}, \mathbf{x})}$ (symmetry).
- $(\alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z}) = \alpha(\mathbf{x}, \mathbf{z}) + \beta(\mathbf{y}, \mathbf{z})$ (linearity).
- $(\mathbf{x}, \mathbf{x}) \geq 0$.
- $(\mathbf{x}, \mathbf{x}) = 0$ iff $\mathbf{x} = 0$.

• If $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i y_i$$

• If $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$, then

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n x_i \bar{y}_i$$

• If $f, g \in \mathbb{P}_n(t)$, then

$$(f, g) = \int_{-1}^1 f \bar{g} dt$$

- The conjugate of a polynomial is the polynomial with the conjugate of the coefficients of the original polynomial. Symbolically, if $f = \sum_{i=0}^n \alpha_i t^i$ is a polynomial, then $\bar{f} = \sum_{i=0}^n \bar{\alpha}_i t^i$.

• It is a fact that

$$\left| \sum_{n=1}^{\infty} a_n \bar{b}_n \right| \leq \|(a_n)_{n \geq 1}\| \|(b_n)_{n \geq 1}\|$$

• Suppose we want to define the inner product between two matrices.

- A common one is

$$(A, B) = \text{tr}(B^* A)$$

where $B^* = \bar{B}^T = \overline{B^T}$ is the conjugate transpose.

- We define the norm as a function $V \rightarrow [0, \infty)$ given by

$$\|\mathbf{x}\| = \sqrt{(\mathbf{x}, \mathbf{x})}$$

- Properties of the norm.

- $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$.
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.
- $\|\mathbf{x}\| = 0$ iff $\mathbf{x} = 0$.

- In \mathbb{R}^n ,

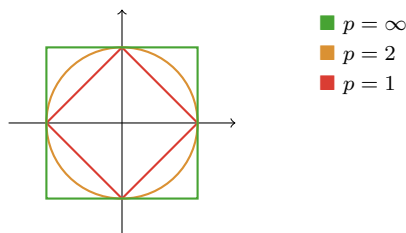


Figure 5.1: The unit ball of norms corresponding to $p = 1, 2, \infty$.

- The standard norm is

$$\|\mathbf{x}\| = \sqrt{\sum |x_i|^2}$$

- We can also define

$$\|\mathbf{x}\|_p = \sqrt[p]{\sum |x_i|^p}$$

- We can even define

$$\|\mathbf{x}\|_\infty = \max |x_i|$$

- And we can prove that all of these are valid norms.
- Only the norm corresponding to ℓ^2 is given by an inner product, but all the other quantities are still norms as defined by the properties (see Treil (2017)).
- Figure 5.1 shows the unit ball of each norm, i.e., the set of all points which have norm 1.

- The parallelogram rule:

$$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$$

- Orthogonality: Given \mathbf{v}, \mathbf{w} , if $\mathbf{v} \perp \mathbf{w}$, then $(\mathbf{v}, \mathbf{w}) = 0$.

- In particular, if $\mathbf{v} \perp \mathbf{w}$, then

$$\|\mathbf{v} + \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2$$

- Let E be a subspace of V . If $\mathbf{v} \perp E$, then $\mathbf{v} \perp \mathbf{e}$ for all $\mathbf{e} \in E$, i.e., $\mathbf{v} \perp$ a set of vectors spanning E .
- Any set of orthogonal vectors is linearly independent. Thus, if V is n dimensional, then $\mathbf{v}_1, \dots, \mathbf{v}_n$ orthogonal is a basis.
- Let E be a subspace of V . Take $\mathbf{v} \in V$. We want to define the projection $P_E \mathbf{v}$ of \mathbf{v} onto E .
 - We have that $P_E \mathbf{v} \in E$ and $\mathbf{v} - P_E \mathbf{v} \perp E$.
 - Additionally, we have that

$$\|\mathbf{v} - P_E \mathbf{v}\| \leq \|\mathbf{v} - \mathbf{e}\|$$

for all $\mathbf{e} \in E$.

- Lastly, we have that $P_E \mathbf{v}$ is unique.
- If we receive a basis of a vector space, how do we create out of that a basis that is orthogonal? The process of doing this is called **Gram-Schmidt orthogonalization**.
 - We keep \mathbf{v}_1 , subtract $P_{\mathbf{v}_1} \mathbf{v}_2$ from \mathbf{v}_2 , subtract $P_{\{\mathbf{v}_1, \mathbf{v}_2\}} \mathbf{v}_3$ from \mathbf{v}_3 , and on and on.
- If we are given a set of orthogonal vectors, we can normalize them by dividing each by its norm. This creates an orthonormal list. The standard basis is orthonormal.
- Let

$$E^\perp = \{v \in V : v \perp E\}$$

- It follows that $V = E \oplus E^\perp$.
- How close can we come to solving $A\mathbf{x} = \mathbf{b}$ if we cannot solve it exactly (i.e., if the columns are not linearly independent)?
 - Let A be an $m \times n$ matrix, and let $\mathbf{b} \in \mathbb{R}^m$.
 - Then the best solution is given by minimizing $\|A\mathbf{x} - \mathbf{b}\|$. We minimize this with projections. A special case of this is least squares regression! More details in Treil (2017).

10/8:

- Soug is gonna send us a hefty amount of reading for the weekend.
- Least square approximation:
 - If we want to minimize $\|A\mathbf{x} - \mathbf{b}\|$, the best we can do is project \mathbf{b} onto the range of A .
 - Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be an orthogonal basis of range A .
 - Then

$$\text{Proj}_{\text{range } A} \mathbf{b} = \sum_{k=1}^k \frac{(\mathbf{b}, \mathbf{v}_k)}{\|\mathbf{v}_k\|^2} \mathbf{v}_k$$

- Matrix equation form:

$$\text{Projection}_{\text{range } A} = A(A^*A)^{-1}A^*$$

if A^*A is invertible, where $A^* = \bar{A}^T$.

■ Soug never uses this though.

- The minimum is found when $\mathbf{b} - A\mathbf{x} \perp \text{range } A$. Implies that $\mathbf{b} - A\mathbf{x} \perp \mathbf{a}_k$ for all k . Implies $(\mathbf{b} - A\mathbf{x}, \mathbf{a}_k) = \bar{\mathbf{a}}_k^T (\mathbf{b} - A\mathbf{x}) = 0$.
- Note that we're letting $\bar{\mathbf{a}}_k^T$ be the row vector

$$\bar{\mathbf{a}}_k^T = (\bar{a}_{1,k} \quad \cdots \quad \bar{a}_{n,k})$$

- We also have $\bar{A}^T (\mathbf{b} - A\mathbf{x}) = 0$, from which it follows that $A^*A\mathbf{x} = A^*\mathbf{b}$, so $\mathbf{x} = (A^*A)^{-1}A^*\mathbf{b}$. Thus, $\text{Proj}_{\text{range } A} = Ax$, so $\text{Proj}_{\text{range } A} = A(A^*A)^{-1}A^*\mathbf{b}$.

- Adjoint of a linear map $T : V \rightarrow W$ is the A^* discussed above.
 - First, we'll do this for matrices. And then we'll do it for any finite-dimensional vector space.
 - Let A be an $m \times n$ matrix. We claim then that

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y})$$

for all $\mathbf{x} \in \mathbb{C}^n, \mathbf{y} \in \mathbb{C}^m$. Proof:

$$\begin{aligned} (A\mathbf{x}, \mathbf{y}) &= \bar{\mathbf{y}}^T A\mathbf{x} \\ &= \mathbf{y}^* A\mathbf{x} \\ &= (A^*\mathbf{y})^* \mathbf{x} \\ &= (\mathbf{x}, A^*\mathbf{y}) \end{aligned}$$

- Properties of the adjoint:

$$(AB)^T = B^T A^T$$

$$(AB)^* = B^* A^*$$

$$(A^*)^* = A$$

- A^* is the unique matrix B such that $(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, B\mathbf{y})$.
- Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a basis of V , and let $\mathbf{w}_1, \dots, \mathbf{w}_m$ be a basis of W .
- Definition of A^* : If $(A\mathbf{x}, \mathbf{y}) = (y, A^*\mathbf{x})$ for all $\mathbf{x} \in V$ and $\mathbf{y} \in W$.
- But it's not enough to define something; we have to check that it exists.
- If $[A]_{AB}$, then $[A^*]_{AB}$.
- More properties (give criteria for solving systems of equations):

$$\ker A^* = (\text{range } A)^\perp$$

$$\ker A = (\text{range } A^*)^\perp$$

$$\text{range } A = (\ker A^*)^\perp$$

$$\text{range } A^* = (\ker A)^\perp$$

■ Soug proves these.

• Isometries and unitary operators.

- $U : X \rightarrow Y$ is an isometry if $\|\mathbf{x}\| = \|U\mathbf{x}\|$ for all $\mathbf{x} \in X$. It is an isometry because it preserves the distance between points.
- It immediately follows that $\|\mathbf{x}_1 - \mathbf{x}_2\| = \|U\mathbf{x}_1 - U\mathbf{x}_2\| = \|U(\mathbf{x}_1 - \mathbf{x}_2)\|$.
- This definition is equivalent to an inner product one: $(\mathbf{x}, \mathbf{y}) = (U\mathbf{x}, U\mathbf{y})$. This follows from the definition of the norm.
- We have

$$(\mathbf{a}, \mathbf{b}) = \frac{1}{4} \sum_{\alpha=\pm 1, \pm i} \alpha \|\mathbf{a} + \alpha \mathbf{b}\|^2$$

■ $(a+b)^2 - (a-b)^2 = 4ab$ for any $a, b \in \mathbb{R}$, so $ab = \frac{1}{4}[(a+b)^2 - (a-b)^2]$. Thus, in a real inner product space,

$$(\mathbf{a}, \mathbf{b}) = \frac{1}{4} \left(\|\mathbf{a} + \mathbf{b}\|^2 - \|\mathbf{a} - \mathbf{b}\|^2 \right)$$

■ It follows that isometries preserve inner products.

- U is an isometry if and only if $U^*U = I$. Proof:

$$(\mathbf{x}, \mathbf{x}) = (U\mathbf{x}, U\mathbf{y}) = (U^*U\mathbf{x}, \mathbf{x})$$

$$(\mathbf{x}, \mathbf{y}) = (U\mathbf{x}, U\mathbf{y}) = (U^*U\mathbf{x}, \mathbf{y})$$

$$(\mathbf{x}, \mathbf{y}) = (U\mathbf{x}, U\mathbf{y}) = (\mathbf{x}, \mathbf{y})$$

for all \mathbf{y} .

- An isometry is unitary if it is invertible.

■ Thus, $U : X \rightarrow Y$ an isometry is unitary iff $\dim X = \dim Y$.

- Note that it follows that $U^* = U^{-1}$ for U an isometry.
- U unitary implies $|\det U| = 1$, so λ an eigenvalue of U implies that $|\lambda| = 1$.
- A is diagonalizable iff it has an orthogonal basis of eigenvectors.

Chapter 6

Structure of Operators on Inner Product Spaces

- 10/11:
- Spectral decomposition of self-adjoint linear maps.
 - Can we write a map in term of the eigenvalues only?
 - Let $A : X \rightarrow X$ be linear and self-adjoint. Where $\dim X < \infty$.
 - Let A have eigenvalues $\lambda_1, \dots, \lambda_n$ and eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. There is an orthonormal basis of X consisting of eigenvectors of A . An operator is self-adjoint if $A = A^*$.
 - If A is self-adjoint, then A can be written as diagonal with the eigenvalues on the diagonal with respect to some orthonormal basis of eigenvectors.
 - Let $\mathbb{F} = \mathbb{C}$.
 - If there exists an orthonormal basis u_1, \dots, u_n of X such that A is triangular, then $A = UTU^*$ where U is unitary and T is upper triangular.
 - Proved with induction on $\dim X$.
 - $\dim X = 1$ is clear.
 - Assume for $\dim X = n - 1$, WTS for $\dim X = n$.
 - The subspace has a basis $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$ such that A has a diagonal form.
 - Let $u \in X$ be linearly independent of $\mathbf{v}_1, \dots, \mathbf{v}_{n-1}$.
 - Let λ be the remaining eigenvalue and u the corresponding eigenvector. Let $E = \text{span}(u)$. Then make the matrix λ in the upper left corner, and block diagonal with “ A_{n-1} ” in the bottom right corner, zeroes everywhere else.
 - **Self-adjoint** (matrix A): A linear map $A : X \rightarrow X$ where $\dim X < \infty$ such that $A = A^*$.
 - Similarly, $(Ax, y) = (x, Ay)$.
 - A self-adjoint implies all eigenvalues are real, eigenvectors corresponding to different eigenvalues are orthogonal.
 - Soug proves this.
 - **Strictly positive** (operator A): A self-adjoint operator $A : X \rightarrow X$ such that $(Ax, x) > 0$ for all $x \neq 0$. Also known as **positive definite**.
 - Implies that all eigenvalues are strictly positive.
 - **Nonnegative** (operator A): A self-adjoint operator $A : X \rightarrow X$ such that $(Ax, x) \geq 0$ for all $x \neq 0$. Also known as **definite**.

- All eigenvalues are nonnegative.
- Suppose $A \geq 0$ is self-adjoint. Then there exists a unique self-adjoint $B \geq 0$ such that $B^2 = A$.
 - A self-adjoint is diagonal (wrt. some basis).
 - A positive means that all eigenvalues (diagonal entries) are positive.
 - Thus, take

$$B = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{\lambda_n} \end{pmatrix}$$

- Suppose $B^2 = A$, $C^2 = A$. Then we have an orthonormal basis corresponding to B and an orthonormal basis corresponding to C . It follows that $B^2 = C^2 = A$. Write B^2x and C^2x in terms of their bases; will necessitate that the bases are the same.

10/13:

- If we get yes/no questions, we don't have to justify.
- Cauchy-Schwarz inequality:

$$|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\| \|\mathbf{y}\|$$

- Real spaces, V vs. (\cdot, \cdot) inner product.
- Proof:

$$\begin{aligned} 0 &\leq \|\mathbf{x} + t\mathbf{y}\|^2 \\ &= t^2 \|\mathbf{y}\|^2 + 2t(\mathbf{x}, \mathbf{y}) + \|\mathbf{x}\|^2 \end{aligned}$$

Thus, the discriminant must be less than zero (because the whole polynomial is positive, so the discriminant [the opposite of the x^0 term of the factored form of the polynomial] must be less than zero so the polynomial doesn't get dragged down to negative values):

$$(\mathbf{x}, \mathbf{y})^2 - \|\mathbf{x}\|^2 \|\mathbf{y}\|^2 \leq 0$$

Taking square roots of both sides proves the desired inequality.

- Recall that if $A^* = A$, then all eigenvalues are real and all eigenvectors of distinct eigenvalues are orthogonal to each other.
- **Normal** (matrix): A matrix N such that $N^*N = NN^*$.
 - Examples: Diagonal, self-adjoint, and unitary operators are all normal.
- Any normal operator in a complex vector space has an orthonormal set of eigenvectors, e.g., $N = UDU^*$.
 - Proof: N is upper triangular wrt. some basis (because all matrices are). WTS any normal upper triangular matrix is diagonal. Done by induction on the dimension of N from $n = 2$.
 - Assume the claim for every $(n-1) \times (n-1)$ normal upper triangular matrix.
 - Let

$$N = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ 0 & & & \\ \vdots & & N_1 & \\ 0 & & & \end{pmatrix}$$

(we know every normal matrix can be written in this upper triangular form)

- Then just compute NN^* and N^*N . Knowing they have to be equal, we have that $a_{12} = \cdots = a_{1n} = 0$.

- We can also prove from the above (block diagonal multiplication) that N_1 is normal. Thus, it's diagonal, too. Therefore, the whole thing is diagonal.
- N is normal if and only if $\|N\mathbf{x}\| = \|N^*\mathbf{x}\|$.
 - Proof: $(N\mathbf{x}, N\mathbf{y}) = (N^*N\mathbf{x}, \mathbf{y}) = (NN^*\mathbf{x}, \mathbf{y}) = (N^*\mathbf{x}, N^*\mathbf{y})$. This is equivalent to the desired condition.
- If A is nonnegative and $(A\mathbf{e}_k, \mathbf{e}_k) = a_{kk}$, then

$$\sum_{i,j=1}^n a_{ij} \mathbf{x}_i \mathbf{x}_j$$

- **Positive definite** (matrix): An $n \times n$ self-adjoint matrix such that $(A\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \in X$.
- Let $A : X \rightarrow Y$, $\dim X = \dim Y$. Then AA^* is positive semidefinite. And there exists a unique square root $R = \sqrt{A^*A}$.
 - Proof: $(A^*A\mathbf{x}, \mathbf{x}) = (A\mathbf{x}, A\mathbf{x}) = \|A\mathbf{x}\|^2 \geq 0$.
- **Modulus** (of A): The matrix $|A| = \sqrt{A^*A}$.
- Check $\| |A|\mathbf{x} \| = \|A\mathbf{x}\|$.

$$\| |A|\mathbf{x} \|^2 = (|A|\mathbf{x}, |A|\mathbf{x}) = (|A|^*|A|\mathbf{x}, \mathbf{x}) = (A^*A\mathbf{x}, \mathbf{x}) = (A\mathbf{x}, A\mathbf{x}) = \|A\mathbf{x}\|^2$$

- Let $A : X \rightarrow X$ be a linear operator. Then $A = U|A|$ where U is unitary.
- Look at singular matrices.
- Recall that if $A : X \rightarrow Y$, we have that A^*A is semidefinite, positive, and self adjoint.
 - Thus, there exists a unique matrix $R = \sqrt{A^*A} \geq 0$, which we define to be $|A| = \sqrt{A^*A}$.
- Polar form of a matrix:

$$A = U|A|$$

- This may not be unique!
- Proof: Suppose $A\mathbf{x} = U(|A|\mathbf{x})$. $A\mathbf{x} \in \text{range } A$, and $|A|\mathbf{x} \in \text{range}(|A|)$. $\mathbf{x} \in \text{range}(|A|)$ implies that there exists $\mathbf{v} \in X$ such that $\mathbf{x} = |A|\mathbf{v}$.
- Define $U\mathbf{x} = A\mathbf{x}$. U is a well-defined linear map.
- $\|U_0\mathbf{x}\| = \|A\mathbf{x}\| = \||A|\mathbf{v}\| = \|\mathbf{x}\|$.
- U is an isometry.
- $\text{range } |A| \rightarrow X$.
- Use $\ker A = \ker |A| = (\text{range } A)^\perp$ to extend U_0 to U : $U = U_0 + U_1$.
- **Singular values** (of a matrix): The eigenvalues of $|A|$.
 - So if $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A^*A , the singular values of A are $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$.
- Let $A : X \rightarrow Y$ be a linear map.
 - Let $\sigma_1, \dots, \sigma_n$ be the singular values of A . Then $\sigma_1, \dots, \sigma_n > 0$.
 - Additionally, if $\mathbf{v}_1, \dots, \mathbf{v}_n$ is an orthonormal basis of eigenvectors of A^*A , then the list of n vectors $\mathbf{w}_1, \dots, \mathbf{w}_n$ defined by $\mathbf{w}_k = 1/\sigma_k A\mathbf{v}_k$ for each $k = 1, \dots, n$ is orthonormal.

10/15:

■ Proof:

$$(\mathbf{w}_k, \mathbf{w}_j) = \frac{1}{\sigma_k \sigma_j} (A\mathbf{v}_k, A\mathbf{v}_j) = \frac{1}{\sigma_k \sigma_j} = \frac{1}{\sigma_k \sigma_j} (A^* A \mathbf{v}_k, \mathbf{v}_j) = \frac{\sigma_k^2}{\sigma_k \sigma_j} (\mathbf{v}_k, \mathbf{v}_j) = 0$$

and

$$\|\mathbf{w}_k\| = \frac{1}{\sigma_k} \|A\mathbf{v}_k\| = \frac{1}{\sigma_k} \||A|\mathbf{v}_k\| = 1$$

– Schmidt decomposition of A :

$$A\mathbf{x} = \sum_{k=1}^r \sigma_k (\mathbf{x}, \mathbf{v}_k) \mathbf{w}_k$$

■ This is because $\mathbf{x} = \sum (\mathbf{x}, \mathbf{v}_k) \mathbf{v}_k$, so by the above,

$$A\mathbf{x} = \sum_{k=1}^n (\mathbf{x}, \mathbf{v}_k) A\mathbf{v}_k = \sum_{k=1}^r \sigma_k (\mathbf{x}, \mathbf{v}_k) \mathbf{w}_k$$

• **Operator norm:** $\|A\| = \max\{\|A\mathbf{x}\| : \|\mathbf{x}\| \leq 1\}$.

• Properties of the operator norm:

- $\|A\mathbf{x}\| \leq \|A\| \|\mathbf{x}\|$.
- $\|\alpha A\| = |\alpha| \|A\|$.
- $\|A + B\| \leq \|A\| + \|B\|$.
- $\|A\| \geq 0$.
- $\|A\| = 0$ iff $A = 0$.

• **Frobenious norm:** The norm $\|A\|_2^2 = \text{tr}(A^* A)$.

• The operator norm is always less than or equal to the Frobenius norm.

• If $A : \mathbb{F}^n \rightarrow \mathbb{F}^n$, then $A = W\Sigma V^*$ where σ is a diagonal matrix of nonzero singular values.

• The operator norm of A is the largest of the singular values.

• An orthogonal matrix can be decomposed to a block-diagonal matrix of rotations.

10/18:

• Soug tests what he teaches and doesn't give super tricky questions.

• Structure of orthogonal matrices.

• **Orthogonal (matrix):** A unitary matrix U with all elements real and $|\det U| = 1$.

• Theorem: Let U be an orthogonal operator on \mathbb{R}^n such that $\det U = 1$. Then there exists an orthonormal basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ such that with respect to this basis,

$$U = \begin{pmatrix} R_{\phi_1} & & & \mathbf{0} \\ & \ddots & & \\ & & R_{\phi_k} & \\ \mathbf{0} & & & I_{n-2k} \end{pmatrix}$$

where each R_{ϕ_i} is a 2×2 rotation matrix.

- If you are in \mathbb{R}^7 for example, you would be able to express U as a composition of at most 3 rotation maps and the identity map.
- Each rotation map acts on two orthonormal vectors.
- Proof: $P(\lambda)$ is the n -degree characteristic polynomial $\det(U - \lambda I) = 0$. The eigenvalues are the roots of it.

- $p(\lambda) = 0$ if and only if $p(\bar{\lambda}) = 0$.
 - $\lambda \in \mathbb{C}$ is an eigenvalue with eigenvector $\mathbf{u} \neq 0$ iff $U\mathbf{u} = \lambda\mathbf{u}$ and $U\bar{\mathbf{u}} = \bar{\lambda}\bar{\mathbf{u}}$.
- Recall that U unitary implies $|\lambda| = 1$.
 - Proof^[1]: $\|U\mathbf{x}\| = \|\mathbf{x}\|$ and $U\mathbf{x} = \lambda\mathbf{x}$. Thus,

$$\|U\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\| = \|\mathbf{x}\|$$

and since $\mathbf{x} \neq 0$, we can divide by $\|\mathbf{x}\|$, so $|\lambda| = 1$.

- Let $\mathbf{u} = \text{Re } \mathbf{u} + i \text{Im } \mathbf{u}$.
- It follows that we may define

$$\mathbf{x} = \text{Re } \mathbf{u} = \frac{\mathbf{u} + \bar{\mathbf{u}}}{2} \qquad \mathbf{y} = \text{Im } \mathbf{u} = \frac{\mathbf{u} - \bar{\mathbf{u}}}{2}$$

- Thus, $\mathbf{u} = \mathbf{x} + i\mathbf{y}$ and $\bar{\mathbf{u}} = \mathbf{x} - i\mathbf{y}$.
- Since $U\mathbf{x} = \frac{U\mathbf{u} + U\bar{\mathbf{u}}}{2} = \frac{\lambda\mathbf{u} + \bar{\lambda}\bar{\mathbf{u}}}{2}$, $U\mathbf{y} = \text{Im}(\lambda\mathbf{u}) = \text{Re}(\lambda\mathbf{u})$.
- Since $|\lambda| = 1$, $\lambda = e^{i\alpha}$ and $\bar{\lambda} = e^{-i\alpha}$.
- It follows that $U\mathbf{x} = (\cos \alpha)\mathbf{x} - (\sin \alpha)\mathbf{y}$ and $U\mathbf{y} = (\cos \alpha)\mathbf{y} + (\sin \alpha)\mathbf{x}$.
- Thus, since $U\mathbf{x} = \text{Re } \lambda\mathbf{u}$, we have that

$$\begin{aligned} \lambda\mathbf{u} &= (\cos \alpha + i \sin \alpha)(\mathbf{x} + i\mathbf{y}) \\ &= (\cos \alpha)\mathbf{x} - (\sin \alpha)\mathbf{y} + i[(\cos \alpha)\mathbf{y} + (\sin \alpha)\mathbf{x}] \end{aligned}$$

- If E_λ is a 2 dimensional space spanned by \mathbf{x} and \mathbf{y} and invariant by U . Thus, any block of the desired matrix leaves its desired sub-block invariant.
- We also know that the eigenvectors of a unitary matrix corresponding to different eigenvalues are orthogonal.
- Thus, $\|\mathbf{x}\| = \|\mathbf{y}\| = \sqrt{2}/2\|\mathbf{u}\|$, $\mathbf{x} \perp \mathbf{y}$.
- Let \mathbf{x}, \mathbf{y} complete the theorem to form a basis of \mathbb{R}^n .
- It will follow that

$$U = \begin{pmatrix} R_\alpha & \mathbf{0} \\ \mathbf{0} & U_1 \end{pmatrix}$$

where U_1 is orthogonal, and we may repeat the process.

¹This would be a good exam question.

Chapter 7

Bilinear and Quadratic Forms

10/18: • **Bilinear form:** A function $L : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ such that

$$L(\alpha \mathbf{x}_1 + \beta \mathbf{x}_2, \mathbf{y}) = \alpha L(\mathbf{x}_1, \mathbf{y}) + \beta L(\mathbf{x}_2, \mathbf{y}) \quad L(\mathbf{x}, \alpha \mathbf{y}_1 + \beta \mathbf{y}_2) = \alpha L(\mathbf{x}, \mathbf{y}_1) + \beta L(\mathbf{x}, \mathbf{y}_2)$$

$$- L(\mathbf{x}, \mathbf{y}) = (A\mathbf{x}, \mathbf{y}).$$

• **Quadratic form:** A bilinear form $L(\mathbf{x}, \mathbf{x})$.

- (\mathbf{x}, \mathbf{x}) is a polynomial of degree 2 in $\mathbf{x}_1, \dots, \mathbf{x}_n$:

$$L(\lambda \mathbf{x}, \lambda \mathbf{x}) = (\lambda \mathbf{x}, \lambda \mathbf{x}) = \lambda^2 (\mathbf{x}, \mathbf{x})$$

• We have that

$$(A\mathbf{x}, \mathbf{x}) = (A\lambda \mathbf{x}, \lambda \mathbf{x}) = \lambda^2 (A\mathbf{x}, \mathbf{x}) = \sum_{j,i=1}^n \alpha_{j,i} \mathbf{x}_i \mathbf{x}_j$$

• The general form of a quadratic form:

- Can any quadratic form on \mathbb{R}^n be written as $(A\mathbf{x}, \mathbf{x})$?

10/20: • Bilinear forms are linear in each argument when keeping the other fixed.

• Quadratic forms $Q(\mathbf{x}) = L(\mathbf{x}, \mathbf{x})$ are quadratic polynomials in the coordinates of \mathbf{x} .

- In particular, $Q(\lambda \mathbf{x}) = |\lambda|^2 Q(\mathbf{x})$.

• If Q quadratic is real, then $Q(\mathbf{x}) = (A\mathbf{x}, \mathbf{x})$ where A is some square matrix.

- If $\mathbf{e}_1, \dots, \mathbf{e}_n$ is an orthonormal basis of \mathbb{R}^n , then there exists a unique $A = A^*$ such that $(A)_{ij} = L(\mathbf{e}_i, \mathbf{e}_j)$.

- Keeping $\mathbf{x} = \sum_{i=1}^n \mathbf{x}_i \mathbf{e}_i$ fixed, we have

$$\begin{aligned} Q(\mathbf{x}) &= L(\mathbf{x}, \mathbf{x}) \\ &= L\left(\sum_{i=1}^n \mathbf{x}_i \mathbf{e}_i, \sum_{j=1}^n \mathbf{x}_j \mathbf{e}_j\right) \\ &= \sum_{i=1}^n \mathbf{x}_i L\left(\mathbf{e}_i, \sum_{j=1}^n \mathbf{x}_j \mathbf{e}_j\right) \\ &= \sum_{i,j=1}^n \mathbf{x}_i \mathbf{x}_j \underbrace{L(\mathbf{e}_i, \mathbf{e}_j)}_{A_{ij}} \end{aligned}$$

- We have that

$$\begin{aligned}(A\mathbf{x}, \mathbf{x}) &= (UDU^{-1}\mathbf{x}, \mathbf{x}) \\ &= (DU^{-1}\mathbf{x}, U^{-1}\mathbf{x}) \\ &= \sum_{i=1}^n \lambda_i \underbrace{(U^{-1}\mathbf{x})_i}_{\mathbf{y}_i} \underbrace{(U^{-1}\mathbf{x})_i}_{\mathbf{y}_i}\end{aligned}$$

- Can we characterize the set $\{\mathbf{x} : (A\mathbf{x}, \mathbf{x}) = 1\}$?
 - Note that this set is equivalent to $\{\mathbf{y} : (D\mathbf{y}, \mathbf{y}) = 1\}$ by the above. This set is a rotation of the previous one. Ellipse?
- Positive quadratic form:
 - Q is positive definite if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$ and Q is positive semidefinite if $Q(\mathbf{x}) \geq 0$ for all $\mathbf{x} \neq \mathbf{0}$.
 - Take a self-adjoint matrix $A = A^*$. It is positive definite if $Q(\mathbf{x}) = (A\mathbf{x}, \mathbf{x})$ is positive definite.
- Theorem: If $A = A^*$, then
 1. A is positive definite if and only if all eigenvalues of A are positive.
 2. A is positive semidefinite if and only if all eigenvalues of A are nonnegative.
 3. A is negative semidefinite if and only if all eigenvalues of A are nonpositive.
 4. A is negative definite if and only if all eigenvalues of A are negative.
 5. A is indefinite if and only if the eigenvalues of A have positive and negative values.
- Theorem: $A = A^*$ is positive definite iff $\det A_k > 0$ for all $k = 1, \dots, n$ where A_k is the upper left $k \times k$ submatrix.
- Minimax representation of eigenvalues of a self-adjoint A .
 - Let E be a subspace of X where $\dim X < \infty$. We define $\text{codim}(E) = \dim E^\perp$.
 - Thus, $\dim E + \text{codim } E = \dim X$.
 - Theorem: Let $A = A^*$, $\lambda_1 \geq \dots \geq \lambda_n$ eigenvalues of A . Then

$$\lambda_k = \max_{\substack{E \text{ subspace} \\ \dim E = k}} \min_{\substack{\mathbf{x} \in E \\ \|\mathbf{x}\|=1}} (A\mathbf{x}, \mathbf{x}) = \min_{\substack{F \text{ subspace} \\ \text{codim } F = k-1}} \max_{\substack{\mathbf{x} \in F \\ \|\mathbf{x}\|=1}} (A\mathbf{x}, \mathbf{x})$$

- Proof: A diagonal equals $(\lambda_1, \dots, \lambda_n)$.
- An orthonormal basis of X such that $\dim E = k$, $\text{codim } F = k-1$, $\dim F = n-k+1$.
- There exists an $\mathbf{x}_0 \neq \mathbf{0}$ such that $\mathbf{x}_0 \in E \cap F$.
- Note that if $B = B^*$, then the max and min of $(B\mathbf{x}, \mathbf{x})$ over the unit sphere is the maximal and minimal eigenvalue of B .
- Thus,

$$\min_{\substack{\mathbf{x} \in E \\ \|\mathbf{x}\|=1}} (A\mathbf{x}, \mathbf{x}) \leq (A\mathbf{x}_0, \mathbf{x}_0) \leq \max_{\substack{\mathbf{x} \in F \\ \|\mathbf{x}\|=1}} (A\mathbf{x}, \mathbf{x})$$

- This is true for any E, F subspaces. $\dim E = k$, $\text{codim } F = k-1$, $E_0 = \text{span}(\mathbf{e}_1, \dots, \mathbf{e}_k)$ and $F_0 = \text{span}(\mathbf{e}_k, \dots, \mathbf{e}_n)$.
- Thus,

$$\min_{\substack{E_0 \\ \|\mathbf{x}\|=1}} (A\mathbf{x}, \mathbf{x}) = \lambda_k = \max_{\substack{F_0 \\ \|\mathbf{x}\|=1}} (A\mathbf{x}, \mathbf{x})$$

■ Additionally,

$$\lambda_{k_1} \leq \max_{\dim E=k} \min_{\mathbf{x}} (A\mathbf{x}, \mathbf{x}) \leq \min_{\substack{F \\ \text{codim } F=k-1}} \max_{\mathbf{x}} (A\mathbf{x}, \mathbf{x}) \leq \lambda_k$$

- Corollary: Let $A = A^* = (a_{jk})_{1 \leq j, k \leq n}$ with eigenvalues $\lambda_1, \dots, \lambda_n$ listed in decreasing order. Let $\tilde{A} = (a_{j,k})_{1 \leq j, k \leq n-1}$ with eigenvalues μ_1, \dots, μ_{n-1} listed in decreasing order. Then $\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \dots \geq \mu_{n-1} \geq \lambda_n$.
 - Consider $(A\mathbf{x}, \mathbf{x})$ on $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, but then restrict yourself to $\mathbf{x} \in \mathbb{R}^{n-1}$ on $\{\mathbf{e}_1, \dots, \mathbf{e}_{n-1}\}$.

Chapter 8

Dual Spaces and Tensors

- 10/22:
- **Functional:** A linear bounded map $L : H \rightarrow F$, where H is finite dimensional (equivalent to \mathbb{R}^n).
 - **Dual space:** The set of bounded linear functionals on H . Denoted by H' , H^* .
 - If $l \leq p < \infty$, then

$$l^p = \left\{ (a_n)_{n \in \mathbb{N}} : \sum_{n=1}^{\infty} |a_n|^p < \infty \right\}$$

- Back to finite dimensions, $H' \approx \mathbb{R}^n$.
- Let $\mathbf{a}_1, \dots, \mathbf{a}_n$ be a basis of H . Then $L\mathbf{x} = (L\mathbf{a}_1, \dots, L\mathbf{a}_n) \approx \mathbb{R}^n$.
- Let $L((a_n)_{n \in \mathbb{N}}) = \sum_{n=1}^{\infty} a_n b_n$. Then $L((a_n)_{n \in \mathbb{N}})$ will be bounded if and only if $(b_n)_{n \in \mathbb{N}} \in l^q$ where $1 < p < q$ where $\frac{1}{q} + \frac{1}{p} = 1$.
- **Young's inequality:** The statement

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

- We have $|\sum a_n b_n| \leq \|a_n\|_p \|b_n\|_q$.
- Conclusion:

$$\sum \frac{|a_n| |b_n|}{\|a_n\|_p \|b_n\|_q} = 1$$

- We can define H'' , too. This contains linear functionals on H' .
- We know that $L(x) = \langle x, L \rangle = x(L)$. $x \in H''$.
- Riesz representation theorem: Let H have an inner product. $L \in H'$ if and only if there exists a unique $y \in H$ such that $L(x) = \langle x, y \rangle$.
 - Gives us a way to identify all bounded linear functionals on H .
 - In finite dimensions, $L(x)$, where $x = \sum_1^n \alpha_i a_i$ gives us $L(x) = \sum_1^n \alpha_i L(a_i)$.

Chapter 9

Advanced Spectral Theory

- 10/22:
- Let $p(z) = \sum_{i=0}^n a_i z^i$ be a polynomial. Let A be an $n \times n$ matrix. We let $p(A) = \sum_{i=0}^n a_i A^i$.
 - Theorem: If A is an $n \times n$ and $p(\lambda) = \det(A - \lambda I)$, then $p(A) = 0$.
 - We know that $p(\lambda) = a(z - \lambda_1) \cdots (z - \lambda_n)$ where $\lambda_1, \dots, \lambda_n$ are the eigenvalues.
 - Thus $p(A) = a(A - \lambda_1 I) \cdots (A - \lambda_n I)$.
 - If you are in \mathbb{R}^n and have this property, you can factorize your matrix.
 - Thus, $p(A)\mathbf{x} = \mathbf{0}$ since \mathbf{x} can be decomposed into a linear combination of eigenvectors of A , which will be taken to 0 one by one by the terms of $p(A)$.
 - $\sigma(B) = \{\text{eigenvalues of } B\}$ is known as the **spectrum** of B .
 - If p is an arbitrary polynomial and A is $n \times n$, then μ is an eigenvalue of $p(A)$ if and only if $\mu = p(\lambda)$ where λ is an eigenvalue of A . In essence, $\sigma(p(A)) = p(\sigma(A))$.
 - Chapter 9 will not be on the exam. We don't have to know the generalization to infinite dimensional spaces.

References

Treil, S. (2017). *Linear algebra done wrong* [http://www.math.brown.edu/streil/papers/LADW/LADW_2017-09-04.pdf].