Math 54 Notes & Key Sheet

Andrew Tao

November 14, 2022

Chapter 1

Introduction

Lorem Ipsum

Chapter 2

Linear Equations

2.1 Linear Systems

Definition 1 (Linear Equations)**.** *Given variables $x_1, x_2, \ldots, x_n$, real or complex numbers b and coefficients $a_1, a_2, \ldots, a_n$, we introduce **Linear Equations** as an equation written in the form*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{2.1}$$

Definition 2 (Linear Systems)**.** *A **System of Linear Equations** or **Linear System** is a collection of m linear equations sharing n variables.*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n} = &\quad b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n} = &\quad b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn} = &\quad b_m \end{aligned} \tag{2.2}$$

Definition 3 (Solution Set)**.** *A **Solution** of a linear system is a list of numbers that satisfies each equation when substituted for $x_1, x_2, \ldots, x_n$. The **Solution Set** is the set of all possible solutions to the linear system. Also note that two linear systems are **equivalent** if they share the same solution set.*

Concept 1 (Existence/Uniqueness of Solutions)**.** *A linear system's solution set has a certain number of elements*

- *none* $\{\}$

- *one* $\{s\}$
- *infinitely many*

Definition 4 (Consistent)**.** *A linear system is* ***Consistent*** *is its solution set is nonempty. Otherwise, it is* ***Inconsistent****.*

Definition 5 (Matrix)**.** *We define a* $m \times n$ ***Matrix*** *as a rectangular array of objects. We can take the coefficients of a linear system as a* ***Coefficient Matrix*** *and a* ***Augmented Matrix*** *if we include the rightmost constants.*

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix} \tag{2.3}$$

Procedure 1 (Elementary Row Operations)**.** *We are allowed to preform three basic operations to create row equivalent augmented matrixes. Note that the new associated linear systems share the same solution set.*

Replacement *Replace row by sum of itself and a scaled version of another row*

Interchange *Swap two rows*

Scaling *Multiply a row by a nonzero constant*

Concept 2 (Fundamental Questions)**.** *We are faced with two fundamental questions when considering linear systems.*

- *Is a system consistent?*
- *If the system is consistent, is the solution unique?*

2.2 Row Reduction

Definition 6 (Echelon Form)**.** *A matrix is in* ***echelon form*** *if it has these three properties*

- *Nonzero rows are above any rows with all zeros*

- *Each leading entry in a row, its location called the* ***pivot position****, is in a column, called a* ***pivot column****, to the right of the leading entry of the row above it*

- *All entries in a column below a leading entry are zeros*

where leading entry refers to the first nonzero entry.
*Example (where $\blacksquare$ refers to a nonzero value and * refers to any value):*

$$\begin{bmatrix} \blacksquare & * & * & * \\ 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & \blacksquare \end{bmatrix} \tag{2.4}$$

A matrix is in ***reduced echelon form*** *is it satisfies*

- *Each nonzero row has only a leading entry of 1*

Note that each matrix is row equivalent to only one reduced echelon matrix.
Example:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.5}$$

Procedure 2 (Row Reduction Algorithm)**.** *We use a set of steps in producing a matrix in first echelon, then reduced echelon form.*

1. *Identify the leftmost nonzero column as the first pivot column. Identify the topmost entry location as the pivot position*

2. *Identify a nonzero entry as the pivot and move it to the pivot position (if not already there)*

3. *Use row operations to turn all entries below the pivot in the pivot column into zeros*

4. *Ignoring the rows containing and above the pivot position just in question, and repeat steps 1-3 on the remaining submatrix. Repeat until there are no more nonzero rows to change*

Steps 1-4 are called the forward phase. Echelon form has been achieved

5. *Using row operations, scale pivots to one and create zeros above each pivot*

   Step 5 is called the backward pase. Row echelon form has been achieved

Definition 7 (Free/Basic variables)**.** *The variables corresponding to pivot columns are called* ***basic variables****. The other variables are called* ***free variables****. With solutions of linear systems, we want to represent basic variables as parametric descriptions with free variables as parameters.*

Concept 3 (Existence/Uniqueness Theorem)**.** *The echelon form of our augmented matrix offers some useful information on existence/uniqueness.*

Existence *A linear system is consistent iff the rightmost column of the augmented matrix is not a pivot column*

Uniqueness *If the linear system is consistent, then the solution set contains one or infinitely many solutions if they have*

One *No free variables*

Infinite *At least one free variable*

Procedure 3 (Methodology for Solving Linear Systems)**.** *In summary, the following is the general procedure for using row reduction to solve linear systems*

1. *Create augmented matrix*
2. *Use row reduction algorithm to reduce to echelon form. If system is consistent, continue. Otherwise, return no solution.*
3. *Continue reducing to reduced echelon form.*
4. *Write parametrized solutions for each basic variable is expressed in terms of the free variables.*

2.3 Vectors

Definition 8 (Vector)**.** *A **vector** is defined as an element of a vector space. In the next few sections we're only going to deal with vectors which are just ordered collections of numbers.*

Definition 9 (Vector Operations)**.** *For scalar k and vectors in $\mathbb{R}^n$*

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

vector addition

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} \tag{2.6}$$

scalar multiplication

$$k\mathbf{v} = \begin{bmatrix} kv_1 \\ kv_2 \\ \vdots \\ kv_n \end{bmatrix} \tag{2.7}$$

Property 1 (Algebraic properties of vectors in $\mathbb{R}^n$)**.** *Given $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ in $\mathbb{R}^n$ and scalars c and d:*

1. $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$
2. $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$
3. $\boldsymbol{u} + \mathbf{0} = \mathbf{0} + \boldsymbol{u} = \boldsymbol{u}$
4. $\boldsymbol{u} + (-\boldsymbol{u})$
5. $c(\boldsymbol{u} + \boldsymbol{v}) = c\boldsymbol{u} + c\boldsymbol{v}$

6. $c(d\boldsymbol{u}) = (cd)\boldsymbol{u}$

7. $1\boldsymbol{u} = \boldsymbol{u}$

Definition 10 (Linear Combinations)**.** *Given vectors $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_p}$ in $\mathbb{R}^n$ and scalars $c_1, c_2, \ldots, c_p$ (called weights), we define the **linear combination***

$$\boldsymbol{y} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \ldots + c_p\boldsymbol{v_p} \tag{2.8}$$

Definition 11 (Span)**.** *For vectors $bmv_1, bmv_2, \ldots, bmv_p$ in $\mathbb{R}^n$, Span$\{bmv_1, bmv_2, \ldots, \boldsymbol{v_p}\}$ is the collection of all vectors that may be represented in the form:*

$$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \ldots + c_p\boldsymbol{v_p} \tag{2.9}$$

2.4 Matrix Equation

Definition 12 (Ax=b)**.** *Given matrix $A_{m\times n}$ and vector $\boldsymbol{x}$ in $\mathbb{R}^n$:*

$$A\boldsymbol{x} = \begin{bmatrix}\boldsymbol{a_1} & \boldsymbol{a_2} & \cdots & \boldsymbol{a_n}\end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{a_1} + x_2\boldsymbol{a_2} + \cdots + x_n\boldsymbol{a_n} = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{bmatrix} \tag{2.10}$$

Note that the row-vector rule for $A\boldsymbol{x}$ means the i-th entry in $A\boldsymbol{x}$ is the dot product of the row i in A and the vector x.

Concept 4. *Important! We now have enough definitions to establish a relationship between three concepts.*

A vector equation of the form

$$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \ldots + c_p\boldsymbol{v_p} = \boldsymbol{b} \tag{2.11}$$

has the same solution set as the L.S with the augmented matrix

$$\begin{bmatrix}\boldsymbol{a_1} & \boldsymbol{a_2} & \cdots & \boldsymbol{a_n} & \boldsymbol{b}\end{bmatrix} \tag{2.12}$$

which has the same solution set as the matrix equation

$$A\boldsymbol{x} = \boldsymbol{b} \tag{2.13}$$

A nice observation from this is that the equation $\boldsymbol{x} = \boldsymbol{b}$ has a solution iff $\boldsymbol{b}$ is a linear combination of the column vectors of A.

Theorem 1. *Given $A_{m\times n}$, the follwing statements are logically equivalent.*

1. *For each $\boldsymbol{b}$ in $\mathbb{R}^m$ the equation $A\boldsymbol{x} = \boldsymbol{b}$ has a solution (is consistent)*
2. *Each $\boldsymbol{b}$ in $\mathbb{R}^m$ is a linear combination of the column vector of A*
3. *The column vectors of A span $\mathbb{R}^m$*
4. *A has a pivot position in every row*

Definition 13 (Identity Matrix)**.** *The I_n matrix is the matrix such that $I_n\boldsymbol{x} = \boldsymbol{x}$ for every x in $\mathbb{R}^n$. It is an $n \times n$ square matrix with 1s on every row position and 0s everywhere else.*

Property 2 (Matrix Operations)**.** *Given matrix $A_{m\times n}$:*

1. $A(\boldsymbol{u} + \boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v}$
2. $A(c\boldsymbol{u}) = c(A\boldsymbol{u})$

2.5 Solution Sets

Definition 14 (Homogenous Linear Systems)**.** *A linear system is **homogenous** if it can be represented in the form $A\boldsymbol{x} = \mathbf{0}$. This always has the trivial solution, where $\boldsymbol{x} = \mathbf{0}$. This also has a nontrivial solution if there exists a nonzero vector $\boldsymbol{x}$ that also satisfies the equation.*

Definition 15 (Homogenous Linear System)**.** *A linear system is **homogenous** if it can be represented in the form $A\boldsymbol{x} = \mathbf{0}$. This always has the trivial solution, where $\boldsymbol{x} = \mathbf{0}$. This also has a nontrivial solution if there exists a nonzero vector $\boldsymbol{x}$ that also satisfies the equation.*

Theorem 2 (Nonhomogenous vs Homogenous Solution Sets)**.** *Given nonhomogenous equation $A\boldsymbol{x} = \boldsymbol{b}$ is consistent with some solution $\boldsymbol{p}$, its solution set is a translation of the solution set of the homogenous equation $A\boldsymbol{x} = \mathbf{0}$ (which passes through the origin) with any solution v_h where we can write each solution as $w = p + v_h$.*

Note, to parameterize the solution set of a system, we write a typical solution x according to the free variables as parameters.

2.6 Linear Dependence/Independence

Definition 16 (Linear Independence/Dependence)**.** *A set of vectors* $\boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$ *in* $\mathbb{R}^n$ *is* ***linearly independent*** *if there exists only the trivial solution to vector equation*

$$x_1\boldsymbol{v_1} + x_2\boldsymbol{v_2} + \ldots + x_p\boldsymbol{v_p} = \mathbf{0} \tag{2.14}$$

Else, if there exists nontrivial solutions, it is called ***linearly dependent****. We can represent the above homogenous equation as an augmented matrix and solve for solutions. The vectors are linear independent iff from this we only find the trivial solution.*

Theorem 3. *A set* $S = \boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$ *is linearly dependent iff any vector is a linear combination of the other vectors.*

Observation 1. *Given a set* $S = \boldsymbol{v_1}, \ldots, \boldsymbol{v_p}$ *in* $\mathbb{R}^n$*:*

1. *S is linearly dependent if $p > n$*

2. *S is linearly dependent if it contains the zero vector*

2.7 Linear/Matrix Transformations

Definition 17 (Transformation)**.** *A* ***transformation*** *(also known as a* ***function*** *or* ***mapping*** *T from $\mathbb{R}^n$ to $\mathbb{R}^m$ assigns some vector $T(x)$ (called the* ***image*** *of x) in $\mathbb{R}^m$ to each vector x in $\mathbb{R}^n$.*

The ***domain*** *is the set of input values for which T is defined. The* ***codomain*** *is the set within which the output values may lie. The* ***range*** *is the set of all images.*

Definition 18 (Linear Transformation)**.** ***Linear transformations*** *must follow the following conditions:*

1. *$T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$ for all $\boldsymbol{u}$, $\boldsymbol{v}$ in the domain of T*

2. *$T(c\boldsymbol{u}) = cT()$ for all scalars c and $\boldsymbol{u}$ in the domain of T*

   From this, we can also draw the conclusions that

3. *$T(\mathbf{0}) = \mathbf{0}$*

4. *$T(c_1\boldsymbol{v_1} + \ldots + c_p\boldsymbol{v_p} = c_1T(\boldsymbol{v_1}) + \ldots + dT(\boldsymbol{v_p})$*

Definition 19 (Matrix Transformations)**.** ***Matrix Multiplications****, denoted by $\boldsymbol{x} \mapsto A\boldsymbol{x}$ are transformations given by $T(x) = A\boldsymbol{x}$ for some matrix $A_{m\times n}$. The domain and codomain are $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively, and the range is the span of the columns vectors of A.*

2.8 Standard Matrix

Theorem 4. *First, note that every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation with* ***standard matrix*** *$A_{m\times n}$ given by*

$$A = \begin{bmatrix} T(\boldsymbol{e}_1) & \cdots & T(\boldsymbol{e}_n) \end{bmatrix} \tag{2.15}$$

Concept 5 (Onto and One-to-one)**.** *Given mapping $T : \mathbb{R}^n \to \mathbb{R}^m$:*

	Onto (surjective)	*One-to-one (injective)*
Definition	*Each $\boldsymbol{b}$ in the codomain is the image of at least one $\boldsymbol{x}$ in the domain*	*Each $\boldsymbol{b}$ in the range is the image of at most one $\boldsymbol{x}$ in the domain*
Relation to Ax=b	*Existence of solution*	*Uniqueness of solution*
Conditions (iff)	*1. columns of A span $\mathbb{R}^m$*	*1. $T(\boldsymbol{x}) = \mathbf{0}$ has only trivial solution* *2. columns of A are linearly independent*

Chapter 3

Matrix Algebra

3.1 Matrix Operations

Before we begin, note that the entries in a matrix $A_{m\times n}$ with m rows and n columns are denoted a_{ij} with i and j referring to the row and column number respectively. Two matrices are equal if their entries are all equal. Addition of two matrices involves adding their corresponding entires. Scalar multiplication involves scaling each entry by the scalar.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1(n-1)} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2(n-1)} & a_{2n} \\ \vdots & & & \vdots & & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{i(n-1)} & a_{1n} \\ \vdots & & & \vdots & & & \vdots \\ a_{(m-1)1} & a_{(m-1)2} & \cdots & a_{(m-1)j} & \cdots & a_{(m-1)(n-1)} & a_{(m-1)n} \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{m(n-1)} & a_{mn} \end{bmatrix} \tag{3.1}$$

Concept 6 (Types of matrices)**.** *A* ***diagonal matrix*** *$A_{m\times n}$ is a square matrix whose non****diagonal entries*** *are zero. A special case is the* ***identity matrix*** *$I_{n\times n}$ which has all diagonal entries as 0. The* ***zero matrix*** *has all entries as 0.*

Property 3 (Properties of matrix operations)**.** *Given A, B, C are matrices of the same size and r & s are scalar:*

Commutative $A + B = B + A$

Associative $(A+B)+C = A+(B+C)$

Identity $A+0 = A$

Scalar Distributive $r(A+B) = rA+rB$

1. $(r+s)A = rA+sA$

Scalar associative $r(sA) = (rs)A$

Concept 7 (Matrix Multiplication)**.** *Given*

$$A_{m\times n} = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix}$$

and

$$B_{n\times p} = \begin{bmatrix} b_{11} & \cdots & b_{1l} & \cdots & b_{1p} \\ \vdots & & \vdots & & \vdots \\ b_{k1} & \cdots & b_{kl} & \cdots & b_{1l} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nl} & \cdots & b_{np} \end{bmatrix}$$

:

We have

$$AB = \begin{bmatrix} a_{11}b_{11}+\ldots+a_{1n}b_{n1} & \cdots & a_{11}b_{1l}+\ldots+a_{1n}b_{nl} & \cdots & a_{11}b_{1p}+\ldots+a_{1n}b_{np} \\ \vdots & & \vdots & & \vdots \\ a_{i1}b_{11}+\ldots+a_{in}b_{n1} & \cdots & a_{i1}b_{1l}+\ldots+a_{in}b_{nl} & \cdots & a_{i1}b_{1p}+\ldots+a_{in}b_{np} \\ \vdots & & \vdots & & \vdots \\ a_{m1}b_{11}+\ldots+a_{mn}b_{n1} & \cdots & a_{m1}b_{1l}+\ldots+a_{mn}b_{nl} & \cdots & a_{m1}b_{1p}+\ldots+a_{mn}b_{np} \end{bmatrix} \tag{3.2}$$

Note that we must have the number of columns of A and the number of rows of B be equal. Also, the (i, j) entry of AB comes out to be:

$$(AB)_{ij} = a_{i1}b_{1j}+a_{i2}b_{2j}+\ldots+a_{i(n-1)}b_{(n-1)j}+a_{in}b_{nj} \tag{3.3}$$

Property 4 (Properties of matrix multiplication). *Given $A_{m\times n}$, $B_{n\times p}$, $C_{p\times q}$:*

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

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

Distributive $(B+C)A = BA + CA$

- $r(AB) = (rA)B = A(rB)$

Identity $I_m A = A = AI_n$

Definition 20 (Transpose). *The* ***transpose*** *of $A_{m\times n}$ is some matrix $(A^T)_{n\times m}$ where each column is formed from the corresponding row of A. For example, the nth column of A^T is the nth row of A.*

Property 5 (Properties of transpositions). *We obtain the following properties for transpositions.*

1. $(A^T)^T = A$
2. $(A+B)^T = A^T + B^T$
3. $(rA)^T = rA^T$
4. $(AB\ldots C)^T = B^T A^T \ldots C^T$

3.2 Matrix Multiplicative Inverse

Definition 21 (Inverse). *We define the* ***inverse*** *of a square matrix A to be the matrix A^{-1} such that*

$$A^{-1}A = AA^{-1} = I \tag{3.4}$$

If such a matrix can be found, then A is called invertible. If not, then it is called singular.

Property 6 (Properties of Operations with inverses). *Given A is an invertible matrix:*

1. $(A^{-1})^{-1} = A$

2. $(AB \dots C)^{-1} = C^{-1}B^{-1} \dots A^{-1}$
3. $(A^T)^{-1} = (A^{-1})^T$

Definition 22 (Elementary Matrix)**.** *An* ***elementary matrix*** *is created by performing one row operation on an identity matrix. Multiplying a matrix A by it performs the same operation on A.*

Procedure 4 (Finding the inverse)**.** *Row reduce the augmented matrix* $\begin{bmatrix} A\ I \end{bmatrix}$ *until we obtain* $\begin{bmatrix} I\ A^{-1} \end{bmatrix}$*. If we can't row reduce A into I, then A is not invertible.*

3.3 More on Invertible Matrices

Theorem 5 (Invertible Matrix Theorem)**.** *Given* $A_{m\times n}$*, the following statements are equivelent:*

1. *A is invertible*
2. *There exists inverse* A^{-1} *such that* $A^{-1}A = AA^{-1} = I$*. Note that A and* A^{-1} *are both inverses of each other*
3. *A is row equivalent to* $I_{n\times n}$
4. *A has n pivot positions*
5. $A\boldsymbol{x} = \mathbf{0}$ *has only the trivial solution*
6. *The column vectors of A are linearly independent*
7. *The linear transformation* $\boldsymbol{x} \mapsto A\boldsymbol{x}$ *from* $\mathbb{R}^n \mapsto \mathbb{R}^m$ *is bijective (one-to-one and onto)*
8. *The equation* $A\boldsymbol{x} = \boldsymbol{b}$*has at least one solution for each* $\boldsymbol{b}$ *in* $\mathbb{R}^n$
9. *The column vectors of A span* $\mathbb{R}^n$
10. A^T *is also an invertible matrix*
11. *The columns of A form a basis of* $\mathbb{R}^n$
12. $Col(A) = \mathbb{R}^n$

13. $rank(A) = dim(Col(A) = n$
14. $Nul(A) = \{\mathbf{0}\}$
15. $dim(Nul(A)) = 0$
16. *The scalar 0 is not an eigenvalue of A*
17. $det(A) \neq 0$
18. $(ColA)^{\perp} = \{0\}$
19. $(NulA)^{\perp}\mathbb{R}^n$
20. $RowA = \mathbb{R}^n$
21. *A has n nonzero singular values.*

Theorem 6. *Linear Transformation T with standard matrix A is invertible if A is invertible. In that case, its inverse is the transformation given by standard matrix* A^{-1}

3.4 Subspaces

Definition 23 (Subspace)**.** *A **subspace** of* $\mathbb{R}^n$ *is any set H in* $\mathbb{R}^n$ *with three properties:*

1. *H contains the zero vector*
2. *For each* $\boldsymbol{u}$ *and* $\boldsymbol{v}$ *in H, the sum* $\approxeq + \succapprox$ *is in H*
3. *For each* $\boldsymbol{u}$ *in H and each scalar c, the vector* $c\boldsymbol{u}$ *is in H*

Definition 24 (Column Space)**.** *The **column space** of a matrix A is the set Col(A) of all linear combinations of the column vectors of A.*

$$ColA = Span\{a_1, \ldots, a_n\} \tag{3.5}$$

Definition 25 (Null Space)**.** *The **null space** of a matrix A is the set Nul(A) of all solutions of* $A\boldsymbol{x} = \mathbf{0}$

$$NulA = \{\boldsymbol{x} : \boldsymbol{x} \text{ is in } \mathbb{R}^n \text{ and } A\boldsymbol{x} = \mathbf{0}\} \tag{3.6}$$

Also note that $Nul(A_{m\times n})$ *is a subspace of* $\mathbb{R}^n$

Definition 26 (Basis)**.** *A **basis** for a subspace H is a linearly independent set in H that spans H*

*The set of vectors $e_1, \ldots, e_n$ is called the **standard basis** for $\mathbb{R}^n$ where the the jth entry in e_j is 1, all else being 0*

$$e_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

Note that the pivot columns of A form a basis for the column space of A. The free variable columns of A form a basis for the null space of A.

3.5 Dimension and Rank

Concept 8 (Coordinate Vectors)**.** *Given a basis $B = b_1, \ldots, b_p$ for subspace H, for each $\boldsymbol{x}$ in H, the coordinates of $\boldsymbol{x}$ relative to B are the weights $c_1, \ldots, c_p$ such that $\boldsymbol{x} = c_1\boldsymbol{b_1} + \ldots + c_p\boldsymbol{b_p}$.*

We call the B-coordinate vector:

$$[x]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_p \end{bmatrix}$$

Definition 27 (Dimension)**.** *The **dimension** of subspac, dim(H), is the number of vectors in the basis for H*

Definition 28 (Rank)**.** *The **rank** of a matrix A, rank(A), is dim(col(A))*

Theorem 7. *Rank Theorem For matrix A with n columns, we have $rank(A) + dim(Nul(A)) = n$*

Theorem 8 (Basis Theorem)**.** *Given H is a p-dimensional subspace of $\mathbb{R}^n$, any linearly dependent set of p elements in H is a basis for H. In addition, any set of p elements that span H are a basis as well.*

Chapter 4

Vector Spaces

4.1 Vector Spaces and Subspaces

Definition 29 (Vector Spaces)**.** *A **vector space** is a nonempty set V of objects, called vectors, for which we define two operations: vector addition (input: 2 vectors, output: 1 vector) and scalar multiplication (input: 1 vector 1 scalar, output: 1 vector). Given vectors $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$, and scalars c and d, the following ten axioms most hold:*

osed under vector addition $\boldsymbol{u}+\boldsymbol{v}$ is in V

1. $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$
2. $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$
3. *We can define a zero vector $\boldsymbol{0}$ s.t $\boldsymbol{u}+\mathbf{0}=\boldsymbol{u}$*
4. *We can define an additive inverse $-\boldsymbol{u}$ for each $\boldsymbol{u}$ in V s.t $\boldsymbol{u}+(-\boldsymbol{u})=\mathbf{0}$*

under scalar multiplication $c\boldsymbol{u}$ is in V

5. $c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$
6. $(c+d)\boldsymbol{u}=c\boldsymbol{u}+d\boldsymbol{u}$
7. $c(d\boldsymbol{u})=(cd)\boldsymbol{u}$
8. $1\boldsymbol{u}=\boldsymbol{u}$

Definition 30 (Subspace)**.** *A* ***subspace*** *of a vector space V is a subset H of V that obeys the following three properties:*

1. *The zero vector of V is in H*
2. *H is closed under vector addition*
3. *H is closed under scalar multiplication*

Note that if $v_1, \ldots, v_p$ are in a vector space V, then $Span v_1, \ldots, v_p$ is a subspace of V.

4.2 Null and Column Spaces

Concept 9 (Comparison of Nul A and Col A for $A_{m \times n}$)**.** *The following table may provide some useful insights on the Null Space and Column Space.*

Nul A	
Nul A is a subspace of $\mathbb{R}^n$	
Each vector $\boldsymbol{v}$ in Nul A satisfies $A\boldsymbol{v} = \mathbf{0}$.	*Ea*
Implicitly defined. You must take time to find vectors that satisfy a given condition.	*Explici*
$Nul A = \{\mathbf{0}\}$ iff the L.T $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is one-to-one.	

Sidenote that when applied to some linear transformation T, the null space of said transformation is also called the kernel of T.

4.3 Bases

Theorem 9 (Spanning Set Theorem)**.** *Let $S = \{v_1, \ldots, v_p\}$ be a set in V, and let $H = Span\{v_1, \ldots, v_p\}$, then the following statements hold true:*

1. *If some vector, v_k, is a L.C of other vectors in S, then the set formed by removing v_k from S still spans H.*
2. *If $H \neq \{\mathbf{0}\}$, then there exists some subset of S that is a basis for H.*

Observation 2 (Finding Bases for Col A and Null A)**.** *The pivot columns of matrix A form a basis of Col A. The columns corresponding to free variables in form a basis of Null A.*

4.4 Coordinate systems

Definition 31 (Coordinates based on Basis)**.** *Given basis for V, $\beta = \{\boldsymbol{b_1}, \boldsymbol{b_2}, \ldots, \boldsymbol{b_n}$, and vector $\boldsymbol{x}$ in V, the coordinates of x relative to the basis β are the weights $c_1, \ldots, c_2$ such that $\boldsymbol{x} = c_1\boldsymbol{b_1} + c_2\boldsymbol{b_2} + \ldots + c_n\boldsymbol{b_n}$. We also define the coordinate vector to be*

$$[\boldsymbol{x}]_\beta = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \tag{4.1}$$

and coordinate mapping to be $\boldsymbol{x} \mapsto [\boldsymbol{x}]_\beta$. Note that this is a one-to-one coordinate mapping from V to $\mathbb{R}^n$. This is referred to as an isomorphism.

Also, given the matrix with the set of bases as its column vectors,

$$P_\beta = [\boldsymbol{b}_1\ \boldsymbol{b}_n] \tag{4.2}$$

we have the vector equation

$$\boldsymbol{x} = P_\beta[\boldsymbol{x}]_\beta \tag{4.3}$$

4.5 Dimension

Definition 32 (Dimension)**.** *If vector space V is spanned by a finite set (where V is referred to as finite-dimensional), the **dimension** of V, dim V, is the number of vectors in the basis of V. If V cannot be spanned by a finite set, then V is infinite-dimensional.*

Note that any set in V with more than dim V vectors must be linearly dependent. Also, every basis in V must contain exactly dim V vectors.

Theorem 10. *Let H be a subspace of finite dimensional vector space V. We can build a basis for H from any linearly independent set in H by adding in more linearly independent vectors. We also know that*

$$dim H \leq dim V \tag{4.4}$$

Theorem 11 (Basis Theorem)**.** *Given p-dimensional vector space V, any set of p vectors that is linearly independent or spanning V is a basis for V.*

Definition 33 (Row Space)**.** *The **Row Space** of a matrix is the span of its row vectors.*

Definition 34 (Rank)**.** $\boldsymbol{rank} = dim(Col(A))$

Theorem 12 (Rank Theorem)**.** *Given* $A\boldsymbol{x} = \boldsymbol{b}Amn$

$$rankA + dimNullA = n \tag{4.5}$$

4.6 Change of Basis

Theorem 13. *Let* $\beta = [b_1, \ldots, b_n$ *and* $C = [c_1, \ldots, c_n$ *be bases of a vector space V. Then there exists a unique* $n/ctimesn$ *matrix* $\underset{C \leftarrow B}{P}$*, such that*

$$[\boldsymbol{x}]_C = \underset{C \leftarrow B}{P}[\boldsymbol{x}]_\beta \tag{4.6}$$

We know that

$$\underset{C \leftarrow B}{P} = \left[[\boldsymbol{b_1}]_C \; [\boldsymbol{b_2}]_C \ldots [\boldsymbol{b_n}]_C\right] \tag{4.7}$$

Also note that $(\underset{C \leftarrow B}{P})^{-1} = \underset{B \leftarrow C}{P}$

Chapter 5

Determinants

5.1 Intro to Determinants

Definition 35 (Recursive definition of Determinant)**.** *The* ***determinant*** *of matrix $A_{n\times n}$, where A_{ij} refers to the matrix A without row i and column j, is*

$$detA = a_{11}detA_{11} - a_{12}detA_{12} + \cdots + (-1)^{1+n}detA_{1n} = \sum_{j=1}^{n}(-1)^{1+j}a_{1j}detA_{1j} \tag{5.1}$$

If we define the (i, j)-cofactor of A (where i, j refers to the row, column) as the number

$$C_{ij} = (-1)^{i+j}detA_{ij} \tag{5.2}$$

we can also write the determinant of $A_{n\times n}$ as the cofactor expansion across the ith row or down the jth column.

$$detA = a_{i1}C_{i1} + a_{i2} + \cdots + a_{in}C_{in} \tag{5.3}$$

and

$$detA = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \tag{5.4}$$

5.2 Properties of Determinants

Property 7 (Row Operations). *For matrix $A_{n\times n}$, we can perform the following row operations to obtain matrix $B_{n\times n}$ s.t*

1. *Adding a multiple of one row to another row gives us $detB = detA$*
2. *Interchanging two rows gives us $detB = -detA$*
3. *Multiplying one row by a factor of k gives us $detB = k \cdot detA$*

Procedure 5 (Row Reduction for Determinant). *If we row reduce matrix $A_{n\times n}$ to row echelon form $\boldsymbol{U}$ using* r *and a total of* K *scaling factors, we know that*

$$detA = \begin{cases} (-1)^r \cdot K \cdot \prod(pivots\ in\ U) & \textit{if A is invertible (triangular)} \\ 0 & \textit{if A is not invertible} \end{cases} \tag{5.5}$$

Property 8 (More Properties). *Given matrix $A_{n\times n}$ and $B_{n\times n}$, we know that*

1. $detA^T = detA$
2. $detAB = (detA)(detB)$

Observation 3. *Interesting side note. If we frame det A as a function of the n column vectors of A, by holding the other columns fixed, we can define another function on some ith column x such that the function is linear.*

5.3 Cramer's Rule

Theorem 14 (Cramer's Rule). *Given matrix $A_{n\times n}$ and any $\boldsymbol{b}$ in $\mathbb{R}^n$, let $A_i(\boldsymbol{b}) = [\boldsymbol{a}_1 \ldots \boldsymbol{b} \ldots \boldsymbol{a}_n]$ where $\boldsymbol{b}$ is located in column i. Then, Cramer's Rule states that the solution $\boldsymbol{x}$ in $A\boldsymbol{x} = \boldsymbol{b}$ has entries given by*

$$x_i = \frac{detA_i(\boldsymbol{b})}{detA} \textit{for i in } \{1,\ 2,\ \ldots,\ n\} \tag{5.6}$$

Theorem 15 (Inverse Formula from Determinants)**.** *Given matrix $A_{n\times n}$, let $C_{n\times n}$ be its cofactor matrix and the (j, i)th cofactor be C_{ji}. Let the adjugate of A be*

$$adj A = C^T = \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{21} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix} \tag{5.7}$$

then we know the inverse of A is given by

$$A^{-1} = \frac{1}{det A} adj A \tag{5.8}$$

Theorem 16 (Geometric Interpretations)**.** *If A is a 2×2 matrix, then the area of the parallelogram determined by its column vectors is det A. If A is a 3×3 matrix, then the volume of the parallelepiped determined by its column vectors is det A.*

If we define a linear transformation, T, using A as the standard matrix:

1. If A is a 2×2 matrix and S is a parallelogram in $\mathbb{R}^2$ then
    - area of T(S) $= det A \cdot$ area of S
2. If A is a 3×3 matrix and S is a parallelepiped in $\mathbb{R}^3$ then
    - volume of T(S) $= det A \cdot$ volume of S

Chapter 6

Eigenvalues and Eigenvectors

6.1 Introduction

Definition 36 (Eigenvalue and Eignvector). *Given the equation $A\boldsymbol{x} = \lambda\boldsymbol{x}$ with matrix A, the* ***eigenvalue*** *is a scalar λ and the* ***eignvector*** *is a vector x that satisfies the equation.*

Theorem 17. *Given set S of eigenvectors $v_1, v_2, \ldots, v_n$ that each corresponds to distinct eigenvalues of matrix A, we know the set S is linearly independent.*

6.2 Characteristic Equation

Definition 37 (Characteristic Equation). *The eigenvalues of A must satisfy the below characteristic equation:*

$$det(A - \lambda I) = 0 \tag{6.1}$$

Procedure 6. *If matrix A is not invertible, the determinant is 0. When it is, we can calculate the determinant of A using row reduction into echelon form U. During row reduction, let r be the number of times we interchange rows, and let k be the cumulative product of the amount we scale rows, the determinant is equal to $\frac{(-1)^r}{k} \cdot (product\, of\, pivots\, of\, U)$*

Property 9. *Given $A_{n\times n}$ and $B_{n\times n}$*

1. *A is invertible iff $detA \neq 0$*

2. $det(AB) = (detA)(detB)$

3. $detA^T = detA$

4. *If A is in row echelon form, the determinant is the product of the diagonals*

5. *row replacement makes no difference, row interchange changes the sign, and row scaling scales by the same scalar factor on the determinant*

Definition 38 (Similarity). *Two matrices A and B are similar iff there is a matrix P s.t $A = P^{-1}BP$ We also know that if A and B are similar, then they have the same characteristic equation and eigenvalues.*

6.3 Diagonalization

Definition 39 (Diagonalization). *A matrix $A_{n\times n}$ is diagonalizable if it can be written in the form $A = P^{-1}DP$ where D is a diagonal matrix. By the Diagonalization theorem, this is also iff A has n linearly independent eigenvectors. In this case, the columns of P are these eigenvectors and the diagonal entries of D are the eigenvalues. Also note that we know A must be diagonalizable if it has n distinct eigenvalues.*

6.4 More on Linear Transformation

For a linear transformation from vector space V to vector space W with bases $B = b_1, \ldots, b_n$ and $C = c_1, \ldots, c_n$, the transformation T can be represented as

$$[T(x)]_C = M[x]_B \tag{6.2}$$

where we have M (called the B-matrix for T if V=W)

$$M = [[T(b_1)]_C \ [T(b_2)]_C \ \cdots \ [T(b_n)]_C] \tag{6.3}$$

Theorem 18 (Diagonal Matrix Representation). *Let $A = PDP^{-1}$ For linear transformation $x \mapsto Ax$ in $\mathbf{R}^n$, if B is the basis for $\mathbf{R}^n$ formed from the columns of P, then D is the B-matrix for the linear transformation.*

6.5 Complex Eigenvalues

Solving the Characteristic Equation may often give us complex eigenvalues. We know that these must come in pairs.

We can break up complex vectors into their real and imaginary parts. We can also define conjugates to be vectors where each entry is the complex entry of its corresponding entry in the original vector.

Chapter 7

Orthogonality and Least Squares

7.1 Introduction

Definition 40 (Inner Product)**.** *The inner product of two n-vectors u and v is the scalar given by $u^T v$. This is also equivalent to the dot product.*

Property 10. *Given u, v, and w are vectors in mathbbR^n and c is a scalar:*

1. $u \cdot v = v \cdot u$
2. $(u + v) \cdot w = u \cdot w + v \cdot w \cdot$
3. $(cu) \cdot v = c(u \cdot v) = u \cdot (cv)$
4. $u \cdot u \geq 0$ *and* $u \cdot u = 0$ *iff* $u = 0$

Definition 41 (Norm)**.** *The length/norm of v is the scalar $\|v\|$*

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2} \tag{7.1}$$

Definition 42 (Distance)**.** *Given u and v in $\mathbb{R}^n$, the distance in between is*

$$dist(u, v) = \|u - v\| \tag{7.2}$$

Definition 43 (Orthogonality)**.** *Let's start with the theorem that two vectors u and v are orthogonal iff $\|u + v\|^2 = \|u\|^2 + \|v\|^2$. From this, we define the vectors to be orthogonal if $u \cdot v = 0$*

Concept 10. *Given a vector space W in* $\mathbb{R}^n$*, we call the vector space* $W^\perp$ *the orthogonal complement of W if every vector in* $W^\perp$ *is orthogonal to every vector in W. Also, we note that* $W^\perp$ *is a subspace of* $\mathbb{R}^n$*.*

Theorem 19. *Given matrix* $A_{n\times n}$*, we know that*

$$(RowA)^\perp = NulA \qquad and \qquad (ColA)^\perp = NulA^T \tag{7.3}$$

Note: If u and v are both vectors in $\mathbb{R}^2 or \mathbb{R}^3$, we also know that $u \cdot v = \|u\| \, \|v\| cos\theta$ where θ is the angle between the two.

7.2 Orthogonal Sets

Definition 44 (Orthogonal Set)**.** *A set of vectors is orthogonal if each pair of vectors from the set is orthogonal.*

Definition 45 (Orthogonal Basis)**.** *An orthogonal basis for subspace W in* $\mathbb{R}^n$ *is a set that is both a basis for W and orthogonal. Note that all orthogonal sets are linearly independent and thus bases for the subspace they span.*

Theorem 20. *Given orthogonal basis* $u_w, \ldots, u_p$ *for subspace W, each y can be represented by the linear combination*

$$y = c_1u_1 + \cdots + c_pu_p \tag{7.4}$$

where the weight c_j *are given by the projection of y onto* c_j

$$c_j = \frac{y \cdot u_j}{u_j \cdot u_j} \tag{7.5}$$

Definition 46 (Orthonormal Set)**.** *We define the orthonormal set and basis as the orthogonal set and basis composed of only unit vectors.*

Theorem 21. *We take a closer look at the matrix* $U_{m\times n}$ *whose columns make up an orthonormal set. We see that its columns are orthonormal iff* $U^TU = I$*.*

Given x and y in $\mathbb{R}^n$*, we also have the following properties:*

1. $\|Ux\| = \|x\|$
2. $(Ux) \cdot (Uy) = x \cdot y$
3. $(Ux) \cdot (Uy) = 0$ *iff* $x \cdot y = 0$

Using these properties, we can see that the linear transformation given by $x \mapsto Ux$ *preserves length and orthogonality.*

7.3 Orthogonal Projections

Theorem 22 (Orthogonal Decomposition Theorem). *Let W be a subspace in $\mathbb{R}^n$. Each y in $\mathbb{R}^n$ can be written uniquely in the form*

$$y = \hat{y} + z \tag{7.6}$$

where $\hat{y}$ is in W and z is in $W^{\perp}$

Theorem 23 (Best Approximation Theorem). *$\hat{y}$, the orthogonal projection of y onto W, is the closest point on W to y (the distance from y to $\hat{y}$.*

Theorem 24. *Given matrix U whose columns are the orthonormal basis for subspace W of $\mathbb{R}^n$, then*

$$proj_W y = UU^T y \tag{7.7}$$

7.4 Gram-Schmidt Process

Procedure 7 (Gram-Schmidt Process). *Using the Gram-Schmidt process we can construct an orthogonal basis for a subspace W from an already given basis.*

Given a basis $x_1, \ldots, x_p$ for subspace W in $\mathbb{R}^n$, we generate new basis vectors by subtracting each already seen orthogonal vector from the old vectors.

$$\begin{aligned}
v_1 &= x_1 \\
v_2 &= x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 \\
v_3 &= x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 \\
&\vdots \\
v_p &= x_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \cdots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}
\end{aligned}$$

To make this into an orthonormal basis we just normalize each vector.

Theorem 25. *We can write matrix $A_{m\times n}$ out with the QR factorization as $A = QR$ where W is the $m \times n$ matrix whose columns are formed by the orthonormal basis for Col A and R is an $m \times n$ upper triangular matrix.*

7.5 The Least-Squares Problem

Theorem 26. *The least squares solution of $A\boldsymbol{x} = \boldsymbol{b}$is the vector $\hat{x}$ such that*

$$\|b - A\hat{x}\| \leq \|b - Ax\| \tag{7.8}$$

This set of least-squares solutions is the also the set of solutions of the normal equations $A^TAx = A^Tb$

Theorem 27. *Given matrix $A_{m\times n}$, the following statements are equivalent*

1. *The equation $A\boldsymbol{x} = \boldsymbol{b}$has a unique least-squares solution for each b in $\mathbb{R}^m$*
2. *The columns of A are linearly independent*
3. *The matrix A^TA is invertible*

Theorem 28. *We can also calculate the least-squares solution of $A\boldsymbol{x} = \boldsymbol{b}$using the QR factorization of $A = QR$.*

$$\hat{x} = R^{-1}Q^Tb \tag{7.9}$$

7.6 Inner Product Spaces

Definition 47 (Inner Product)**.** *An **inner product** on a vector space V is a function that, for each u and v on V, associates a real number ¡u, v¿ that satisfies the following axioms (where u, v, w are vectors and c is a scalar)*

1. $< u, v >=< v, u >$
2. $< u + v, w >=< u, w > + < v, w >$
3. $< cu, v >= c < u, v >$
4. $< u, u >\geq 0$ *and* $< u, u >= 0$ *iff* $u = 0$

A vector space with an inner product is called an inner product space.

Many of the same properties explored in earlier sections such as vector length and Gram-Schmidt still hold.

Theorem 29. *The Cauchy-Schwarz Inequality states that for all u, v in V,*

$$| < u, v > | \leq \|u\| \|v\|| \tag{7.10}$$

Theorem 30. *The Triangle Inequality states that for all u, v in V,*

$$\|u + v\| \leq \|u\| + \|v\|| \tag{7.11}$$

We can also commonly consider the inner product space for C[a, b] defined as the vector space of all continuous polynomial functions where the inner product for two vectors p and q in $\mathbb{P}_n$ is

$$< p, q >= p(t_0)q(t_0) + p(t_1)q_(t_1) + \cdots p(t_n)q(t_n) \tag{7.12}$$

Chapter 8

Symmetric Matrices and Quadratic Forms

Definition 48 (Symmetric Matrix)**.** *A* ***symmetric matrix*** *is a matrix A such that $A^T = A$*

Theorem 31. *If A is a symmetric matrix, then any two eigenvectors from different eigenspaces are orthogonal.*

Theorem 32. *A matrix $A_{n\times n}$ is orthogonally diagonalizable if there is an orthogonal matrix P such that*

$$A = PDP^T = PDP^{-1} \tag{8.1}$$

Also, A is orthogonally diagonalizable iff A is a symmetric matrix.

Theorem 33. *The Spectral Theorem for Symmetric Matrices states that matrix $A_{n\times n}$ has the following properties*

1. *A has n real eigenvalues (counting multiplicity)*
2. *The dimension for each eigenspace corresponding to λ is the multiplicity of λ*
3. *The eigenspaces are mutually orthogonal*
4. *A is orthogonally diagonalizable*

Theorem 34. *With the spectral decomposition, we can break up a diagonalizable matrix $A = PDP^{-1}$, where the columns of P are orthonormal eigenvectors $u_1, u_2, \ldots, u_n$ and*

$$A = \lambda_1 u_1 u_q^T + \lambda_2 u_2 u_2^T + \ldots + \lambda_n u_n u_n^T \tag{8.2}$$

8.1 Singular Value Decomposition

Theorem 35. *Let the r singular values of matrix A be the square roots of the eigenvalues of A^TA. Then $\{Av_1, Av_2, \ldots, Av_r\}$, where v_i are the eigenvectors of A^TA, is the orthogonal basis of $ColA$ and $rankA = r$*

Theorem 36 (Singular Value Decomposition)**.** *Let A be any matrix $A_{m\times n}$ with rank r. Then, we can decompose A into $\Sigma_{m\times n}$ for which the first r diagonal entries are the singular values of A in nonincreasing order, and into orthogonal matrices $U_{m\times m}$ (A times corresponding eigenvectors) and $V_{n\times n}$ (corresponding eigenvectors) such that*

$$A = U\Sigma V^T \tag{8.3}$$

Procedure 8.

1. *Find orthogonal diagonalization of A^TA*
2. *Construct Σ and U and V from eigenvalues/eigenvectors*

Chapter 9

Linear Second Order Differential Equations

9.1 Homogeneous Linear Equations

Theorem 37. *We start by defining linear second order differential equations in the form*

$$ay'' + by' + cy = 0 \tag{9.1}$$

There exists a unique solution to the above equation. Because we are looking for functions that are similar to their first and second order derivatives, we consider solutions to be expressions of the form $y = ce^{rt}$, where r_1 and r_2 are of the form $\frac{-b \pm \sqrt{b^2-4ac}}{2a}$. We can find c_1 and c_2 by considering initial values.

If there is a repeated root r, then c_1e^{rt} and c_2te^{rt} are solutions.

Definition 49 (Linear Independence)**.** *A pair of functions is linearly independent on interval I iff neither of them is a constant multiple of the other on all of I. Otherwise, they are linearly dependent.*

If y_1 and y_2 are solutions to a linear second order differential equation and have a solution τ on the interval $(-\infty, -\infty)$ for the equation

$$y_1(\tau)y_2'(\tau) - y_1'(\tau)y_2(\tau) \tag{9.2}$$

Theorem 38. *Now let's consider harmonic equations - a special case of differential equations with auxiliary equations with complex roots.*

Before we begin, note that by Euler's Formula:

$$e^{i\theta} = cos\theta + isin\theta \tag{9.3}$$

Thus, solutions to the auxiliary equation $\alpha \pm i\beta$ gives a solutions for the differential equation $c_1 e^{\alpha} cos(\beta)$ and $c_2 e^{\alpha} sin(\beta)$

Procedure 9 (Method of Undetermined Coefficients)**.** *With differential equations of the form*

$$ay'' + by' + cy = P_m(t)e^{rt} \tag{9.4}$$

We have solutions of the form

$$t^s(A_m t^m + \cdots + A_1 t + A_0)e^{rt} \tag{9.5}$$

where

1. *$s = 0$ if r is not a root of the auxiliary equation*
2. *$s = 1$ if r is a single root of the auxiliary equation*
3. *$s = 2$ if r is a double root of the auxiliary equation*

With differential equations of the form

$$ay'' + by' + cy = Ct^m e^{\alpha t} cos(\beta t) \tag{9.6}$$

or

$$ay'' + by' + cy = Ct^m e^{\alpha t} sin(\beta t) \tag{9.7}$$

We have solutions of the form

$$t^s(A_m t^m + \cdots A_1 t + A_0)e^{\alpha t} cos(\beta t) \tag{9.8}$$

or

$$t^s(A_m t^m + \cdots A_1 t + A_0)e^{\alpha t} sin(\beta t) \tag{9.9}$$

where

1. *s=0 if $\alpha + i\beta$ is not a root of the auxiliary equation*
2. *s=1 if $\alpha + i\beta$ is a root of the auxiliary equation*

Theorem 39. *By the Superposition Principle, if y_1 is a solution to the differential equation with left side $f_1(t)$ and y_2 is a solution to the differential equation with left side $f_2(t)$, then $k_1y_1 + k_2y_2$ is a solution to the differential equation with left side $k_1f_1(t) + k_2f_2(t)$*

Theorem 40. *Given nonhomogenous differential equation $ay'' + by' + cy = f(t)$ with some solution $y_p(t)$ and corresponding homogenous differential equation $ay'' + by' + cy = 0$ with linearly independent solutions $y_1(t)$ and $y_2(t)$, there exists a unique solution $y_p(t) + c_1y_1(t) + c_2y_2(t)$*

9.2 Linear Systems and Differential Equations

Theorem 41. *The first fundamental idea is to consider functions as vectors with each element as a term, and derivatives as linear transformations with associated matrices.*

Definition 50 (Normal Form)**.** *Thus, we can represent a system of n linear differential equations in its normal form:*

$$x'(t) = A(t)x(t) + f(t) \tag{9.10}$$

where $x(t) = col(x_1(t), \ldots, x_n(t))$, $f(t) = col(f_1(t), \ldots, f_n(t))$, $A(t) = [a_{ij}(t)]$

Definition 51 (Wronskian)**.** *Linear dependence/independence of vector functions follows much the same principle as before.*

The Wronskian of n vector functions $x_1 = col(x_{1,1}, \ldots, x_{n,1}, \ldots, x_n = col(x_{1,n}, \ldots, x_{n,n}, \ldots,$ is the function:

$$W[x_1, \ldots, x_n](t) = \begin{vmatrix} x_{1,1}(t) & x_{1,2}(t) & \ldots & x_{1,n}(t) \\ x_{2,1}(t) & x_{2,2}(t) & \ldots & x_{2,n}(t) \\ \vdots & \ldots & & \ldots \\ x_{n,1}(t) & x_{n,2}(t) & \ldots & x_{n,n}(t) \end{vmatrix} \tag{9.11}$$

Procedure 10. *Procedure to find the general solution to the homogenous system $x' = Ax$.*

1. *Find fundamental, linearly independent solution set $x_1, \ldots, x_n$*

2. *Form the linear combination*

$$\mathbf{x} = c_1x_1 + \ldots + c_nx_n \tag{9.12}$$

Procedure to find the general solution to the nonhomogenous system $x' = Ax + f$

1. *Find one particular solution x_p*

2. *General solution is sum of x_p and general solution for corresponding homogeneous system from before $x = x_p + c_1 x_1 + \ldots + c_n x_n$*

The fundamental linearly independent solution set for matrix $A_{n\times n}$ is:

$$e^{r_1 t}u_1, e^{r_2 t}u_2, \ldots, e^{r_n t}u_n \tag{9.13}$$

where r_i is the eigenvalue corresponding to eigenvector u_i. Note also that the eigenvectors associated with distinct eigenvalues must be linearly independent, thereby making a fundamental solution set (if there are n distinct eigenvalues).

Also note that real symmetric matrices must satisfy this property.

Theorem 42. *Now, we consider complex eigenvalues. If real matrix A has complex conjugate eigenvalues $\alpha \pm i\beta$ and corresponding eigenvectors $a \pm ib$, then two LI vector solutions to $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$are $e^{\alpha t}cos(\beta t)a - e^{\alpha t}sin(\beta t)b$ and $e^{\alpha t}sin(\beta t)a - e^{\alpha t}cos(\beta t)b$*

9.3 Fourier Series

Definition 52 (Fourier Series)**.** *Given f is a piecewise continuous function on the interval [-L,L]. The Fourier Series of f is the trigonometric series:*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\{a_n cos(\frac{n\pi x}{L}) + b_n sin(\frac{n\pi x}{L})\} \tag{9.14}$$

Where a_n b_n are given by the formulas

$$a_n \sim \frac{1}{L}\int_{-L}^{L} f(x)cos(\frac{n\pi x}{L})dx,\ n = 0, 1, 2, \ldots \tag{9.15}$$

$$b_n \sim \frac{1}{L}\int_{-L}^{L} f(x)sin(\frac{n\pi x}{L})dx,\ n = 0, 1, 2, \ldots \tag{9.16}$$

Definition 53 (Orthogonal Expansions)**.** *A set of functions* $\{f_n x)\}^{\{} infty_{n=1}$ *is an orthogonal system with respect to the weight function* $w(x)$ *on the interval* $[a, b]$ *if*

$$\sum_a^b f_m(x) f_n(x) w(x) dx = 0, \ wheneverm \neq n \tag{9.17}$$

It is an orthonormal set if each element is also normal

$$\|f\| = \sqrt{\int_a^b f^2(x) w(x) dx} = 1 \tag{9.18}$$

Now, we can define a generalized fourier series or orthogonal expansion for such an orthonormal set:

$$f(x) = c_1 f_1(x) + c_2 f_2(x) + \dots \tag{9.19}$$

We find each constant is

$$c_m = \frac{\int_c^b f(x) f_m(x) w(x) dx}{\|f_m\|^2}, \ n = 1, 2, 3, \dots \tag{9.20}$$

Theorem 43. *If f and f' are piecewise continuous on interval* $[-L, L]$*, then for any x on* $(-L, L)$*, the fourier series converges to* $\frac{1}{2}[f(x^+) + f(x^-)]$*. When* $x = \pm L$*, the series converges to* $\frac{1}{2}[f(-L^+) + f(L^-)]$

Differentiation and Integration of a Fourier Series can all occur term-wise.

Definition 54 (Fourier Sine and Cosine Series)**.** *Given f(x) is a piecewise continuous function on the interval [0, L]. The Fourier Cosine Series on that same interval is*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n cos(\frac{n\pi x}{L} \tag{9.21}$$

$$a_n = \frac{2}{L} \int_0^L f(x) cos(\frac{n\pi x}{L} dx), l : n = 0, 1, 2, \dots \tag{9.22}$$

The Fourier Sine Series on that same interval is

$$\sum_{n=1}^{\infty} b_n sin(\frac{n\pi x}{L} \tag{9.23}$$

$$b_n = \frac{2}{L} \int_0^L f(x) sin(\frac{n\pi x}{L} dx), l : n = 0, 1, 2, \dots \tag{9.24}$$