SUMMARY NOTES – HARVARD MATH 21B, LINEAR ALGEBRA & DIFFERENTIAL EQUATIONS

Contents

1. Linear Equations

1.1. Introduction to Linear Systems.

Definition 1 – Linear Combination
A **linear combination** of $x_1, \ldots, x_n$ has the form

$$a_1 x_1 + a_2 x_2 + a_3 x_3 + \cdots + a_n x_n$$

where the numbers $a_1, \ldots, a_n \in \mathbb{R}$ are the combination's **coefficients**[a].

---

[a]Sometimes we replace $\mathbb{R}$ with another field like $\mathbb{Q}$ or $\mathbb{C}$.

Definition 2 – Linear Equation
A **linear equation** in the variables $x_1, \ldots, x_n$ has the form $a_1 x_1 + a_2 x_2 + a_3 x_3 + \cdots + a_n x_n = d$ where $d \in \mathbb{R}$ is the **constant**.
An n-tuple $(s_1, s_2, \ldots, s_n) \in \mathbb{R}^n$ is a **solution** of, or **satisfies**, that equation if substituting the numbers $s_1, \ldots, s_n$ for the variables gives a true statement: $a_1 s_1 + a_2 s_2 + \cdots + a_n s_n = d$.
A **system of linear equations**

$$a_{1,1} x_1 + a_{1,2} x_2 + \cdots + a_{1,n} x_n = d_1$$
$$a_{2,1} x_1 + a_{2,2} x_2 + \cdots + a_{2,n} x_n = d_2$$
$$\vdots$$
$$a_{m,1} x_1 + a_{m,2} x_2 + \cdots + a_{m,n} x_n = d_m$$

has the solution $(s_1, s_2, \ldots, s_n)$ if that n-tuple is a solution of all of the equations.

Theorem 3 – Gauss's Method a.k.a Gauss-Jordan or Row Reduction
If a linear system is changed to another by one of these operations
   (1) an equation is swapped with another
   (2) an equation has both sides multiplied by a nonzero constant
   (3) an equation is replaced by the sum of itself and a multiple of another
then the two systems have the same set of solutions.

Definition 4 – Reduced row-echelon form
A matrix is said to be in **reduced row-echelon form (rref)** if it satisfies all of the following conditions
   (1) If a row has nonzero entries, then the first nonzero entry is a 1, called **the leading 1 (or pivot)** in this row.
   (2) If a column contains a leading 1, then all the other entries in that column are 0 .
   (3) If a row contains a leading 1, then each row above it contains a leading 1 further to the left.
The last condition implies that rows of 0's, if any, appear at the bottom of the matrix.

Definition 5 – Elementary Reduction Operations
The three operations from Theorem 3 are the **elementary reduction operations**, or **row operations**, or **Gaussian operations**. They are:
   • **swapping**
   • **multiplying by a scalar** (or **rescaling**)
   • **row combination**

Definition 6 – Echelon Form
A system is in **echelon form** if each leading variable (first nonzero variable in each equation) is to the right of the leading variable in the row above it, with any all-zero rows at the bottom.

Definition 7 – Free Variables
In an echelon form linear system the variables that are not leading are **free**.

Corollary 8 – Solution Set Types
Solution sets of linear systems are either
- *empty,*
- *have one element, or*
- *have infinitely many elements.*

Theorem 9 – General Solution Form
Any linear system's solution set has the form $\{p + c_1\beta_1 + \cdots + c_k\beta_k \mid c_1, \ldots, c_k \in \mathbb{R}\}$ where p is any particular solution and k equals the number of free variables.

Definition 10 – Matrix
An $m \times n$ **matrix** is a rectangular array of numbers with m rows and n columns. Each number is an **entry**.

Definition 11 – Vector
A **column vector** is a matrix with a single column. A **row vector** has a single row. Entries are called **components**.

Definition 12 – Vector Sum
The **vector sum** of $\vec{u}$ and $\vec{v}$ is the vector of the sums (component-wise addition).

Definition 13 – Scalar Multiplication
The **scalar multiplication** of real number r and vector $\vec{v}$ is the vector of the multiples.

Definition 14 – Homogeneous Equation
A **homogeneous equation** is a linear equation with constant of zero: $a_1 x_1 + \cdots + a_n x_n = 0$.

Definition 15 – Nonsingular Matrix
A square matrix is **nonsingular** if it is the matrix of coefficients of a homogeneous system with a unique solution. Otherwise it is **singular**.

Definition 16 – Linear System
A **linear system** in variables $x_1, \ldots, x_n$ is a collection of equations $a_{i1}x_1 + \cdots + a_{in}x_n = b_i$. Its solution set is an affine subspace of $\mathbb{R}^n$ (possibly empty).

Definition 17 – Matrix Form $Ax = b$
The **coefficient matrix** $A \in \mathbb{R}^{m \times n}$, **unknown vector** $x \in \mathbb{R}^n$, and **right-hand side** $b \in \mathbb{R}^m$ define
the system $Ax = b$.

Theorem 18 – Existence and Uniqueness
*The system $Ax = b$ has a solution iff b lies in the column space of A (the image). If a solution exists, it
is unique iff $\ker(A) = \{0\}$, equivalently $\mathrm{rank}(A) = n$.*

1.2. **Matrices, Vectors, and Gauss–Jordan Elimination.**

Definition 19 – Vectors and vector spaces
A matrix with only one column is called a **column vector**, or simply a **vector**. The entries of a vector
are called its **components**. The set of all column vectors with n components is denoted by $\mathbb{R}^n$; we will
refer to $\mathbb{R}^n$ as a vector space.
A matrix with only one row is called a **row vector**. In this text, the term vector refers to column
vectors, unless otherwise stated.

Definition 20 – Elementary Row Operations
(1) Swap two rows.
(2) Multiply a row by a nonzero scalar.
(3) Replace a row by itself plus a multiple of another row. These preserve the solution set of $Ax = b$.

Theorem 21 – Reduced Row-Echelon Form (RREF)
*Every matrix is row-equivalent to a unique RREF. **Pivot** columns correspond to **leading variables**;
non-pivot columns correspond to **free variables** that parametrize the solution set.*

1.3. **On the Solutions of Linear Systems; Matrix Algebra.**

Theorem 22 – Number of solutions of a linear system
*A system of equations is said to be consistent if there is at least one solution; it is inconsistent if there
are no solutions.*
*A linear system is inconsistent if (and only if) the reduced row-echelon form of its augmented matrix
contains the row $\begin{bmatrix} 0 & 0 & \cdots & 0 & | & 1 \end{bmatrix}$, representing the equation $0 = 1$.*
If a linear system is consistent, then it has either
- infinitely many solutions (if there is at least one free variable), or
- exactly one solution (if all the variables are leading).

Definition 23 – Rank of a Matrix (preliminary definition)
The rank of a matrix A is the number of leading 1 's in $\mathrm{rref}(A)$, denoted $\mathrm{rank}(A)$.

Theorem 24 – Number of equations vs. number of unknowns
*If a linear system has exactly one solution, then there must be at least as many equations as there are
variables ($m \le n$ with the notation from Definition 2).*
Equivalently, we can formulate the contrapositive:
*A linear system with fewer equations than unknowns $(n < m)$ has either no solutions or infinitely many
solutions.*

Theorem 25 – The product $A\vec{x}$ in terms of the columns of A

If the column vectors of an $n \times m$ matrix A are $\vec{v}_1, \ldots, \vec{v}_m$ and $\vec{x}$ is a vector in $\mathbb{R}^m$ with components $x_1, \ldots, x_m$, then

$$A\vec{x} = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_m \\ | & & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix} = x_1 \vec{v}_1 + \cdots + x_m \vec{v}_m$$

Definition 26 – Linear combinations

A vector $\vec{b}$ in $\mathbb{R}^n$ is called a linear combination of the vectors $\vec{v}_1, \ldots, \vec{v}_m$ in $\mathbb{R}^n$ if there exist scalars $x_1, \ldots, x_m$ such that

$$\vec{b} = x_1 \vec{v}_1 + \cdots + x_m \vec{v}_m$$

Definition 27 – Matrix Operations

Matrix addition, scalar multiplication, and multiplication $(AB)_{ij} = \sum_k A_{ik} B_{kj}$; the identity I_n satisfies $I_n x = x$.

Theorem 28 – Rank–Nullity

For $A \in \mathbb{R}^{m \times n}$, $\operatorname{rank}(A) + \dim \ker(A) = n$.

2. LINEAR TRANSFORMATIONS

2.1. **Introduction to Linear Transformations and Their Inverses.**

Definition 29 – Linear Transformation

A map $T : V \to W$ is **linear** if $T(u + v) = T(u) + T(v)$ and $T(\alpha v) = \alpha T(v)$. It is invertible (has a linear inverse) iff it is bijective.

Theorem 30 – Invertibility Criteria on $\mathbb{R}^n$

For $T : \mathbb{R}^n \to \mathbb{R}^n$ with matrix A, the following are equivalent:
(i) T invertible;
(ii) $\det A \neq 0$;
(iii) $\ker(A) = \{0\}$;
(iv) $\operatorname{Im}(A) = \mathbb{R}^n$.

Theorem 31 – The columns of the matrix of a linear transformation

Consider a linear transformation T from $\mathbb{R}^m$ to $\mathbb{R}^n$. Then, the matrix of T is

$$A = \begin{bmatrix} | & | & & | \\ T(\vec{e}_1) & T(\vec{e}_2) & \cdots & T(\vec{e}_m) \\ | & | & & | \end{bmatrix}, \quad where \quad \vec{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i\ th$$

2.2. **Linear Transformations in Geometry.**

Proposition 32 – Geometric Examples

Rotations, reflections, orthogonal projections, and shears on $\mathbb{R}^2$ or $\mathbb{R}^3$ are linear and each is represented by a matrix in the standard basis.

Definition 33 – Orthogonal Projections

Consider a line L in the coordinate plane, running through the origin. Any vector $\vec{x}$ in $\mathbb{R}^2$ can be written uniquely as

$$\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp},$$

where $\vec{x}^{\parallel}$ is parallel to line L, and $\vec{x}^{\perp}$ is perpendicular to L. The transformation $T(\vec{x}) = \vec{x}^{\parallel}$ from $\mathbb{R}^2$ to $\mathbb{R}^2$ is called the orthogonal projection of $\vec{x}$ onto L, often denoted by $\mathrm{proj}_L(\vec{x})$. If $\vec{w}$ is a nonzero vector parallel to L, then
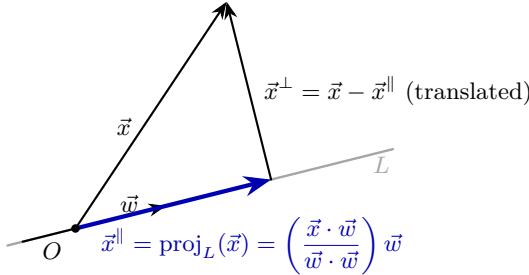
$$\mathrm{proj}_L(\vec{x}) = \left(\frac{\vec{x} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \right) \vec{w}.$$

In particular, if $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is a unit vector parallel to L, then

$$\mathrm{proj}_L(\vec{x}) = (\vec{x} \cdot \vec{u})\vec{u}.$$

The transformation $T(\vec{x}) = \mathrm{proj}_L(\vec{x})$ is linear, with matrix

$$P = \frac{1}{w_1^2 + w_2^2} \begin{bmatrix} w_1^2 & w_1 w_2 \\ w_1 w_2 & w_2^2 \end{bmatrix} = \begin{bmatrix} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{bmatrix}.$$

$\vec{x}^{\perp} = \vec{x} - \vec{x}^{\parallel}$ (translated)

L

$\vec{x}$

$\vec{w}$

O $\vec{x}^{\parallel} = \mathrm{proj}_L(\vec{x}) = \left(\dfrac{\vec{x} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \right) \vec{w}$

Definition 34 – Reflections
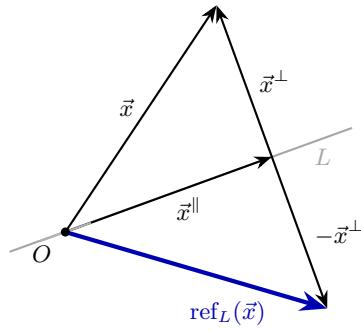
Consider a line L in the coordinate plane, running through the origin, and let $\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$ be a vector in $\mathbb{R}^2$. The linear transformation $T(\vec{x}) = \vec{x}^{\parallel} - \vec{x}^{\perp}$ is called the reflection of $\vec{x}$ about L, often denoted by $\mathrm{ref}_L(\vec{x})$:

$$\mathrm{ref}_L(\vec{x}) = \vec{x}^{\parallel} - \vec{x}^{\perp}$$

We have a formula relating $\mathrm{ref}_L(\vec{x})$ to $\mathrm{proj}_L(\vec{x})$:

$$\mathrm{ref}_L(\vec{x}) = 2\,\mathrm{proj}_L(\vec{x}) - \vec{x} = 2(\vec{x} \cdot \vec{u})\vec{u} - \vec{x} \qquad \text{where } \vec{u} \text{ is a unit vector on } L.$$

The matrix of T is of the form $\begin{bmatrix} a & b \\ b & -a \end{bmatrix}$, where $a^2 + b^2 = 1$. Conversely, any matrix of this form represents a reflection about a line.

2.3. Matrix Products.

Definition 35 – Matrix multiplication
a. Let B be an $n \times p$ matrix and A a $q \times m$ matrix. The product BA is defined if (and only if) $p = q$.
b. If B is an $n \times p$ matrix and A a $p \times m$ matrix, then the product BA is defined as the matrix of the linear transformation $T(\vec{x}) = B(A\vec{x})$. This means that $T(\vec{x}) = B(A\vec{x}) = (BA)\vec{x}$, for all $\vec{x}$ in the vector space $\mathbb{R}^m$. The product BA is an $n \times m$ matrix.

Theorem 36 – Composition $\leftrightarrow$ Product
If $S : U \to V$ and $T : V \to W$ have matrices $[S]$ and $[T]$ in chosen bases, then $[T \circ S] = [T][S]$.

Theorem 37 – The columns of the matrix product
Let B be an $n \times p$ matrix and A a $p \times m$ matrix with columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$. Then, the product BA is

$$BA = B \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ B\vec{v}_1 & B\vec{v}_2 & \cdots & B\vec{v}_m \\ | & | & & | \end{bmatrix}.$$

To find BA, we can multiply B by the columns of A and combine the resulting vectors.

Theorem 38 – The entries of the matrix product
Let B be an $n \times p$ matrix and A a $p \times m$ matrix. The ij th entry of BA is the dot product of the i th row of B with the j th column of A.

$$BA = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{ip} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pj} & \cdots & a_{pm} \end{bmatrix}$$

is the $n \times m$ matrix whose ij th entry is

$$b_{i1}a_{1j} + b_{i2}a_{2j} + \cdots + b_{ip}a_{pj} = \sum_{k=1}^{p} b_{ik}a_{kj}.$$

Proposition 39 – Matrix Properties
Multiplying with the identity matrix
For an $n \times m$ matrix A,
$$AI_m = I_n A = A$$
Matrix multiplication is associative
$$(AB)C = A(BC)$$
Distributive property for matrices
If A and B are $n \times p$ matrices, and C and D are $p \times m$ matrices, then
$$A(C+D) = AC + AD, \quad \text{and}$$
$$(A+B)C = AC + BC$$
If A is an $n \times p$ matrix, B is a $p \times m$ matrix, and k is a scalar, then
$$(kA)B = A(kB) = k(AB).$$

Warning 40 – Not as simple as numbers
If A, B, C are $n \times n$ matrices, in general
- *$AB \neq BA$ (when they are equal we say that A and B commute)*
- *$(A+B)^2 = A^2 + AB + BA + B^2 \neq A^2 + 2AB + B^2$*
- *We could have $A \neq 0, B \neq 0$ but $AB = 0$*
- *We could have $AB = 0$ and $BA \neq 0$*
- *$AC = AB$ does not imply that $B = C$ (but it does when A is non singular/invertible).*

2.4. **The Inverse of a Linear Transformation.**

Definition 41 – Invertible matrices
A square matrix A is said to be invertible if the linear transformation $\vec{y} = T(\vec{x}) = A\vec{x}$ is invertible. In this case, the matrix of T^{-1} is denoted by A^{-1}. If the linear transformation $\vec{y} = T(\vec{x}) = A\vec{x}$ is invertible, then its inverse is $\vec{x} = T^{-1}(\vec{y}) = A^{-1}\vec{y}$.

Theorem 42 – Invertibility
An $n \times n$ matrix A is invertible if and only if
$$\mathrm{rref}(A) = I_n$$
or, equivalently, if and only if
$$\mathrm{rank}(A) = n.$$

Technics 43 – Computing A^{-1}
If A is invertible, then A^{-1} is obtained by Gauss–Jordan on $\begin{bmatrix} A \,|\, I \end{bmatrix}$. Moreover, $[T^{-1}] = [T]^{-1}$ for invertible T.

Theorem 44 – Invertibility and linear systems
Let A be an $n \times n$ matrix.
(1) Consider a vector $\vec{b}$ in $\mathbb{R}^n$. If A is invertible, then the system $A\vec{x} = \vec{b}$ has the unique solution $\vec{x} = A^{-1}\vec{b}$. If A is noninvertible, then the system $A\vec{x} = \vec{b}$ has infinitely many solutions or none.
(2) Consider the special case when $\vec{b} = \vec{0}$. The system $A\vec{x} = \vec{0}$ has $\vec{x} = \vec{0}$ as a solution. If A is invertible, then this is the only solution. If A is noninvertible, then the system $A\vec{x} = \vec{0}$ has infinitely many solutions.

Warning 45 – The inverse of a product of matrices
If A and B are invertible $n \times n$ matrices, then BA is invertible as well, and
$$(BA)^{-1} = A^{-1}B^{-1}$$
Pay attention to the order of the matrices. (Order matters!)

Theorem 46 – A criterion for invertibility
Let A and B be two $n \times n$ matrices such that
$$BA = I_n$$
Then
a. A and B are both invertible,
b. $A^{-1} = B$ and $B^{-1} = A$, and
c. $AB = I_n$.

Corollary 47 – Inverse of a 2×2 matrix
The 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if (and only if) $ad - bc \neq 0$. Quantity $ad - bc$ is called the
determinant of A, written $\det(A)$
$$\det(A) = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc.$$
If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible, then
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{\det(A)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Theorem 48 – Various characterizations of invertible matrices
*For an $n \times n$ matrix A, the following statements are equivalent; that is, for a given A, they are either
all true or all false*
    i. A is invertible.
    ii. The linear system $A\vec{x} = \vec{b}$ has a unique solution $\vec{x}$, for all $\vec{b}$ in $\mathbb{R}^n$.
    iii. $\mathrm{rref}(A) = I_n$.
    iv. $\mathrm{rank}(A) = n$.
    v. $\mathrm{im}(A) = \mathbb{R}^n$.
    vi. $\ker(A) = \{\vec{0}\}$.
    vii. $\det A \neq 0$.

3. SUBSPACES OF $\mathbb{R}^n$ AND THEIR DIMENSIONS

3.1. **Image and Kernel of a Linear Transformation.**

Definition 49 – Span
Consider the vectors $\vec{v}_1, \ldots, \vec{v}_m$ in $\mathbb{R}^n$. The set of all linear combinations $c_1\vec{v}_1 + \cdots + c_m\vec{v}_m$ of the vectors
$\vec{v}_1, \ldots, \vec{v}_m$ is called their span:
$$\mathrm{span}\,(\vec{v}_1, \ldots, \vec{v}_m) = \{c_1\vec{v}_1 + \cdots + c_m\vec{v}_m : c_1, \ldots, c_m \text{ in } \mathbb{R}\}$$

Definition 50 – Kernel and Image
For linear $T : V \to W$, $\ker(T) = \{v \in V : Tv = 0\}$ and $\mathrm{im}(T) = \{Tv : v \in V\}$; both are subspaces.

Theorem 51 – Image of a linear transformation
The image of a linear transformation $T(\vec{x}) = A\vec{x}$ is the span of the column vectors of A.

Theorem 52 – Rank–Nullity for Linear Maps
If $\dim V < \infty$, then $\dim \ker(T) + \dim \operatorname{im}(T) = \dim V$.

3.2. Subspaces of $\mathbb{R}^n$; Bases and Linear Independence.

Definition 53 – Subspaces of $\mathbb{R}^n$
A subset W of the vector space $\mathbb{R}^n$ is called a (linear) subspace of $\mathbb{R}^n$ if it has the following three properties:
  a. W contains the zero vector in $\mathbb{R}^n$.
  b. W is closed under addition: If $\vec{w}_1$ and $\vec{w}_2$ are both in W, then so is $\vec{w}_1 + \vec{w}_2$.
  c. W is closed under scalar multiplication: If $\vec{w}$ is in W and k is an arbitrary scalar, then $k\vec{w}$ is in W.

Theorem 54 – Image and kernel are subspaces
If $T(\vec{x}) = A\vec{x}$ is a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$, then
  - *$\ker(T) = \ker(A)$ is a subspace of $\mathbb{R}^m$*
  - *$\operatorname{im}(T) = \operatorname{im}(A)$ is a subspace of $\mathbb{R}^n$.*

Definition 55 – Redundant vectors; linear independence; basis
Consider vectors $\vec{v}_1, \ldots, \vec{v}_m$ in $\mathbb{R}^n$.
  a. We say that a vector $\vec{v}_i$ in the list $\vec{v}_1, \ldots, \vec{v}_m$ is **redundant** if $\vec{v}_i$ is a linear combination of the preceding vectors $\vec{v}_1, \ldots, \vec{v}_{i-1}$.
  b. The vectors $\vec{v}_1, \ldots, \vec{v}_m$ are called **linearly independent** if none of them is redundant. Otherwise, the vectors are called **linearly dependent** (meaning that at least one of them is redundant).
  c. We say that the vectors $\vec{v}_1, \ldots, \vec{v}_m$ in a subspace V of $\mathbb{R}^n$ **form a basis** of V if they span V and are linearly independent.

Theorem 56 – Dimension
*Any two bases of a finite-dimensional vector space have the same cardinality, called the **dimension**.*

3.3. The Dimension of a Subspace of $\mathbb{R}^n$.

Definition 57 – Row/Column/Null Spaces
For $A \in \mathbb{R}^{m \times n}$, the row space (in $\mathbb{R}^n$), column space (in $\mathbb{R}^m$), and null space are subspaces. Their dimensions satisfy $\operatorname{rank}(A) = \dim(\text{row}) = \dim(\text{col})$.

3.4. Coordinates.

Definition 58 – Coordinates and Change of Basis
Given a basis $\mathcal{B} = \{v_1, \ldots, v_n\}$ of V, each $v \in V$ has unique coordinates $[v]_\mathcal{B} \in \mathbb{R}^n$ with $v = \sum_i ([v]_\mathcal{B})_i v_i$.

If P changes coordinates from $\mathcal{B}$ to $\mathcal{C}$, then $[v]_\mathcal{C} = P\,[v]_\mathcal{B}$ and $[T]_\mathcal{C} = P[T]_\mathcal{B}P^{-1}$.

Definition 59 – Similar matrices
Consider two $n \times n$ matrices A and B. We say that A is **similar** to B if there exists an invertible matrix S such that
$$AS = SB, \quad \text{or} \quad B = S^{-1}AS.$$

4. Linear Spaces

4.1. Introduction to Linear Spaces.

Definition 60 – Vector Space
A (real) vector space is a set V with operations $+$ and scalar multiplication satisfying the vector space axioms.

4.2. Linear Transformations and Isomorphisms.

Definition 61 – Isomorphism
A bijective linear map is an **isomorphism**. Finite-dimensional vector spaces are isomorphic iff they have the same dimension.

4.3. The Matrix of a Linear Transformation.

Definition 62 – Matrix in Chosen Bases
Given bases $\mathcal{B}$ of V and $\mathcal{C}$ of W, the matrix $[T]_{\mathcal{C} \leftarrow \mathcal{B}}$ has columns $[T(v_i)]_{\mathcal{C}}$.

5. ORTHOGONALITY AND LEAST SQUARES

5.1. **Orthogonal Projections and Orthonormal Bases.**

Definition 63 – Orthogonality, length, unit vectors
a. Two vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^n$ are called perpendicular or orthogonal [1] if $\vec{v} \cdot \vec{w} = 0$.
b. The length (or magnitude or norm) of a vector $\vec{v}$ in $\mathbb{R}^n$ is $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$.
c. A vector $\vec{u}$ in $\mathbb{R}^n$ is called a unit vector if its length is 1, (i.e., $\|\vec{u}\| = 1$, or $\vec{u} \cdot \vec{u} = 1$).

Definition 64 – Orthonormal vectors
The vectors $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ in $\mathbb{R}^n$ are called orthonormal if they are all unit vectors and orthogonal to one another:
$$\vec{u}_i \cdot \vec{u}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Proposition 65 – Properties of orthonormal vectors
a. Orthonormal vectors are linearly independent.
b. Orthonormal vectors $\vec{u}_1, \ldots, \vec{u}_n$ in $\mathbb{R}^n$ form a basis of $\mathbb{R}^n$.

Proposition 66 – Linear Decomposition on a Basis
Consider an orthonormal basis $\vec{u}_1, \ldots, \vec{u}_n$ of $\mathbb{R}^n$. Then, for all $\vec{x}$ in $\mathbb{R}^n$,
$$\vec{x} = (\vec{u}_1 \cdot \vec{x})\, \vec{u}_1 + \cdots + (\vec{u}_n \cdot \vec{x})\, \vec{u}_n.$$

Proposition 67 – Orthogonal projection
Consider a vector $\vec{x}$ in $\mathbb{R}^n$ and a subspace V of $\mathbb{R}^n$. Then we can write
$$\vec{x} = \vec{x}^{\parallel} + \vec{x}^{\perp}$$
where $\vec{x}^{\parallel}$ is in V and $\vec{x}^{\perp}$ is perpendicular to V, and this representation is unique.
The vector $\vec{x}^{\parallel}$ is called the orthogonal projection of $\vec{x}$ onto V, denoted by $\mathrm{proj}_V \vec{x}$.
The transformation $T(\vec{x}) = \mathrm{proj}_V \vec{x} = \vec{x}^{\parallel}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$ is linear.
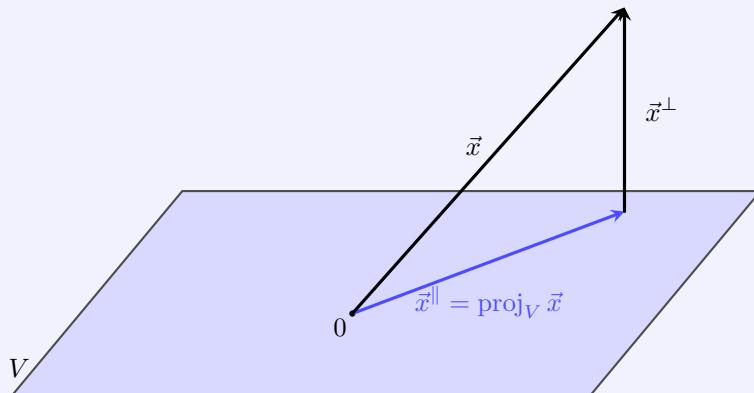
Theorem 68 – Formula for the orthogonal projection
If V is a subspace of $\mathbb{R}^n$ with an **orthonormal basis** $\vec{u}_1, \ldots, \vec{u}_m$, then
$$\operatorname{proj}_V \vec{x} = \vec{x}^{\parallel} = (\vec{u}_1 \cdot \vec{x})\, \vec{u}_1 + \cdots + (\vec{u}_m \cdot \vec{x})\, \vec{u}_m.$$
for all $\vec{x}$ in $\mathbb{R}^n$.

Warning 69   *The above is not true without an **orthonormal basis**.*

Definition 70 – Orthogonal complement
Consider a subspace V of $\mathbb{R}^n$. The orthogonal complement $V^{\perp}$ of V is the set of those vectors $\vec{x}$ in $\mathbb{R}^n$ that are orthogonal to all vectors in V :
$$V^{\perp} = \{\vec{x} \text{ in } \mathbb{R}^n : \vec{v} \cdot \vec{x} = 0, \text{ for all } \vec{v} \text{ in } V\}$$
Note that $V^{\perp}$ is the kernel of the orthogonal projection onto V.

Theorem 71 – Properties of the orthogonal complement
Consider a subspace V of $\mathbb{R}^n$.
a. *The orthogonal complement $V^{\perp}$ of V is a subspace of $\mathbb{R}^n$.*
b. *The intersection of V and $V^{\perp}$ consists of the zero vector: $V \cap V^{\perp} = \{\vec{0}\}$.*
c. $\dim(V) + \dim\left(V^{\perp}\right) = n$.
d. $\left(V^{\perp}\right)^{\perp} = V$.

Proposition 72 – Orthogonal Projection
*If Q has orthonormal columns spanning S, the orthogonal projector onto S is $P = QQ^{\top}$.*

5.2. **Gram–Schmidt Process.**

Figure 1

$$\vec{v}_2^{\perp} = \vec{v}_2 - (\vec{u}_1 \cdot \vec{v}_2)\,\vec{u}_1$$
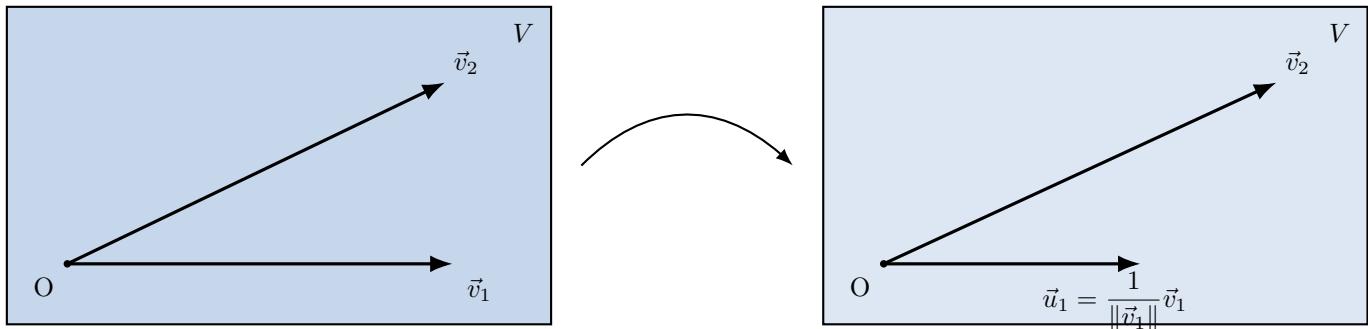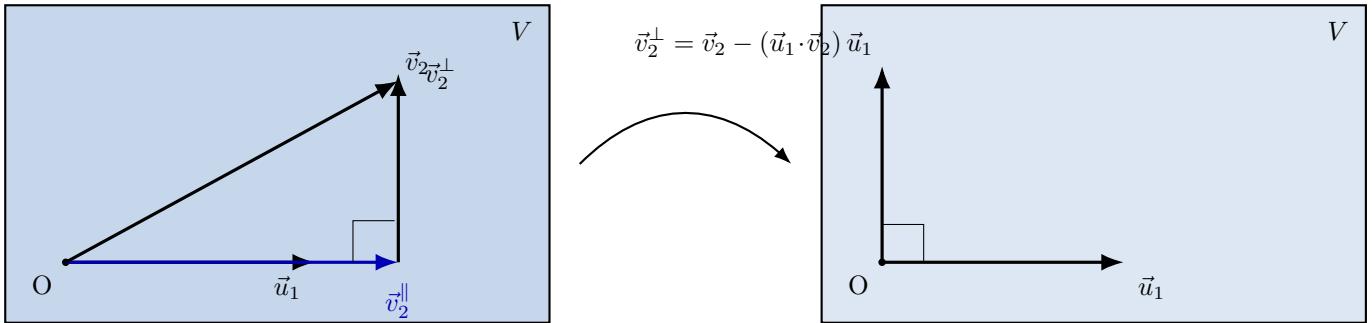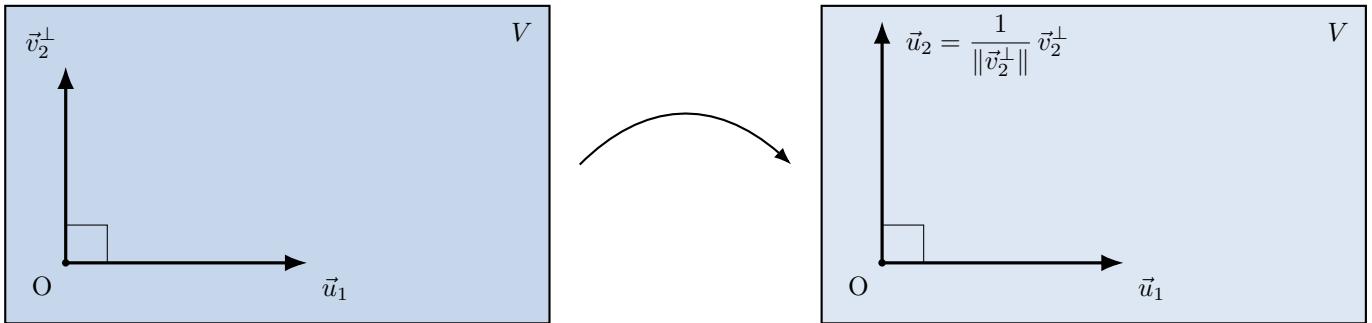
Figure 2

Figure 3

$$\vec{u}_3 = \frac{1}{\|\vec{v}_3^{\perp}\|}\,\vec{v}_3^{\perp}$$
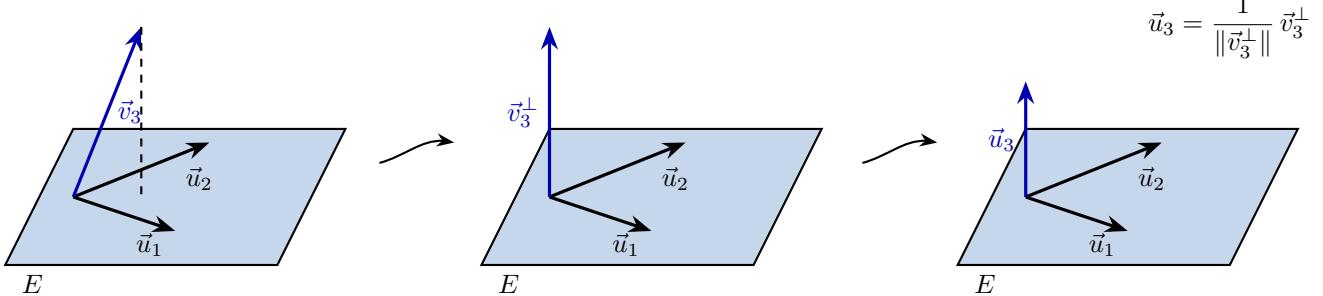
Figure 4

> **Theorem 73 – The Gram-Schmidt process**
> *Consider a basis $\vec{v}_1, \ldots, \vec{v}_m$ of a subspace V of $\mathbb{R}^n$. For $j = 2, \ldots, m$, we resolve the vector $\vec{v}_j$ into its components parallel and perpendicular to the span of the preceding vectors, $\vec{v}_1, \ldots, \vec{v}_{j-1}$:*
> $$\vec{v}_j = \vec{v}_j^{\|} + \vec{v}_j^{\perp}, \quad \text{with respect to } \operatorname{span}(\vec{v}_1, \ldots, \vec{v}_{j-1})$$
> *Then*
> $$\vec{u}_1 = \frac{1}{\|\vec{v}_1\|}\vec{v}_1, \quad \vec{u}_2 = \frac{1}{\|\vec{v}_2^{\perp}\|}\vec{v}_2^{\perp}, \ldots, \quad \vec{u}_j = \frac{1}{\|\vec{v}_j^{\perp}\|}\vec{v}_j^{\perp}, \ldots, \quad \vec{u}_m = \frac{1}{\|\vec{v}_m^{\perp}\|}\vec{v}_m^{\perp}$$
> *is an orthonormal basis of V. By Theorem 68, we have*
> $$\vec{v}_j^{\perp} = \vec{v}_j - \vec{v}_j^{\|} = \vec{v}_j - (\vec{u}_1 \cdot \vec{v}_j)\,\vec{u}_1 - \cdots - (\vec{u}_{j-1} \cdot \vec{v}_j)\,\vec{u}_{j-1}.$$

5.3. **Orthogonal Transformations and Orthogonal Matrices.**

Definition 74 – Orthogonal transformations and orthogonal matrices
A linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$ is called orthogonal if it preserves the length of vectors:
$$\|T(\vec{x})\| = \|\vec{x}\|, \quad \text{for all } \vec{x} \text{ in } \mathbb{R}^n.$$
If $T(\vec{x}) = A\vec{x}$ is an orthogonal transformation, we say that A is an orthogonal matrix.
Alternatively, a real $n \times n$ matrix Q is **orthogonal** if $Q^\top Q = I_n$.

Theorem 75
If Q in an $n \times n$ orthogonal matrix, then $Q^{-1} = Q^\top$ and $\|Qx\| = \|x\|$; orthogonal maps preserve inner products and angles 9thus orthogonality).

Theorem 76 – Orthogonal transformations and orthonormal bases
a. A linear transformation T from $\mathbb{R}^n$ to $\mathbb{R}^n$ is orthogonal if (and only if) the vectors $T(\vec{e}_1), T(\vec{e}_2), \ldots, T(\vec{e}_n)$ form an orthonormal basis of $\mathbb{R}^n$.
b. An $n \times n$ matrix A is orthogonal if (and only if) its columns form an orthonormal basis of $\mathbb{R}^n$.

Proposition 77 – Products and inverses of orthogonal matrices
a. The product AB of two orthogonal $n \times n$ matrices A and B is orthogonal.
b. The inverse A^{-1} of an orthogonal $n \times n$ matrix A is orthogonal.

Definition 78 – The transpose of a matrix; symmetric and skew-symmetric matrices
Consider an $m \times n$ matrix A. The transpose $A^\top$ of A is the $n \times m$ matrix whose ij th entry is the ji th entry of A : The roles of rows and columns are reversed.
We say that a square matrix A is symmetric if $A^\top = A$, and A is called skew-symmetric if $A^\top = -A$.

Proposition 79
If $\vec{v}$ and $\vec{w}$ are two (column) vectors in $\mathbb{R}^n$, then
$$\begin{array}{cc} \vec{v} \cdot \vec{w} & = \vec{v}^\top \vec{w} \\ \uparrow & \uparrow \\ Dot & Matrix \\ product & product \end{array}$$

Proposition 80
Consider an $n \times n$ matrix A. The matrix A is orthogonal if (and only if) $A^\top A = I_n$ or, equivalently, if $A^{-1} = A^\top$.

Theorem 81 – The matrix of an orthogonal projection
Consider a subspace V of $\mathbb{R}^n$ with orthonormal basis $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$. The matrix P of the orthogonal projection onto V is
$$P = QQ^\top, \quad \text{where} \quad Q = \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \ldots & \vec{u}_m \\ | & | & & | \end{bmatrix}.$$
Pay attention to the order of the factors ($QQ^\top$ as opposed to $Q^\top Q$). Note that matrix P is symmetric, since $P^\top = \left(QQ^\top\right)^\top = \left(Q^\top\right)^\top Q^\top = QQ^\top = P$.

Proposition 82 – Orthogonal matrices summary

Consider an $n \times n$ matrix A. Then the following statements are equivalent:

- *A is an orthogonal matrix.*
- *The transformation $L(\vec{x}) = A\vec{x}$ preserves length; that is, $\|A\vec{x}\| = \|\vec{x}\|$ for all $\vec{x}$ in $\mathbb{R}^n$.*
- *The columns of A form an orthonormal basis of $\mathbb{R}^n$.*
- *$A^\top A = I_n$.*
- *$A^{-1} = A^\top$.*
- *A preserves the dot product, meaning that $(A\vec{x}) \cdot (A\vec{y}) = \vec{x} \cdot \vec{y}$ for all $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^n$.*

5.4. **Least Squares and Data Fitting.**

Proposition 83 – Image, Kernel and Transpose

For any matrix A,
$$(\operatorname{im} A)^\perp = \ker\left(A^\top\right).$$

Proposition 84 – A special case

If A is an $n \times m$ matrix, then $\ker(A) = \ker\left(A^\top A\right)$.

If A is an $n \times m$ matrix with $\ker(A) = \{\vec{0}\}$, then $A^\top A$ is invertible.

Proposition 85 – An Alternative Characterization of Orthogonal Projections

Consider a vector $\vec{x}$ in $\mathbb{R}^n$ and a subspace V of $\mathbb{R}^n$. Then the orthogonal projection $\operatorname{proj}_V \vec{x}$ is the vector in V closest to $\vec{x}$, in that
$$\|\vec{x} - \operatorname{proj}_V \vec{x}\| < \|\vec{x} - \vec{v}\|$$
for all $\vec{v}$ in V different from $\operatorname{proj}_V \vec{x}$.

Definition 86 – Least-squares solution

Consider a linear system
$$A\vec{x} = \vec{b},$$
where A is an $n \times m$ matrix. A vector $\vec{x}^*$ in $\mathbb{R}^m$ is called a **least-squares solution** of this system if $\left\|\vec{b} - A\vec{x}^*\right\| \leq \|\vec{b} - A\vec{x}\|$ for all $\vec{x}$ in $\mathbb{R}^m$.

Theorem 87 – Normal Equations

The least-squares solutions of the system
$$A\vec{x} = \vec{b}$$
are the exact solutions of the (consistent) system
$$A^\top A\vec{x} = A^\top \vec{b}.$$
*The system $A^\top A\vec{x} = A^\top \vec{b}$ is called the **normal equation** of $A\vec{x} = \vec{b}$.*

If A has full column rank, the solution is unique.

Proposition 88 – The matrix of an orthogonal projection

Consider a subspace V of $\mathbb{R}^n$ with basis $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$. Let

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix}.$$

Then the matrix of the orthogonal projection onto V is

$$A \left(A^\top A \right)^{-1} A^\top.$$

Note that we are not required to find an orthonormal basis of V here, unlike in Theorem 81.

5.5. Inner Product Spaces.

Definition 89 – Inner Product Space

On a real vector space V, an inner product is a positive-definite symmetric bilinear form. It induces the norm $\|v\| = \sqrt{\langle v, v \rangle}$ and orthogonality.

6. DETERMINANTS

6.1. Introduction to Determinants.

Definition 90 – Determinant of a 3×3 matrix, in terms of the columns

If $A = \begin{bmatrix} \vec{u} & \vec{v} & \vec{w} \end{bmatrix}$, then

$$\det A = \vec{u} \cdot (\vec{v} \times \vec{w}).$$
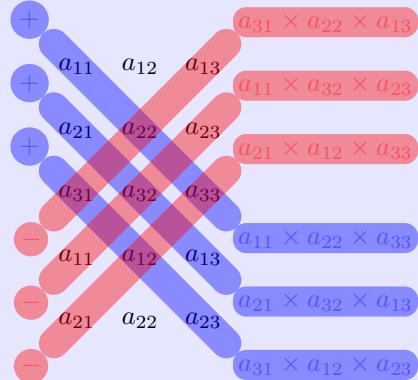
A 3×3 matrix A is invertible if (and only if) $\det A \neq 0$.

Theorem 91 – Sarrus's rule

To find the determinant of a 3×3 matrix A, write the first two columns of A to the right of A. Then multiply the entries along the six diagonals shown below.

Add or subtract these diagonal products, as shown in the diagram:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = (a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{33}) - (a_{13}a_{22}a_{31} + a_{23}a_{32}a_{11} + a_{33}a_{12}a_{31})$$

Definition 92 – Determinant
$\det A$ is the alternating multilinear function of rows (or columns) normalized by $\det I = 1$; equivalently by the Leibniz formula $\det A = \sum\limits_{\sigma \in S_n} \text{sgn}(\sigma) \prod\limits_i a_{i,\sigma(i)}$.

Theorem 93 – Determinant of a triangular matrix
The determinant of an (upper or lower) triangular matrix is the product of the diagonal entries of the matrix.
In particular, the determinant of a diagonal matrix is the product of its diagonal entries.

Theorem 94 – Determinant of a block matrix
If $M = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$, *where A and C are square matrices (not necessarily of the same size), then*

$$\det \begin{bmatrix} A & B \\ 0 & C \end{bmatrix} = (\det A)(\det C).$$

Likewise,

$$\det \begin{bmatrix} A & 0 \\ B & C \end{bmatrix} = (\det A)(\det C).$$

6.2. **Properties of the Determinant.**

Theorem 95 – Basic Properties
Row operations affect $\det$ as follows:
- *swapping rows flips sign;*
- *scaling a row scales $\det$;*
- *adding a multiple of one row to another leaves $\det$ unchanged;*
- *Also $\det(AB) = \det A \cdot \det B$.*

Theorem 96 – Determinant of the transpose
If A is a square matrix, then
$$\det\left(A^\top\right) = \det A.$$

Theorem 97 – Linearity of the determinant in the rows and columns
Consider fixed row vectors $\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n$ with n components. Then the function

$$T(\vec{x}) = \det \begin{bmatrix} - & \vec{v}_1 & - \\ & \vdots & \\ - & \vec{v}_{i-1} & - \\ - & \vec{x} & - \\ - & \vec{v}_{i+1} & - \\ & \vdots & \\ - & \vec{v}_n & - \end{bmatrix} \quad \text{from } \mathbb{R}^{1 \times n} \text{ to } \mathbb{R}$$

*is a linear transformation. This property is referred to as **linearity of the determinant in the ith row**. Likewise, the determinant is **linear in all the columns**.*

Theorem 98 – Elementary row operations and determinants
a. If B is obtained from A by dividing a row of A by a scalar k, then $\det B = (1/k)\det A$.
b. If B is obtained from A by a row swap, then

$$\det B = -\det A.$$

We say that the determinant is alternating on the rows.
c. If B is obtained from A by adding a multiple of a row of A to another row, then

$$\det B = \det A.$$

Analogous results hold for elementary column operations.

Theorem 99 – Invertibility and determinant
A square matrix A is invertible if and only if $\det A \neq 0$.

Theorem 100 – Determinants of products and powers
If A and B are $n \times n$ matrices and m is a positive integer, then
a. $\det(AB) = (\det A)(\det B)$, and
b. $\det(A^m) = (\det A)^m$.

Theorem 101 – Determinants of similar matrices
If matrix A is similar to B, then $\det A = \det B$.

Warning 102   *The converse is not necessarily true. For instance Let $A = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$.*
Then $\det(A) = \det(B)$. However, A and B are not similar.
Similar matrices must have the same eigenvalues (with the same algebraic multiplicities). Here

$$Spec(A) = \{1,1\}, \qquad Spec(B) = \{2, \tfrac{1}{2}\},$$

so A and B do not have the same eigenvalues, and therefore cannot be similar.

Theorem 103 – Determinant of an inverse
If A is an invertible matrix, then

$$\det\left(A^{-1}\right) = \frac{1}{\det A} = (\det A)^{-1}.$$

6.3. **Geometrical Interpretations; Cramer's Rule.**

Proposition 104 – Geometry and Cramer
$|\det A|$ equals the volume scaling of $x \mapsto Ax$. For invertible A, Cramer's rule gives $x_i = \dfrac{\det A_i}{\det A}$.

7. EIGENVALUES AND EIGENVECTORS

7.1. **Diagonalization.**

Definition 105 – Diagonalizable Matrix

Consider a linear transformation $T(\vec{x}) = A\vec{x}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$. Then A (or T) is said to be **diagonalizable** if the matrix B of T with respect to some basis is diagonal.

Equivalently, A is **diagonalizable** if $A = SDS^{-1}$ with D diagonal. From Definition 59, this means A is similar to D.

Equivalently there exists a basis of eigenvectors (called eigenbasis).

Definition 106 – Eigenvectors, eigenvalues, and eigenbases

Consider a linear transformation $T(\vec{x}) = A\vec{x}$ from $\mathbb{R}^n$ to $\mathbb{R}^n$. A **nonzero** vector $\vec{v}$ in $\mathbb{R}^n$ is called an **eigenvector** of A (or T) if

$$A\vec{v} = \lambda\vec{v}$$

for some scalar λ. This λ is called the **eigenvalue** associated with eigenvector $\vec{v}$.

A basis $\vec{v}_1, \ldots, \vec{v}_n$ of $\mathbb{R}^n$ is called an **eigenbasis** for A (or T) if the vectors $\vec{v}_1, \ldots, \vec{v}_n$ are eigenvectors of A, meaning that $A\vec{v}_1 = \lambda_1\vec{v}_1, \ldots, A\vec{v}_n = \lambda_n\vec{v}_n$ for some scalars $\lambda_1, \ldots, \lambda_n$.

Theorem 107 – Eigenbases and diagonalization

The matrix A is diagonalizable if (and only if) there exists an eigenbasis for A.

If $\vec{v}_1, \ldots, \vec{v}_n$ is an eigenbasis for A, with $A\vec{v}_1 = \lambda_1\vec{v}_1, \ldots, A\vec{v}_n = \lambda_n\vec{v}_n$, then the matrices

$$S = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \quad and \quad B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

will diagonalize A, meaning that $S^{-1}AS = B$.

Conversely, if the matrices S and B diagonalize A, then the column vectors of S will form an eigenbasis for A, and the diagonal entries of B will be the associated eigenvalues.

Theorem 108 – Eigenvalues of an orthogonal matrix

*The possible **real** eigenvalues of an orthogonal matrix are 1 and -1.*

Theorem 109 – Various characterizations of invertible matrices

For an $n \times n$ matrix A, the following statements are equivalent.

- *(i) A is invertible.*
- *(ii) The linear system $A\vec{x} = \vec{b}$ has a unique solution $\vec{x}$, for all $\vec{b}$ in $\mathbb{R}^n$.*
- *(iii) $\operatorname{rref} A = I_n$.*
- *(iv) $\operatorname{rank} A = n$.*
- *(v) $\operatorname{im} A = \mathbb{R}^n$.*
- *(vi) $\ker A = \{\vec{0}\}$.*
- *(vii) The column vectors of A form a basis of $\mathbb{R}^n$.*
- *(viii) The column vectors of A span $\mathbb{R}^n$.*
- *(ix) The column vectors of A are linearly independent.*
- *(x) $\det A \neq 0$.*
- *(xi) 0 fails to be an eigenvalue of A.*

Theorem 110 – Discrete dynamical systems

Consider the dynamical system

$$\vec{x}(t+1) = A\vec{x}(t) \quad with \quad \vec{x}(0) = \vec{x}_0.$$

Then $\vec{x}(t) = A^\top \vec{x}_0$. Suppose we can find an eigenbasis $\vec{v}_1, \ldots, \vec{v}_n$ for A, with $A\vec{v}_1 = \lambda_1 \vec{v}_1, \ldots, A\vec{v}_n = \lambda_n \vec{v}_n$. Find the coordinates $c_1, \ldots, c_n$ of the vector $\vec{x}_0$ with respect to the eigenbasis $\vec{v}_1, \ldots, \vec{v}_n$:

$$\vec{x}_0 = c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n.$$

Then

$$\vec{x}(t) = A^\top \vec{x}_0 = c_1 A^\top \vec{v}_1 + \cdots + c_n A^\top \vec{v}_n = c_1 \lambda_1^\top \vec{v}_1 + \cdots + c_n \lambda_n^\top \vec{v}_n.$$

7.2. Finding the Eigenvalues of a Matrix.

Theorem 111 – Characteristic Polynomial

If A is an $n \times n$ matrix, then $\det(A - \lambda I_n)$ is a polynomial of degree n, of the form

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

This is called the characteristic polynomial of A, denoted by $f_A(\lambda)$.
Some define the characteristic polynomial of A as $\det(\lambda I_n - A)$, which is a multiplication by $(-1)^n$.

Theorem 112 – Eigenvalues and determinants; characteristic equation

Consider an $n \times n$ matrix A and a scalar λ. Then λ is an eigenvalue of A if and only if

$$\det(A - \lambda I_n) = 0$$

This is called the characteristic equation (or the secular equation) of matrix A.

Definition 113 – Algebraic multiplicity of an eigenvalue

We say that an eigenvalue λ_0 of a square matrix A has **algebraic multiplicity** k if λ_0 is a root of multiplicity k of the characteristic polynomial $f_A(\lambda)$, meaning that we can write

$$f_A(\lambda) = (\lambda_0 - \lambda)^k g(\lambda)$$

for some polynomial $g(\lambda)$ with $g(\lambda_0) \neq 0$. We write $\mathrm{almu}(\lambda_0) = k$.

Theorem 114 – Number of eigenvalues

An $n \times n$ matrix has at most n real eigenvalues, even if they are counted with their algebraic multiplicities.
If n is odd, then an $n \times n$ matrix has at least one real eigenvalue.

Theorem 115 – Eigenvalues, determinant, and trace

If an $n \times n$ matrix A has the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, listed with their algebraic multiplicities, then

$$\det A = \lambda_1 \lambda_2 \cdots \lambda_n, \quad the\ product\ of\ the\ eigenvalues$$

and

$$\operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n, \quad the\ sum\ of\ the\ eigenvalues.$$

7.3. Finding the Eigenvectors of a Matrix.

Definition 116 – Eigenspace

Consider an eigenvalue λ of an $n \times n$ matrix A. Then the kernel of the matrix $A - \lambda I_n$ is called the **eigenspace** associated with λ, denoted by E_λ :

$$E_\lambda = \ker(A - \lambda I_n) = \{\vec{v} \text{ in } \mathbb{R}^n : A\vec{v} = \lambda\vec{v}\}.$$

Definition 117 – Geometric multiplicity

Consider an eigenvalue λ of an $n \times n$ matrix A. The dimension of eigenspace $E_\lambda = \ker(A - \lambda I_n)$ is called the **geometric multiplicity** of eigenvalue λ, denoted gemu(λ). Thus,

$$\text{gemu}(\lambda) = \text{nullity}(A - \lambda I_n) = n - \text{rank}(A - \lambda I_n).$$

Theorem 118 – Eigenbases and geometric multiplicities

a. Consider an $n \times n$ matrix A. If we find a basis of each eigenspace of A and concatenate all these bases, then the resulting eigenvectors $\vec{v}_1, \ldots, \vec{v}_s$ will be linearly independent. (Note that s is the sum of the geometric multiplicities of the eigenvalues of A.) This result implies that $s \leq n$.
b. Matrix A is diagonalizable if (and only if) the geometric multiplicities of the eigenvalues add up to n (meaning that $s = n$ in part a).

Theorem 119 – An $n \times n$ matrix with n distinct eigenvalues

If an $n \times n$ matrix A has n distinct eigenvalues, then A is diagonalizable. We can construct an eigenbasis by finding an eigenvector for each eigenvalue.

Theorem 120 – The eigenvalues of similar matrices

Suppose matrix A is similar to B. Then
a. Matrices A and B have the same characteristic polynomial; that is, $f_A(\lambda) = f_B(\lambda)$.
b. rank $A =$ rank B and nullity $A =$ nullity B.
c. Matrices A and B have the same eigenvalues, with the same algebraic and geometric multiplicities. (However, the eigenvectors need not be the same.)
d. Matrices A and B have the same determinant and the same trace: $\det A = \det B$ and $\operatorname{tr} A = \operatorname{tr} B$. (However, A and B have the same determinant and the same trace and not be similar.)

Theorem 121 – Algebraic versus geometric multiplicity

If λ is an eigenvalue of a square matrix A, then

$$\text{gemu}(\lambda) \leq \text{almu}(\lambda).$$

Technics 122 – Strategy for Diagonalization

Suppose we are asked to determine whether a given $n \times n$ matrix A is diagonalizable. If so, we wish to find an invertible matrix S such that $S^{-1}AS = B$ is diagonal.

We can proceed as follows.

a. Find the eigenvalues of A by solving the characteristic equation

$$f_A(\lambda) = \det(A - \lambda I_n) = 0.$$

b. For each eigenvalue λ, find a basis of the eigenspace

$$E_\lambda = \ker(A - \lambda I_n).$$

c. Matrix A is diagonalizable if (and only if) the dimensions of the eigenspaces add up to n. In this case, we find an eigenbasis $\vec{v}_1, \ldots, \vec{v}_n$ for A by concatenating the bases of the eigenspaces we found in part b. Let

$$S = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}. \text{ Then } S^{-1}AS = B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix},$$

where λ_j is the eigenvalue associated with $\vec{v}_j$.

Theorem 123 – Cayley-Hamilton   *A square matrix satisfies its characteristic polynomial. That is*

$$f_A(A) = 0.$$

7.4. **More on Dynamical Systems.**

Proposition 124 – Discrete Linear Systems

For $x_{k+1} = Ax_k$, solutions are combinations of $\lambda_i^k v_i$ along eigen-directions; stability depends on $|\lambda_i|$.

Theorem 125 – Powers of a diagonalizable matrix

If

$$S^{-1}AS = B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

then

$$A^\top = SB^\top S^{-1} = S \begin{bmatrix} \lambda_1^\top & 0 & \cdots & 0 \\ 0 & \lambda_2^\top & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n^\top \end{bmatrix} S^{-1}$$

7.5. **Complex Eigenvalues.**

Theorem 126 – Fundamental theorem of algebra

Any polynomial $p(\lambda)$ with complex coefficients splits; that is, it can be written as a product of linear factors

$$p(\lambda) = k\,(\lambda - \lambda_1)\,(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)\,,$$

for some complex numbers $\lambda_1, \lambda_2, \ldots, \lambda_n$, and k. (The λ_i need not be distinct.) Therefore, a polynomial $p(\lambda)$ of degree n has precisely n complex roots if they are properly counted with their multiplicities.

Proposition 127 – Complex Pairs

Real matrices may have complex conjugate eigenpairs; real dynamics then mix rotations and scalings.

Theorem 128

For a 2×2 matrix A with complex eigenvalues $\lambda = \alpha \pm i\beta$, the matrix exponential is given by the formula:

$$e^{At} = e^{\alpha t}\left(\cos(\beta t)I_2 + \frac{\sin(\beta t)}{\beta}(A - \alpha I_2)\right).$$

One proof uses Cayley-Hamilton's theorem, 123.

7.6. **Stability.**

Definition 129 – Stability Criteria/Equilibrium

Consider a dynamical system

$$\vec{x}(t+1) = A\vec{x}(t).$$

We say that $\overrightarrow{0}$ is an (asymptotically) stable equilibrium for this system if

$$\lim_{t\to\infty} \vec{x}(t) = \lim_{t\to\infty}\left(A^t\vec{x}_0\right) = \overrightarrow{0}$$

for all vectors $\vec{x}_0$ in $\mathbb{R}^n$.

Theorem 130 – Stability and eigenvalues

Consider a dynamical system $\vec{x}(t+1) = A\vec{x}(t)$. The zero state is asymptotically stable if (and only if) the modulus of all the complex eigenvalues of A is less than 1.

Theorem 131 – Dynamical systems with complex eigenvalues

Consider the dynamical system $\vec{x}(t+1) = A\vec{x}(t)$, where A is a real 2×2 matrix with eigenvalues

$$\lambda_{1,2} = p \pm iq = r(\cos(\theta) \pm i\sin(\theta)), \quad \text{where} \quad q \neq 0.$$

Let $\vec{v} + i\vec{w}$ be an eigenvector of A with eigenvalue $p + iq$. Then

$$\vec{x}(t) = r^t S \begin{bmatrix} \cos(\theta t) & -\sin(\theta t) \\ \sin(\theta t) & \cos(\theta t) \end{bmatrix} S^{-1}\vec{x}_0, \quad \text{where} \quad S = \begin{bmatrix} \vec{w} & \vec{v} \end{bmatrix}.$$

Note that $S^{-1}\vec{x}_0$ is the coordinate vector of $\vec{x}_0$ with respect to basis $\vec{w}, \vec{v}$.

Theorem 132 – Phase portrait of a system with complex eigenvalues

Consider a dynamical system

$$\vec{x}(t+1) = A\vec{x}(t),$$

where A is a real 2×2 matrix with eigenvalues $\lambda_{1,2} = p \pm iq$ (where $q \neq 0$). Let

$$r = |\lambda_1| = |\lambda_2| = \sqrt{p^2 + q^2}$$

If $r = 1$, then the points $\vec{x}(t)$ are located on an ellipse; if r exceeds 1 , then the trajectory spirals outward; and if r is less than 1 , then the trajectory spirals inward, approaching the origin.

8. Symmetric Matrices and Quadratic Forms

8.1. Symmetric Matrices.

Theorem 133 – Spectral Theorem
A matrix A is orthogonally diagonalizable (i.e., there exists an orthogonal S such that $S^{-1}AS = S^{\top}AS$ is diagonal) if and only if A is symmetric (i.e., $A^{\top} = A$).

Theorem 134
Consider a symmetric matrix A. If $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors of A with distinct eigenvalues λ_1 and λ_2, then $\vec{v}_1 \cdot \vec{v}_2 = 0$; that is, $\vec{v}_2$ is orthogonal to $\vec{v}_1$.

Theorem 135
A symmetric $n \times n$ matrix A has n real eigenvalues if they are counted with their algebraic multiplicities.

Theorem 136   *Orthogonal diagonalization of a symmetric matrix A*
a. Find the eigenvalues of A, and find a basis of each eigenspace.
b. Using the Gram-Schmidt process, find an orthonormal basis of each eigenspace.
c. Form an orthonormal eigenbasis $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ for A by concatenating the orthonormal bases you found in part b, and let

$$
S = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}.
$$

S is orthogonal, and $S^{-1}AS$ will be diagonal.

8.2. Quadratic Forms.

Definition 137 – Quadratic Forms
A function $q(x_1, x_2, \ldots, x_n)$ from $\mathbb{R}^n$ to $\mathbb{R}$ is called a **quadratic form** if it is a linear combination of functions of the form $x_i x_j$ (where i and j may be equal). A quadratic form can be written as

$$
q(\vec{x}) = \vec{x} \cdot A\vec{x} = \vec{x}^{\top} A\vec{x}
$$

for a unique symmetric $n \times n$ matrix A, called the matrix of q.

Theorem 138 – Diagonalizing a quadratic form
Consider a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, where A is a symmetric $n \times n$ matrix. Let $\mathfrak{B}$ be an orthonormal eigenbasis for A, with associated eigenvalues $\lambda_1, \ldots, \lambda_n$. Then

$$
q(\vec{x}) = \lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2,
$$

where the c_i are the coordinates of $\vec{x}$ with respect to $\mathfrak{B}$.[2]

Definition 139 – Definiteness of a quadratic form

Consider a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, where A is a symmetric $n \times n$ matrix. We say that A is **positive definite** if $q(\vec{x})$ is positive for all nonzero $\vec{x}$ in $\mathbb{R}^n$, and we call A **positive semidefinite** if $q(\vec{x}) \geq 0$, for all $\vec{x}$ in $\mathbb{R}^n$.

Negative definite and negative semidefinite symmetric matrices are defined analogously.

Finally, we call A indefinite if q takes positive as well as negative values.

Theorem 140 – Eigenvalues and definiteness

A symmetric matrix A is positive definite if (and only if) all of its eigenvalues are positive. The matrix A is positive semidefinite if (and only if) all of its eigenvalues are positive or zero.

Theorem 141 – Principal submatrices and definiteness

Consider a symmetric $n \times n$ matrix A. For $m = 1, \ldots, n$, let $A^{(m)}$ be the $m \times m$ matrix obtained by omitting all rows and columns of A past the m-th. These matrices $A^{(m)}$ are called the principal submatrices of A.

The matrix A is positive definite if (and only if) $\det\left(A^{(m)}\right) > 0$, for all $m = 1, \ldots, n$.

Definition 142 – Principal axes

Consider a quadratic form $q(\vec{x}) = \vec{x} \cdot A\vec{x}$, where A is a symmetric $n \times n$ matrix with n distinct eigenvalues. Then the eigenspaces of A are called the **principal axes** of q. (Note that these eigenspaces will be one-dimensional.)

Theorem 143 – Ellipses and hyperbolas

Consider the curve C in $\mathbb{R}^2$ defined by

$$q(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2 = 1.$$

Let λ_1 and λ_2 be the eigenvalues of the matrix $\begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$ of q.

*If both λ_1 and λ_2 are positive, then C is an **ellipse**.*
*If one eigenvalue is positive and the other is negative, then C is a **hyperbola**.*

8.3. **Singular Values.**

Definition 144 – Singular values

The singular values of an $n \times m$ matrix A are the square roots of the eigenvalues of the symmetric $m \times m$ matrix $A^\top A$, listed with their algebraic multiplicities. It is customary to denote the singular values by $\sigma_1, \sigma_2, \ldots, \sigma_m$ and to list them in decreasing order:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m \geq 0$$

Theorem 145 – Singular Value Decomposition

For any $A \in \mathbb{R}^{m \times n}$, $A = U\Sigma V^\top$ with U, V orthogonal and $\Sigma = \mathrm{diag}(\sigma_i)$, $\sigma_i \geq 0$. Moreover $\sigma_i = \sqrt{\lambda_i(A^\top A)}$.

9. LINEAR DIFFERENTIAL EQUATIONS

9.1. **An Introduction to Continuous Dynamical Systems.**

Theorem 146 – Exponential growth and decay
Consider the linear differential equation
$$\frac{dx}{dt} = kx$$
with initial value x_0 (k is an arbitrary constant). The solution is
$$x(t) = e^{kt}x_0.$$
The quantity x will grow or decay exponentially (depending on the sign of k).

Theorem 147 – Linear dynamical systems: Discrete versus continuous
A linear dynamical system can be modeled by
$$\vec{x}(t+1) = B\vec{x}(t) \quad (\text{ discrete model })$$
or
$$\frac{d\vec{x}}{dt} = A\vec{x} \quad (\text{continuous model})$$
A and B are $n \times n$ matrices, where n is the number of components of the system.

Theorem 148 – Continuous dynamical systems
Consider the system $d\vec{x}/dt = A\vec{x}$. Suppose there is a real eigenbasis $\vec{v}_1, \ldots, \vec{v}_n$ for A, with associated eigenvalues $\lambda_1, \ldots, \lambda_n$. Then the general solution of the system is
$$\vec{x}(t) = c_1 e^{\lambda_1 t}\vec{v}_1 + \cdots + c_n e^{\lambda_n t}\vec{v}_n$$
The scalars $c_1, c_2, \ldots, c_n$ are the coordinates of $\vec{x}_0$ with respect to the basis $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$.
We can write the preceding equation in matrix form as

$$\vec{x}(t) = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

$$= S \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix} S^{-1}\vec{x}_0,$$

where $S = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$.

Definition 149 – Matrix Exponential
For $\dot{x} = Ax$, the solution is $x(t) = e^{tA}x(0)$ with $e^{tA} = \sum_{k \geq 0} \frac{t^k}{k!}A^k$.

9.2. **The Complex Case: Euler's Formula.**

Theorem 150 – Euler
$e^{i\theta} = \cos\theta + i\sin\theta$. *Complex eigenpairs produce oscillatory real solutions via real/imag parts.*

Theorem 151 – Continuous dynamical systems with complex eigenvalues

Consider a linear system
$$\frac{d\vec{x}}{dt} = A\vec{x}.$$
Suppose there exists a complex eigenbasis $\vec{v}_1, \ldots, \vec{v}_n$ for A, with associated complex eigenvalues $\lambda_1, \ldots, \lambda_n$. Then the general complex solution of the system is
$$\vec{x}(t) = c_1 e^{\lambda_1 t}\vec{v}_1 + \cdots + c_n e^{\lambda_n t}\vec{v}_n,$$
where the c_i are arbitrary complex numbers.
We can write this solution in matrix form, as in Theorem 148

Theorem 152 – Stability of a continuous dynamical system

For a system
$$\frac{d\vec{x}}{dt} = A\vec{x},$$
the zero state is an asymptotically stable equilibrium solution if (and only if) the real parts of all eigenvalues of A are negative.

Theorem 153 – Determinant, trace, and stability

Consider the system
$$\frac{d\vec{x}}{dt} = A\vec{x}$$
where A is a real 2×2 matrix. Then the zero state is an asymptotically stable equilibrium solution if (and only if) $\operatorname{tr} A < 0$ and $\det A > 0$.

Theorem 154 – Continuous dynamical systems with eigenvalues $p \pm iq$

Consider the linear system
$$\frac{d\vec{x}}{dt} = A\vec{x},$$
where A is a real 2×2 matrix with complex eigenvalues $p \pm iq$ (and $q \neq 0$). Consider an eigenvector $\vec{v} + i\vec{w}$ with eigenvalue $p + iq$. Then
$$\vec{x}(t) = e^{pt}S \begin{bmatrix} \cos(qt) & -\sin(qt) \\ \sin(qt) & \cos(qt) \end{bmatrix} S^{-1}\vec{x}_0,$$
where $S = \begin{bmatrix} \vec{w} & \vec{v} \end{bmatrix}$. Recall that $S^{-1}\vec{x}_0$ is the coordinate vector of $\vec{x}_0$ with respect to basis $\vec{w}, \vec{v}$. The trajectories are either ellipses (linearly distorted circles), if $p = 0$, or spirals, spiraling outward if p is positive and inward if p is negative. In the case of an ellipse, the trajectories have a period of $2\pi/q$.

9.3. **Linear Differential Operators and Linear Differential Equations.**

Definition 155 – Linear differential operators and linear differential equations

A transformation T from C^∞ to C^∞ of the form
$$T(f) = f^{(n)} + a_{n-1}f^{(n-1)} + \cdots + a_1 f' + a_0 f$$
is called an nth-order **linear differential operator**. Here $f^{(k)}$ denotes the kth derivative of function f, and the coefficients a_k are complex scalars.
If T is an nth-order linear differential operator and g is a smooth function, then the equation
$$T(f) = g \quad \text{or} \quad f^{(n)} + a_{n-1}f^{(n-1)} + \cdots + a_1 f' + a_0 f = g$$
is called an nth-order **linear differential equation** (DE). The DE is called **homogeneous** if $g = 0$ and **inhomogeneous** otherwise.

Theorem 156

Consider a linear transformation T from V to W, where V and W are arbitrary linear spaces. Suppose we have a basis $f_1, f_2, \ldots, f_n$ of the kernel of T. Consider an equation $T(f) = g$ with a particular solution f_p. Then the solutions f of the equation $T(f) = g$ are of the form

$$f = c_1 f_1 + c_2 f_2 + \cdots + c_n f_n + f_p,$$

where the c_i are arbitrary constants.

Theorem 157

The kernel of an nth-order linear differential operator is n-dimensional.

Technics 158 – Strategy for solving linear differential equations

To solve an nth-order linear DE $T(f) = g$, we have to find
a. a basis $f_1, \ldots, f_n$ of $\ker(T)$, and
b. a particular solution f_p of the DE $T(f) = g$.
Then the solutions f are of the form

$$f = c_1 f_1 + \cdots + c_n f_n + f_p,$$

where the c_i are arbitrary constants.

Definition 159 – Eigenfunctions

Consider a linear differential operator T from C^∞ to C^∞. A smooth function f is called an **eigenfunction** of T if $T(f) = \lambda f$ for some complex scalar λ; this scalar λ is called the eigenvalue associated with the eigenfunction f.

Definition 160 – Characteristic polynomial

Consider the linear differential operator

$$T(f) = f^{(n)} + a_{n-1} f^{(n-1)} + \cdots + a_1 f' + a_0 f$$

The **characteristic polynomial** of T is defined as

$$p_T(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0.$$

Theorem 161

If T is a linear differential operator, then $e^{\lambda t}$ is an eigenfunction of T, with associated eigenvalue $p_T(\lambda)$, for all λ :

$$T\left(e^{\lambda t}\right) = p_T(\lambda)e^{\lambda t}.$$

In particular, if $p_T(\lambda) = 0$, then $e^{\lambda t}$ is in the kernel of T.

Theorem 162 – The kernel of a linear differential operator

Consider an nth-order linear differential operator T whose characteristic polynomial $p_T(\lambda)$ has n distinct roots $\lambda_1, \ldots, \lambda_n$. Then the exponential functions

$$e^{\lambda_1 t}, e^{\lambda_2 t}, \ldots, e^{\lambda_n t}$$

form a basis of the kernel of T; that is, they form a basis of the solution space of the homogeneous DE

$$T(f) = 0.$$

Theorem 163 – With Complex Roots

Consider a differential equation

$$T(f) = f'' + af' + bf = 0,$$

where the coefficients a and b are real. Suppose the zeros of $p_T(\lambda)$ are $p \pm iq$, with $q \neq 0$. Then the solutions of the given DE are

$$f(t) = e^{pt} \left(c_1 \cos(qt) + c_2 \sin(qt) \right),$$

where c_1 and c_2 are arbitrary constants. The special case when $a = 0$ and $b > 0$ is important in many applications. Then $p = 0$ and $q = \sqrt{b}$, so that the solutions of the DE

$$f'' + bf = 0$$

are

$$f(t) = c_1 \cos\left(\sqrt{b}t\right) + c_2 \sin\left(\sqrt{b}t\right).$$

Theorem 164

Recall the notation $Df = f'$ for the derivative operator.
We let $D^m = \underbrace{D \circ D \circ \cdots \circ D}_{m \ times}$; that is, $D^m f = f^{(m)}$. Then the operator

$$T(f) = f^{(n)} + a_{n-1} f^{(n-1)} + \cdots + a_1 f' + a_0 f$$

can be written more succinctly as

$$T = D^n + a_{n-1} D^{n-1} + \cdots + a_1 D + a_0,$$

the characteristic polynomial $p_T(\lambda)$ "evaluated at D".
An n th-order linear differential operator T can be expressed as the composite of n first-order linear differential operators:

$$T = D^n + a_{n-1} D^{n-1} + \cdots + a_1 D + a_0$$
$$= (D - \lambda_1)(D - \lambda_2) \cdots (D - \lambda_n)$$

where the λ_i are complex numbers.

Theorem 165

Consider the linear differential equation

$$f''(t) + af'(t) + bf(t) = C \cos(\omega t)$$

where a, b, C, and ω are real numbers. Suppose that $a \neq 0$ or $b \neq \omega^2$. This DE has a particular solution of the form

$$f_p(t) = P \cos(\omega t) + Q \sin(\omega t)$$

Now use Technics 158 and Theorems 162 to find all solutions f of the DE.

Technics 166 – Strategy for linear differential equations

Suppose you have to solve an n th-order linear differential equation $T(f) = g$.

Step 1 *Find n linearly independent solutions of the DE $T(f) = 0$.*

- *Write the characteristic polynomial $p_T(\lambda)$ of T [replacing $f^{(k)}$ by λ^k].*
- *Find the solutions $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the equation $p_T(\lambda) = 0$.*
- *If λ is a solution of the equation $p_T(\lambda) = 0$, then $e^{\lambda t}$ is a solution of $T(f) = 0$.*
- *If λ is a solution of $p_T(\lambda) = 0$ with multiplicity m, then $e^{\lambda t}, te^{\lambda t}, t^2 e^{\lambda t}, \ldots, t^{m-1}e^{\lambda t}$ are solutions of the DE $T(f) = 0$.*
- *If $p \pm iq$ are complex solutions of $p_T(\lambda) = 0$, then $e^{pt}\cos(qt)$ and $e^{pt}\sin(qt)$ are real solutions of the DE $T(f) = 0$.*

Step 2 *If the DE is inhomogeneous (i.e., if $g \neq 0$), find one particular solution f_p of the DE $T(f) = g$.*

- *If g is of the form $g(t) = A\cos(\omega t) + B\sin(\omega t)$, look for a particular solution of the same form, $f_p(t) = P\cos(\omega t) + Q\sin(\omega t)$.*
- *If g is constant, look for a constant particular solution $f_p(t) = c$.[7]*
- *If the DE is of first order, of the form $f'(t) - af(t) = g(t)$, use the formula $f(t) = e^{at}\int e^{-at}g(t)dt$.*
- *If none of the preceding techniques applies, write $T = (D - \lambda_1)(D - \lambda_2)\cdots(D - \lambda_n)$, and solve the corresponding first-order DEs.*

Step 3 *The solutions of the DE $T(f) = g$ are of the form*

$$f(t) = c_1 f_1(t) + c_2 f_2(t) + \cdots + c_n f_n(t) + f_p(t)$$

where $f_1, f_2, \ldots, f_n$ are the solutions from step 1 and f_p is the solution from step 2.