Linear Algebra Course Notes

David Yang

Swarthmore College, Fall 2020

1 Chapter 1: Vectors, Matrices, and Linear Systems

Preliminary Definitions

Definition 1.1. *The* ***dot product*** *is defined as follows: for* $\vec{v} = [v_1, v_2, ..., v_n], \vec{w} = [w_1, w_2, ..., w_n] \in \mathbb{R}^n$,

$$\vec{v} \cdot \vec{w} = \sum_{i=1}^{n} v_i w_i$$

Equivalently, $\vec{v} \cdot \vec{w} = \|\vec{v}\| \|\vec{w}\| \cos\theta$ *where* θ *is the angle between* $\vec{v}, \vec{w}$.

Theorem 1.1 (Properties of Dot Product)**.** *Let* $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ *and* $r \in \mathbb{R}$. *Then*
(1: Commutativity) $\vec{v} \cdot \vec{w} = \vec{w} \cdot \vec{v}$
(2: Distributivity) $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$
(3: Homogeneity) $r(\vec{v} \cdot \vec{w}) = (r\vec{v}) \cdot \vec{w} = \vec{v} \cdot (r\vec{w})$
(4: Positivity) $\vec{v} \cdot \vec{v} \geq 0$ *and* $\vec{v} \cdot \vec{v} = 0 \iff \vec{v} = 0$

Definition 1.2. *The* ***norm/magnitude/length*** *of a vector* $\vec{v}$ *is* $\sqrt{\vec{v} \cdot \vec{v}} = \|\vec{v}\| = \sqrt{v_1^2 + v_2^2 + ...v_n^2}$ *for* $\vec{v} = [v_1, v_2, ..., v_n]$.

Dot Product Theorems

Definition 1.3. *Two vectors* $\vec{v}, \vec{w} \in \mathbb{R}^n$ *are* ***orthogonal*** *(aka perpendicular) if* $\vec{v} \cdot \vec{w} = 0$.

Theorem 1.2 (Schwarz Inequality/Cauchy-Schwarz Inequality)**.** $|\vec{v} \cdot \vec{w}| \leq \|\vec{v}\| \|\vec{w}\|$

Theorem 1.3 (Triangle Inequality)**.** $\|\vec{v}\| + \|\vec{w}\| \geq \|\vec{v} + \vec{w}\|$

Matrices and Their Algebra

Definition 1.4. *A **m x n matrix** is an array of m rows, n columns.*
$M_{m,n}(\mathbb{R})$ is our notation for the set of all m x n matrices whose entities are real numbers.

To multiply matrices, $(AB)_{ij} = \sum_{l=1}^{n} a_{il}b_{lj}$

Definition 1.5. *Let $A = [a_{ij}] = M_{m,n}$. We define the **transpose** of A to be the matrix whose (ij)th entry is the (ji)th entry of A. The rows of A become the columns of A^T.*

Definition 1.6. *$I_n \in M_{nn}$, also called the **Identity Matrix**, is* $\begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \\ 0 & & 1 \end{bmatrix}$

Theorem 1.4 (Associated Matrix Properties)**.** $\boldsymbol{A+B=B+A, AB \neq BA, IA=AI=A}$**.**
Scalar Multiplication Properties
$r(\boldsymbol{A}+\boldsymbol{B}) = r\boldsymbol{A} + r\boldsymbol{B}$
$(r+s)\boldsymbol{A} = r\boldsymbol{A} + s\boldsymbol{A}$
$r(s\boldsymbol{A}) = (rs)\boldsymbol{A}$
$1\boldsymbol{A} = \boldsymbol{A}$

Transpose Properties
$(\boldsymbol{A}^T)^T = \boldsymbol{A}$
$(\boldsymbol{A}+\boldsymbol{B})^T = \boldsymbol{A}^T + \boldsymbol{B}^T$
$(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$

Solving Systems of Linear Equations

Definition 1.7. *A vector $\vec{s} \in \mathbb{R}^n$ is a solution to $\boldsymbol{A}\vec{x} = \vec{b}$ if $\boldsymbol{A}\vec{s} = \vec{b}$. There are 3 flavors of solutions: unique, infinite, or no solutions to a given equation of this form.*

*Linear systems $\boldsymbol{A}\vec{x} = \vec{b}$ that have one or more solutions are called **consistent systems** while linear systems that have no solutions are called **inconsistent systems**.*

*A key observation is that the equation $\boldsymbol{A}\vec{x} = \vec{b}$ is consistent if and only if $\vec{b}$ is in the **span**, or the set of all linear combinations, of the columns of $\boldsymbol{A}$.*

Definition 1.8. *A matrix $\boldsymbol{M}$ is in **row-echelon form** if:*
(1) any rows of $\boldsymbol{M}$ without pivots are below all rows of $\boldsymbol{M}$ with pivots
(2) a pivot in row i of $\boldsymbol{M}$ is in a column to the right of a pivot in row $i-1$ of $\boldsymbol{M}$.

Definition 1.9 (**Elementary Row Operations** of a Matrix)**.** *(1) interchange two rows*
(2) multiply all entries of a row by the same nonzero scalar
(3) replace a row by the sum of itself and a multiple of another row (i.e. add a multiple of one row to another)

Gaussian Elimination *is the process of using elementary row operations to transform a matrix into row-echelon form. Every matrix can be transformed into row echelon form with elementary row operations.*

Definition 1.10 (More on Elementary Row Operations)**.** *A matrix $\boldsymbol{M}$ is* ***row equivalent*** *to a matrix $\boldsymbol{N}$ if you can get from $\boldsymbol{M}$ to $\boldsymbol{N}$ by performing a sequence of elementary row operations. This is denoted $\boldsymbol{M} \sim \boldsymbol{N}$.*

Theorem 1.5. *If $[\boldsymbol{A}|\vec{b}] \sim [\boldsymbol{H}|\vec{c}]$ then the solutions to the system $\boldsymbol{A}\vec{x} = \vec{b}$ are the same solutions to $\boldsymbol{H}\vec{x} = \vec{c}$.*

Definition 1.11. *A matrix that is the result of performing one elementary row operation to an identity matrix $\boldsymbol{I}_m$ is called an* ***elementary matrix****.*

Definition 1.12. *A matrix is in* ***reduced row-echelon form (rref)*** *if it is in row echelon form AND every pivot is 1 and the pivot is the only nonzero entry in its column.*

The process of putting a matrix into rref is called the ***Gauss-Jordan Method****.*

Theorem 1.6. *Let $\boldsymbol{A}$ be a m x n matrix and $\boldsymbol{E}$ be an elementary matrix of size m x m. Then $\boldsymbol{EA}$ is the matrix obtained by transforming $\boldsymbol{A}$ by the same elementary row operations that was needed to create $\boldsymbol{E}$.*

Theorem 1.7 (Solutions of $\boldsymbol{A}\vec{x} = \vec{b}$)**.** *Let $\boldsymbol{A}\vec{x} = \vec{b}$ be a linear system with $[\boldsymbol{A}|\vec{b}] \sim [\boldsymbol{H}|\vec{c}]$ where $\boldsymbol{H}$ is in row-echelon form. Then*
(1) If $[\boldsymbol{H}|\vec{c}]$ has a pivot in the last column then $\boldsymbol{A}\vec{x} = \vec{b}$ is inconsistent.
(2) If $[\boldsymbol{H}|\vec{c}]$ does not have a pivot in the last column then (a) if every column of $\boldsymbol{H}$ has a pivot, then $\boldsymbol{A}\vec{x} = \vec{b}$ has a ***unique solution****, (b) if some columns of $\boldsymbol{H}$ have no pivot, then $\boldsymbol{A}\vec{x} = \vec{b}$ has an* ***infinite number of solutions.***

Inverses of Square Matrices

Definition 1.13. *A nxn matrix $\boldsymbol{A}$ is **invertible** (or nonsingular) if there is a nxn matrix $\boldsymbol{C}$ such that $\boldsymbol{CA} = \boldsymbol{AC} = \boldsymbol{I}$. We say that $\boldsymbol{C}$ is the **inverse** of $\boldsymbol{A}$, denoted as $\boldsymbol{A^{-1}}$,*

Theorem 1.8 (Unique Inverse)**.** *If $\boldsymbol{A}$ has an inverse, it is unique.*

Theorem 1.9. *Let $\boldsymbol{A}, \boldsymbol{B} \in M_n$ be invertible matrices. Then $\boldsymbol{AB}$ is invertible and $\boldsymbol{AB^{-1}} = \boldsymbol{B^{-1}A^{-1}}$.*

Theorem 1.10. *Let $\boldsymbol{A}, \boldsymbol{C} \in M_n$. Then if $\boldsymbol{AC} = \boldsymbol{I}$ then $\boldsymbol{CA} = \boldsymbol{I}$.*

Lemma 1.11. *Let $\boldsymbol{A} \in M_n$. If $\boldsymbol{A}\vec{x} = \vec{b}$ is consistent for any choice of $\vec{b} \in \mathbb{R}^n$, then $\boldsymbol{A} \sim \boldsymbol{I}$.*

Corollary 1.11.1. *If $\boldsymbol{A} \sim \boldsymbol{I}$, then $\boldsymbol{A}$ is invertible.*

Theorem 1.12. *If $\boldsymbol{A}$ is invertible, then $\boldsymbol{A}\vec{x} = \vec{b}$ is consistent for all $\vec{b}$.*

Theorem 1.13 (TFAE)**.** *Let $\boldsymbol{A} \in M_n$. The following are equivalent:*
(1) $\boldsymbol{A}$ is invertible.
(2) $\boldsymbol{A} \sim \boldsymbol{I}$
(3) $\boldsymbol{A}\vec{x} = \vec{b}$ is consistent for all $\vec{b} \in \mathbb{R}^n$.
(4) The span of the columns of $\boldsymbol{A}$ is $\mathbb{R}^n$ / the column space of A is $\mathbb{R}^n$ / the columns form a basis for $\mathbb{R}^n$.
(5) $\boldsymbol{A}$ is a product of elementary matrices.
(6) rank($\boldsymbol{A}$) = n.
(7) nullity($\boldsymbol{A}$) = 0.
(8) $\det(\boldsymbol{A}) = 0$,
(9) No eigenvalue of $\boldsymbol{A}$ is 0.

Homogeneous Linear Systems

Definition 1.14. *A **homogeneous linear system** is a linear system of the form $\boldsymbol{A}\vec{x} = 0$. These are always consistent (as a solution is the zero vector).*

Theorem 1.14. *If $\vec{u}, \vec{v}$ are solutions to a homogeneous linear system, so is $r\vec{u} + s\vec{v}$ for any $r, s \in \mathbb{R}$.*

Theorem 1.15. *Let $\vec{p}$ be a particular solution to $\boldsymbol{A}\vec{x} = \vec{b}$ and $\vec{s}$ be a solution to $\boldsymbol{A}\vec{x} = 0$. Then*
(1) $\vec{p} + \vec{s}$ is a solution to $\boldsymbol{A}\vec{x} = \vec{b}$.
(2) Every solution set to $\boldsymbol{A}\vec{x} = \vec{b}$ has the form $\vec{p} + \vec{h}$ where $\boldsymbol{A}\vec{h} = 0$.

Subspaces

Definition 1.15. *Let $\boldsymbol{W} \subseteq \mathbb{R}^n$. We say that $\boldsymbol{W}$ is closed under addition if $\vec{u}, \vec{v} \in \boldsymbol{W}$ then $\vec{u} + \vec{v} \in \boldsymbol{W}$. We say that $\boldsymbol{W}$ is closed under scalar multiplication if $\vec{u} \in \boldsymbol{W}, r \in \mathbb{R}$ then $r\vec{u} \in \boldsymbol{W}$.*

Definition 1.16 (Definition of a **Subspace**)**.** *Let $\boldsymbol{W} \subseteq \mathbb{R}^n$. We say that $\boldsymbol{W}$ is a **subspace** if it is closed under addition and scalar multiplication.*

Definition 1.17 (Important Subspaces)**.** *Let $\boldsymbol{A}$ be a m x n matrix. Some important subspaces associated with A. The **nullspace** of $\boldsymbol{A}$ is the set of solutions $\boldsymbol{A}\vec{x} = \vec{0}$. The **row space** of $\boldsymbol{A}$ is the span of the rows of $\boldsymbol{A}$. The **column space** of $\boldsymbol{A}$ is the span of columns of $\boldsymbol{A}$.*

Definition 1.18 (Basis of a Subspace)**.** *Let $\boldsymbol{W}$ be a subspace of $\mathbb{R}^n$. A subset $\vec{w}_1, \vec{w}_2, ..., \vec{w}_k$ is a **basis** for $\boldsymbol{W}$ if for all $\vec{w} \in \boldsymbol{W}$, there are unique scalars $r_1, r_2, ..., r_k$ such that $r_1\vec{w}_1 + r_2\vec{w}_2 + ... + r_k\vec{w}_k = \vec{w}$.*

A couple notes on a basis:
A basis of $\boldsymbol{W}$ is the smallest possible set that generates $\boldsymbol{W}$.
Let $\boldsymbol{A}$ be a matrix w/ column vectors $\vec{w}_1, \vec{w}_2, ..., \vec{w}_k$. Let $\boldsymbol{W}$ be the column space of $\boldsymbol{A}$. The column vectors form a basis for $\boldsymbol{W}$ iff $\boldsymbol{A}\vec{x} = \vec{b}$ has a unique solution for all $\vec{b} \in \boldsymbol{W}$.

Theorem 1.16. *$\boldsymbol{A} \in \boldsymbol{M_n}, \boldsymbol{A} \sim \boldsymbol{I} \iff \boldsymbol{A}\vec{x} = \vec{b}$ has a unique solution for all $\vec{b} \in \mathbb{R}^n$.*

Theorem 1.17. *$\vec{w}_1, ..., \vec{w}_k \subseteq \mathbb{R}^n$ is a basis for $\boldsymbol{W} = sp(\vec{w}_1, ..., \vec{w}_k)$ iff $r_1\vec{w}_1 + r_2\vec{w}_2 + ... + r_k\vec{w}_k = 0 \implies r_1 = r_2 = ... = r_k = 0$.*

Theorem 1.18. *A homogeneous linear system $\boldsymbol{A}\vec{x} = \vec{0}$ that has fewer equations (rows) than unknowns (columns) has an infinite number of solutions.*

Theorem 1.19. *Let $\boldsymbol{A}$ be a n x k matrix. TFAE:*
(1) The column vectors of $\boldsymbol{A}$ form a basis for the column space of $\boldsymbol{A}$.
(2) The rref of $\boldsymbol{A}$ is I_k followed by $(n-k)$ rows of zeros.
(3) Each consistent system $\boldsymbol{A}\vec{x} = \vec{b}$ has a unique solution.

2 Chapter 2: Dimension, Rank, and Linear Transformations

Independence and Dimension

Definition 2.1. *Let* $\{\vec{w}_1, ..., \vec{w}_k\} \subseteq \mathbb{R}^n$. *If the only linear combination of these vectors that gives* $\vec{0}$ *is the trivial one, we say that the set is* ***linearly independent****. If there is a nontrivial linear combination, the set is* ***linearly dependent****.*

Theorem 2.1 (New Definition of a Basis)**.** *Let* $\boldsymbol{W}$ *be a subspace of* $\mathbb{R}^n$. *A subset* $\{\vec{w}_1, ..., \vec{w}_k\} \subseteq \boldsymbol{W}$ *is a* ***basis*** *for* $\boldsymbol{W}$ *iff* $\{\vec{w}_1, ..., \vec{w}_k\}$ *is linearly independent and* $sp(\vec{w}_1, ..., \vec{w}_k) = \boldsymbol{W}$. *In other words, a basis of* $\boldsymbol{W}$ *is a linearly independent, spanning set of* $\boldsymbol{W}$.

Theorem 2.2. *Let* $\boldsymbol{W}$ *be a subspace of* $\mathbb{R}^n$. *Let* $\{\vec{w}_1, ..., \vec{w}_k\} \subseteq \boldsymbol{W}$ *be a spanning set of* $\boldsymbol{W}$. *Let* $\vec{v}_1, ...\vec{v}_m \subseteq \boldsymbol{W}$ *be a linearly independent set. Then* $k \geq m$.

In other words, no spanning set has fewer vectors than any linearly independent set.

Corollary 2.2.1. *Any 2 bases of* $\boldsymbol{W}$ *have the same size. Similarly, ever basis of* $\mathbb{R}^n$ *has* n *vectors.*

Definition 2.2. *Let* $\boldsymbol{W}$ *be a subspace of* $\mathbb{R}^n$. *The number of elements in a basis of* $\boldsymbol{W}$ *is called the* ***dimension*** *of* $\boldsymbol{W}$. *In particular,* $\dim(\mathbb{R}^n) = n$.

NOTE: How exactly do we find the basis for span$(\vec{w}_1, ..., \vec{w}_k)$*?*
(1) Form matrix $\boldsymbol{A}$ *whose columns are* $\vec{w}_1, ..., \vec{w}_k$.
(2) Use Gaussian Elimination on $\boldsymbol{A}$ *to transform to* $\boldsymbol{H}$, *a matrix in rref.*
(3) The set $\{\vec{w}_j\}$ *where the jth column of* $\boldsymbol{H}$ *has a pivot is a basis for span*$(\vec{w}_1, ..., \vec{w}_k)$.

Theorem 2.3. *Every subspace of* $\mathbb{R}^n$ *has a basis, and that basis has no more than* n *vectors.*

Theorem 2.4. *Every linearly independent set in a subspace* $\boldsymbol{W} \subseteq \mathbb{R}^n$ *can be enlarged (if necessary) to obtain a basis for* $\boldsymbol{W}$.

Theorem 2.5. *Let* $\boldsymbol{W}$ *be a subspace of* $\mathbb{R}^n$ *of dim*$(\boldsymbol{W}) = k$.
(1) every linearly independent set of k *vectors in* $\boldsymbol{W}$ *must span* $\boldsymbol{W}$.
(2) every spanning set of $\boldsymbol{W}$ *with* k *vectors is linearly independent.*

Finding Bases of Special Subspaces

Nullspaces

Every free variable corresponds to an element in the basis.
Example: if nullspace(A) $= [2r - s, r, s, 0]$ for $r, s \in \mathbb{R}$, then $[2r - s, r, s, 0] = r[2, 1, 0, 0] + s[-1, 0, 1, 0]$, and the basis is $\{[2, 1, 0, 0], [-1, 0, 1, 0]\}$.

Column Spaces

If $\boldsymbol{A} \sim \boldsymbol{H}$, where $\boldsymbol{H}$ is in row echelon form, then the columns of $\boldsymbol{A}$ that correspond to the columns of $\boldsymbol{H}$ with pivots form a basis for the column space.
The dimension of the column space is the number of columns in the rref of $\boldsymbol{A}$ with a pivot.

Row Spaces

If we have already transformed $\boldsymbol{A}$ into row echelon form $\boldsymbol{H}$, we can simply use nonzero rows of $\boldsymbol{H}$ as the basis for the row space of $\boldsymbol{A}$.

The Rank of a Matrix

Theorem 2.6 (Dimension of Row and Column Spaces)**.** *The dimension of the rowspace of $\boldsymbol{A}$ is the same as the dimension of the column space of $\boldsymbol{A}$, i.e.* $\dim(rowsp(\boldsymbol{A})) = \dim(colsp(\boldsymbol{A}))$

Definition 2.3. *Let $\boldsymbol{A}$ be a n x n matrix. The **rank** of $\boldsymbol{A}$ is the dimension of the column and rowspaces of $\boldsymbol{A}$. The **nullity** is the dimension of the nullspace of $\boldsymbol{A}$.*

Theorem 2.7 (**Rank-Nullity Theorem**)**.** *Let $\boldsymbol{A}$ be a m x n matrix. Then $rank(\boldsymbol{A}) + nullity(\boldsymbol{A}) = n$.*

Linear Transformations

Definition 2.4 (Domain and Codomain)**.** *For $f : X \to Y$, X is called the **domain** of f, and Y is called the **codomain** of f.*

Definition 2.5. *Let $H \subseteq X$. The **image** of H under f is the set $\{f(x) \mid h \in H\}$. This is denoted $f[H]$, or the set of elements in Y that are mapped to when the elements of H are plugged into f.*

Definition 2.6. *$f[X]$ is called the **range** of f (the set of all elements in Y that are mapped to by f).*

Definition 2.7. *Let $K \subseteq Y$. The **preimage** (or **inverse image**) of K is $\{x \in X \mid f(x) \in K\}$ and is denoted $f^{-1}[K]$ (the set of all elements of X that map to elements of K).*

Definition 2.8. *$T : \mathbb{R}^n \to \mathbb{R}^n$ is a **linear transformation** if for all $\vec{u}, \vec{v} \in \mathbb{R}^n, r \in \mathbb{R}$,*
(1) $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ (preserves structure of addition)
(2) $T(r\vec{u}) = rT(\vec{u})$ (preserves structure of scalar multiplication)

Theorem 2.8. *$f : \mathbb{R} \to \mathbb{R}$ is a linear transformation iff $f(x) = kx$ where $k \in \mathbb{R}$.*

Theorem 2.9. *Let $B = \{\vec{b}_1, ...\vec{b}_n\}$ be a basis for $\mathbb{R}^n$. Let $\vec{v} \in \mathbb{R}^n$. Then $T(\vec{v}) = r_1T(\vec{b}_1) + ... + r_nT(\vec{b}_n)$ where $r_1, ..., r_n$ are the unique scalars such that $\vec{v} = r_1\vec{b}_1 + ... + r_n\vec{b}_n$. This means $T(\vec{v})$ is determined by $T(\vec{b}_1), ..., T(\vec{b}_n)$.*

Corollary 2.9.1. *Let S, T be linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$. Let $\{\vec{b}_1, ..., \vec{b}_n\}$ be a basis for $\mathbb{R}^n$. If $S(\vec{b}_1) = T(\vec{b}_1)$ for all $i = 1, 2, ...n$ then $S = T$.*

Corollary 2.9.2 (Very Important Corollary)**.** *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Let $\boldsymbol{A}$ be the matrix whose ith column is $T(\vec{e}_i)$; i.e.*

$$\boldsymbol{A} = \begin{bmatrix} | & | & \cdots & | \\ T(\vec{e}_1) & T(\vec{e}_2) & \cdots & T(\vec{e}_n) \\ | & | & \cdots & | \end{bmatrix}$$

which is a m x n matrix. Then, for any $\vec{v} \in \mathbb{R}^n, T(\vec{v}) = \boldsymbol{A}\vec{v}$. Note that all linear transformations are matrix transformations.

The matrix $\boldsymbol{A}$ represented above is called the **standard matrix representation of $\boldsymbol{A}$**.

Definition 2.9 (Important Relations between a Linear Transformation $\boldsymbol{T}$ and the Standard Matrix Representation of $\boldsymbol{A}$)**.**
*(1) The **column space** of $\boldsymbol{A}$, $\{\boldsymbol{A}\vec{x} \mid \vec{x} \in \mathbb{R}^n\}$, is the **range** of $\boldsymbol{T}$, $\{\boldsymbol{T(x)} \mid \vec{x} \in \mathbb{R}^n\}$.*
*(2) The **nullspace** of $\boldsymbol{A}$, $\{\vec{x} \in \mathbb{R}^n \mid \boldsymbol{A}\vec{x} = \vec{0}\}$, is the **kernel** of $\boldsymbol{T}$, $\{\vec{x} \in \mathbb{R}^n \mid \boldsymbol{T(\vec{x})} = \vec{0}\}$.*
*(3) The **rank** of $\boldsymbol{A}$ is the **rank** of $\boldsymbol{T}$.*
*(4) The **nullity** of $\boldsymbol{A}$ is the **nullity** of $\boldsymbol{T}$.*

Theorem 2.10. *For $\boldsymbol{T} : \mathbb{R}^n \to \mathbb{R}^m$, then $rank(\boldsymbol{T}) + nullity(\boldsymbol{T}) = n$.*

Compositions of Linear Transformations

Definition 2.10. *For $T : \mathbb{R}^n \to \mathbb{R}^m, S : \mathbb{R}^m \to \mathbb{R}^k$, the **composition** of S and T is $S \circ T : \mathbb{R}^n \to \mathbb{R}^k$ is defined as $(S \circ T)\vec{x} = S(T(\vec{x}))$.*

Definition 2.11. *If the standard matrix representation $\boldsymbol{A}$ of a linear transformation T is invertible, and S is the linear transformation associated with $\boldsymbol{A}^{-1}$, $S \circ T = T \circ S =$ **identity function**.*

*We say that T has an inverse function S and denote S as T^{-1}. T is an **invertible linear transformation** (which corresponds to invertible matrices).*

Theorem 2.11. *Let $\boldsymbol{T} : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then*
*(1) If $\boldsymbol{W}$ is a subspace of $\mathbb{R}^n$, then **T[w]** is a subspace of $\mathbb{R}^n$.*
(2) If $\boldsymbol{u}$ is a subspace of $\mathbb{R}^m$, then $\boldsymbol{T}^{-1}[\boldsymbol{u}]$ is a subspace of $\mathbb{R}^n$.

3 Chapter 3: Vector Spaces

Preliminary Definitions and Properties of Vector Spaces

Definition 3.1. *Let $\boldsymbol{V}$ be a set of objects (called vectors) that is closed under addition and scalar multiplication. We say that $\boldsymbol{V}$ is a* ***vector space*** *if the following 8 properties hod for all $\vec{u}, \vec{v}, \vec{w} \in \boldsymbol{V}$ and scalars r, s:*

(A1, Associativity of +): $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
(A2, Commutativity of +): $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
(A3, Additive Identity): $\exists \vec{z} \in \boldsymbol{V}$ such that $\vec{z} + \vec{v} = \vec{v}$
(A4, Additive Inverse): For all $\vec{v}, \exists \vec{w} \in \boldsymbol{V}$ such that $\vec{v} + \vec{w} = \vec{0}$.

(S1, Distributivity): $r(\vec{v} + \vec{w}) = r\vec{v} + r\vec{w}$
(S2, Distributivity): $(r + s)\vec{v} = r\vec{v} + s\vec{v}$
(S3, Associativity): $r(s\vec{v}) = (rs)\vec{v}$
(S4, Preservation of Scalar): $1\vec{v} = \vec{v}$

Theorem 3.1. *Let $\boldsymbol{V}$ be a vector space. The following are true:*
(1, Uniqueness of $\vec{0}$) If $\vec{x} + \vec{v} = \vec{v}$ for all $\vec{v} \in \boldsymbol{V}$, then $\vec{x} = 0$.
(2, Uniqueness of Additive Inverse) For every $\vec{v} \in \boldsymbol{V}$, there is only one vector $\vec{x}$ with $\vec{x} + \vec{v} = 0$.
(3, Cancellation Property) If $\vec{u} + \vec{v} = \vec{u} + \vec{w}$, then $\vec{v} = \vec{w}$.
(4) $0\vec{v} = \vec{0}$ for any $\vec{v} \in \boldsymbol{V}$.
(5) $r\vec{0} = \vec{0}$ for any $r \in \mathbb{R}$
(6) $(-1)\vec{v}$ is the additive inverse of $\vec{v}, -\vec{v}$.

Basic Concepts of Vector Spaces

Definition 3.2. *Let $\boldsymbol{V}$ be a vector space with vectors $\vec{v}_1, ...\vec{v}_k$. A* ***linear combination*** *of these vectors is any vector of the form $\vec{v} = r_1\vec{v}_1 + ... + r_k\vec{v}_k$ for scalars $r_1, r_2, ...r_k$.*

Definition 3.3. *Let $\boldsymbol{V}$ be a vector space and let $\boldsymbol{X} \subseteq \boldsymbol{V}$. The* ***span of X*** *(denoted sp($\boldsymbol{X}$)) is the set of all finite linear combinations of elements of $\boldsymbol{X}$. If $\boldsymbol{W} = sp(\boldsymbol{X})$, we say that $\boldsymbol{X}$ generates $\boldsymbol{W}$.*

Definition 3.4. *If a vector space $\boldsymbol{V} = sp(\boldsymbol{X})$ for some finite subset $\boldsymbol{X}$, then $\boldsymbol{V}$ is* ***finitely generated****. Otherwise if no finite subset exists, $\boldsymbol{V}$ is* ***infinitely generated****.*

Definition 3.5. *A subset $\boldsymbol{W}$ of vector space $\boldsymbol{V}$ is a* ***subspace*** *of $\boldsymbol{V}$ if, using the same definitions of addition and scalar multiplication as $\boldsymbol{V}, \boldsymbol{W}$ is itself a vector space. All vector spaces $\boldsymbol{V}$ have $\{\vec{0}\}$ and $\boldsymbol{V}$ as subspaces. $\{\vec{0}\}$ is the* ***trivial subspace****. All other subspaces of $\boldsymbol{V}$ are called* ***proper subspaces****.*

Theorem 3.2. *A nonempty subset $\boldsymbol{W}$ of a vector space $\boldsymbol{V}$ is a subspace of $\boldsymbol{V}$ if it is closed under addition/scalar multiplication.*

Corollary 3.2.1. *Let $\boldsymbol{V}$ be a vector space and suppose $\boldsymbol{X}$ is a subset of $\boldsymbol{V}$. Then $sp(\boldsymbol{X})$ is a subspace of $\boldsymbol{V}$.*

Definition 3.6. *Let $\boldsymbol{V}$ be a vector space and suppose $\boldsymbol{X}$ is a subset of $\boldsymbol{V}$. A **dependence relation** in $\boldsymbol{X}$ is an equation $r_1\vec{v}_1 + r_2\vec{v}_2 + ... + r_k\vec{v}_k = 0$ where $\{\vec{v}_i\} \subseteq \boldsymbol{X}$ and some $r_i \neq 0$. If $\exists$ a dependence relation, then $\boldsymbol{X}$ is **linearly dependent**; if $\nexists$ dependence relation, then $\boldsymbol{X}$ is **linearly independent**.*

Definition 3.7. *A **basis for a vector space $\boldsymbol{V}$** is a set $\boldsymbol{X}$ such that (1) $sp(\boldsymbol{X}) = \boldsymbol{V}$ and (2) $\boldsymbol{X}$ is linearly independent.*

Theorem 3.3. *A subset $\boldsymbol{X}$ of a vector space $\boldsymbol{V}$ is a **basis for $\boldsymbol{V}$** iff every element in $\boldsymbol{V}$ can be uniquely expressed as a linear combination of elements of $\boldsymbol{X}$.*

Theorem 3.4. *Let $\boldsymbol{V}$ be a vector space. Let $\{\vec{w}_1, ...\vec{w}_k\} \subseteq \boldsymbol{V}$ with $sp(\vec{w}_1, ...\vec{w}_k) = \boldsymbol{V}$. Let $\{\vec{v}_1, ...\vec{v}_m\}$ be a linearly independent set. Then $k \geq m$.*

Corollary 3.4.1. *Any two bases of a vector space $\boldsymbol{V}$ have the same size/dimension.*

Definition 3.8. *The **dimension** of a vector space $\boldsymbol{V}$ is the number of elements in the bases.*

Theorem 3.5. *Any linearly independent subset of vector space $\boldsymbol{V}$ can be enlarged to create a basis for $\boldsymbol{V}$; any spanning set of $\boldsymbol{V}$ can be also be trimmed to create a basis.*

Theorem 3.6. *Let $\boldsymbol{V}$ be a vector space of dimension k. Then*
(1) any linearly independent subset of $\boldsymbol{V}$ with k elements is a basis for $\boldsymbol{V}$.
(2) any subset of k elements of $\boldsymbol{V}$ that spans $\boldsymbol{V}$ is a basis for $\boldsymbol{V}$.

Theorem 3.7. *Let $\boldsymbol{V}$ be finitely generated (i.e. $\exists$ finite set $\boldsymbol{X}$ with $sp(\boldsymbol{X}) = \boldsymbol{V}$). Then $\boldsymbol{V}$ has a basis.*

Theorem 3.8. *Every vector space has a basis.*

Theorem 3.9 (Axiom of Choice)**.** *Given a collection of disjoint nonempty sets, there is a set (called the "choice set") that contains exactly 1 element from each set in the collection.*

Theorem 3.10 (Banach-Tarski Paradox)**.** *Given a solid ball, there is a way to cut it into a finite number of pieces and reassemble into 2 solid balls of the same size.*

Coordinization of Vector Spaces

Definition 3.9. *Let $\boldsymbol{V}$ be a finite dimensional vector space. An* ***ordered basis*** *of $\boldsymbol{V}$ is a basis of $\boldsymbol{V}$ with a fixed order, denoted by $(\vec{b}_1, \vec{b}_2, ..., \vec{b}_n)$.*
Ex: The ***standard ordered basis*** *of $\mathbb{R}^n : (\vec{e}_1, ..., \vec{e}_{n)}$.*

Definition 3.10 (Coordinate Vector relative to an Ordered Basis)**.** *Let $\boldsymbol{V}$ be a vector space with ordered basis $\boldsymbol{B} = (\vec{b}_1, ...\vec{b}_n)$. Let $\vec{v} \in \boldsymbol{V}$. The* ***coordinate vector of $\vec{v}$ relative to $\boldsymbol{V}$*** *is the vector $\vec{v}_{\boldsymbol{B}} = [r_1, ...r_n] \in \mathbb{R}^n$, where $\vec{v} = r_1\vec{b}_1 + r_2\vec{b}_2 + ... + r_n\vec{b}_n$.*

Definition 3.11. *We define a* ***map*** *$\boldsymbol{T_B} : \boldsymbol{V} \to \mathbb{R}^n$ with $\boldsymbol{T_B}(\vec{v}) = \vec{v}_{\boldsymbol{B}}$.*
All the important features of $\boldsymbol{V}$ are preserved by $\boldsymbol{T_B}$ (ex. linear independence is preserved).

Definition 3.12. *Let $\boldsymbol{V}, \boldsymbol{W}$ be vector spaces. A map $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{W}$ is a* ***linear transformation*** *if*
(1) For all $\vec{u}, \vec{v} \in \boldsymbol{V}, \boldsymbol{T}(\vec{u} + \vec{v}) = \boldsymbol{T}(\vec{u}) + \boldsymbol{T}(\vec{v})$.
(2) For all $\vec{u} \in \boldsymbol{V}, r \in \mathbb{R}, \boldsymbol{T}(r\vec{v}) = r\boldsymbol{T}(\vec{v})$.

Theorem 3.11. *Let $\boldsymbol{B} : (\vec{b}_1, ...\vec{b}_n)$ be an ordered basis for a vector space $\boldsymbol{V}$. Then the map $\boldsymbol{T_B} : \boldsymbol{V} \to \mathbb{R}^n$ is a linear transformation.*

Lemma 1: $\boldsymbol{T}(\vec{0}_V) = \boldsymbol{T}(0 * \vec{0}_V) = 0\boldsymbol{T}(\vec{0}_V) = \vec{0}_W$, for $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{W}$, a linear transformation.

Lemma 2: $ker(\boldsymbol{T_B}) = \{\vec{0}_V\}$ (i.e. if $\boldsymbol{T_B}(\vec{x}) = \vec{0}$ then $\vec{x} = \vec{0}_V$).

Theorem 3.12. *Let $\boldsymbol{V}$ be a vector space with ordered basis $\boldsymbol{B}$ with n elements. Then $\{\vec{v}_1, ..., \vec{v}_k\}$ linearly independent in $\boldsymbol{V} \iff \{\boldsymbol{T_B}(\vec{v}), ...\boldsymbol{T_B}(\vec{v}_k)\}$ linearly independent in $\mathbb{R}^n$.*

Linear Transformations for Vector Space

Definition 3.13. *Let $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{W}$ be a linear transformation of vector spaces $\boldsymbol{V}, \boldsymbol{W}$. The* ***kernel of $\boldsymbol{T}$*** *is the set of vectors $\vec{v} \in \boldsymbol{V}$ that map to $\vec{0}_{\boldsymbol{W}}$, i.e. $ker(\boldsymbol{T}) = \boldsymbol{T}^{-1}[\{\vec{0}_{\boldsymbol{W}}\}] = \{\vec{v} \in \boldsymbol{V} \mid \boldsymbol{T}(\vec{v}) = \vec{0}_{\boldsymbol{W}}\}$.*

Theorem 3.13. *$\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{W}$, a linear transformation of vector spaces. Then:*
(1: Zero maps to zero) $\boldsymbol{T}(\vec{0}_{\boldsymbol{V}}) = \vec{0}_{\boldsymbol{W}}$
(2: Subspaces map to subspaces) If $\boldsymbol{U}$ is a subspace of $\boldsymbol{V}$, then $\boldsymbol{T}[\boldsymbol{U}]$ is a subspace of $\boldsymbol{W}$.
(3: Inverse Images of Subspaces are Subspaces) If $\boldsymbol{X}$ is a subspace of $\boldsymbol{W}$, then $\boldsymbol{T}^{-1}[\boldsymbol{X}]$ is a subspace of $\boldsymbol{V}$.
(4: Linear Transformations are uniquely determined by what they do to a basis): Let $\boldsymbol{B}$ be a basis for a vector space $\boldsymbol{V}$. If $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{V'}$ and $\boldsymbol{S} : \boldsymbol{V} \to \boldsymbol{V'}$ are linear transformations such that $\boldsymbol{T}(\vec{b}) = \boldsymbol{S}(\vec{b})$ for all $\vec{b} \in \boldsymbol{B}$, then $\boldsymbol{S} = \boldsymbol{T}$.

Definition 3.14 (Injective Function). *$f : \boldsymbol{X} \to \boldsymbol{Y}$ is **one-to-one** (aka **injective**) if $f(x_1) = f(x_2) \implies x_1 = x_2$ i.e. different elements must map to different places.*

Definition 3.15 (Surjective Function). *$f : \boldsymbol{X} \to \boldsymbol{Y}$ is **onto** (aka **surjective**) if for all $y \in Y$ there is an $x \in X$ with $f(x) = y$ i.e. $range(f) = Y; range(f) =$ codomain; every element in Y is mapped to.*

Definition 3.16 (Bijective Function). *$f : \boldsymbol{X} \to \boldsymbol{Y}$ is **bijective** (aka a **bijection**) if it is both one-to-one/onto (or both injective and surjective).*

Definition 3.17 (Invertible Function). *$f : \boldsymbol{X} \to \boldsymbol{Y}$ is **invertible** if there is a function $g : Y \to X$ with $(g \circ f)(x) = x$ for all $x \in Y$ and $(f \circ g)(y) = y$ for all $y \in Y$. Denote y by f^{-1}.*

Theorem 3.14. *A function $f : \boldsymbol{X} \to \boldsymbol{Y}$ is invertible $\iff$ f is bijective.*

Corollary 3.14.1. *Let $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{W}$, a linear transformation of vector spaces. Then T bijective $\iff$ there is a linear transformation $\boldsymbol{S}$ such that $\boldsymbol{S} = \boldsymbol{T}^{-1}$ i.e. if $\boldsymbol{T}$ is invertible, then $\boldsymbol{T}^{-1}$ is a linear transformation.*

Definition 3.18. *A function $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{W}$ between vector spaces V and W is an **isomorphism** iff $\boldsymbol{T}$ is a linear transformation and a bijection.*

Definition 3.19. *$\boldsymbol{V}$ and $\boldsymbol{W}$ are **isomorphic vector spaces** if there is an isomorphism $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{W}$. Write $\boldsymbol{V} \cong \boldsymbol{W}$// isomorphic means "essentially the same"*

Corollary 3.14.2. *A linear transformation $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{W}$ is an isomorphism iff $\boldsymbol{T}$ is invertible.*

Theorem 3.15. *Let $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{W}$ be a linear transformation. $\boldsymbol{T}$ is one-to-one $\iff$ $ker(\boldsymbol{T}) = \{\vec{0}\}$.*

Theorem 3.16. *Let $\boldsymbol{V}$ be a vector space of dimension n and let $\boldsymbol{B}$ be an ordered basis for $\boldsymbol{V}$. The coordinization map $\boldsymbol{T_B} : \boldsymbol{V} \to \mathbb{R}^n$ is defined by $\boldsymbol{T_B}(\vec{v}) = \vec{v}_{\boldsymbol{B}}$ is an isomorphism.*

Corollary 3.16.1. *If $\dim(\boldsymbol{V}) = n, V \cong \mathbb{R}^n$ (isomorphic to $\mathbb{R}^n$).*

Definition 3.20. *Let $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{V'}$ be a linear transformation, $\boldsymbol{B} = (\vec{b}_1, ...\vec{b}_n)$ be an ordered basis of $\boldsymbol{V}$, and $\boldsymbol{B'}$ be an ordered basis of $\boldsymbol{V'}$. Then the matrix $\boldsymbol{A}$ given by*

$$\boldsymbol{A} = \begin{bmatrix} | & | & \cdots & | \\ \boldsymbol{T}(\vec{b}_1)_{\boldsymbol{B'}} & \boldsymbol{T}(\vec{b}_2)_{\boldsymbol{B'}} & \cdots & \boldsymbol{T}(\vec{b}_n)_{\boldsymbol{B'}} \\ | & | & \cdots & | \end{bmatrix}$$

*is called the **matrix representation of $\boldsymbol{T}$ relative to $\boldsymbol{B}, \boldsymbol{B'}$**.*

Theorem 3.17. *Let $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{V'}$ be a linear transformation, $\boldsymbol{B}$ be an ordered basis of $\boldsymbol{B}$, $\boldsymbol{B'}$ be an ordered basis of $\boldsymbol{V'}$, and $\boldsymbol{A}$ be the matrix representation of $\boldsymbol{T}$ with respect to $\boldsymbol{B}, \boldsymbol{B'}$. Then $\boldsymbol{T}$ is invertible $\iff$ $\boldsymbol{A}$ is invertible.*

<u>**Inner Product Spaces**</u> An **inner product** is a generalization of the dot product.

Definition 3.21. *Let $\boldsymbol{V}$ be a vector space with scalars from $\mathbb{R}$. An **inner product on $\boldsymbol{V}$** is a map from an ordered pair of vectors $\vec{v}, \vec{w}$ to $\mathbb{R}$, denoted $\langle \vec{v}, \vec{w} \rangle$, such that for all $\vec{u}, \vec{v}, \vec{w} \in \boldsymbol{V}$ and $r \in \mathbb{R}$,*
(1, Symmetry) $\langle \vec{v}, \vec{w} \rangle = \langle \vec{w}, \vec{v} \rangle$
(2, Additivity/Distributivity) $\langle \vec{u}, \vec{v} + \vec{w} \rangle = \langle \vec{u}, \vec{v} \rangle + \langle \vec{u}, \vec{w} \rangle$
(3, Homogeneity) $r\langle \vec{v}, \vec{w} \rangle = \langle r\vec{v}, \vec{w} \rangle = \langle \vec{v}, r\vec{w} \rangle$
(4, Positivity) $\langle \vec{v}, \vec{v} \rangle \geq 0,$ *and* $\langle \vec{v}, \vec{v} \rangle = 0 \iff \vec{v} = \vec{0}$

Definition 3.22. *A vector space with scalars from $\mathbb{R}$ that has an inner product is called **a (real) inner product space**.*

Definition 3.23. *Let $\boldsymbol{V}$ be an inner product space. We define*
*the **norm** (aka **length**) of $\vec{v} \in \boldsymbol{V}$ as* $\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle}$
*the **distance** between two vectors $\vec{x}, \vec{y}$ in $\boldsymbol{V}$ as* $\big\|(\vec{x} - \vec{y})\big\|$
*the **angle** between $\vec{x}, \vec{y}$ in $\boldsymbol{V}$ as* $\arccos \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\|\|\vec{y}\|}$.

Theorem 3.18. *Let $\boldsymbol{V}$ be an inner product space. Then* $\vec{x} \perp \vec{y} \iff \langle \vec{x}, \vec{y} \rangle = 0$.

Additionally, theorems proved about the dot product apply to inner product spaces.

4 Chapter 4: Determinants

Preliminary Definitions: Every n x n matrix $\boldsymbol{A}$ has a number associated with it called its **determinant**, denoted $\det(\boldsymbol{A})$. We define the $\det(\boldsymbol{A})$ recursively, i.e. the def of $\det(\boldsymbol{A})$ will involve determinants of smaller matrices.

Definition 4.1. *Let* $\boldsymbol{A} = [a_{1,1}]$. *a 1 x 1 matrix.* $\det(\boldsymbol{A}) = a_{1,1}$.

Definition 4.2. *Let* $\boldsymbol{A}$ *be a n x n matrix. The matrix obtained by deleting the* i*th row and* j*th column of* $\boldsymbol{A}$ *is called a* ***minor*** *of* $\boldsymbol{A}$ *and is denoted* $\boldsymbol{A}_{ij}$.

Definition 4.3. *Let* $\boldsymbol{A} = [a_{ij}]$ *be an nxn matrix. The* ***cofactor*** *of* a_{ij} *is* $(-1)^{i+j}\det(\boldsymbol{A}_{ij})$ *and is denoted* a'_{ij}.

Definition 4.4. *Let* $\boldsymbol{A} = [a_{ij}]$ *be a nxn matrix, with* $n > 1$. *The* ***determinant*** *of* $\boldsymbol{A}$, *denoted* $\det(\boldsymbol{A})$, *is*

$$\sum_{j=1}^{n} a_{1j}a'_{1j} + ... + a_{1n}a'_{1n} = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(\boldsymbol{A}_{ij})$$

Sarrus' Rule/Mnemonic *can also be used to find determinants of* $3x3$ *matrices.*

Theorem 4.1 (General Expansion by Minors)**.** *Let* $\boldsymbol{A} \in \boldsymbol{M}_n$ *and* $r, s \in \{1, 2, ..., n\}$.
Then $\det(\boldsymbol{A}) = a_{r1}a'_{r1} + a_{r2}a'_{r2} + ... + a_{rn}a'_{rn} = a_{1s}a'_{1s} + ... + a_{ns}a'_{ns}$,
i.e. the determinant of $\boldsymbol{A}$ *is equal to an expansion along row* r *or column* s.

Corollary 4.1.1. $\boldsymbol{A} \in \boldsymbol{M}_n$. *If* $\boldsymbol{A}$ *has an entire row/column with only* 0 *entries,* $\det(\boldsymbol{A}) = 0$.

Definition 4.5. *Let* $\boldsymbol{A} = [a_{ij}] \in \boldsymbol{M}_n$. *We say that* $\boldsymbol{A}$ *is* ***upper triangular*** *if* $a_{ij} = 0$ *when* $i > j$. *We say that* $\boldsymbol{A}$ *is* ***lower triangular*** *if* $a_{ij} = 0$ *when* $i < j$.

Theorem 4.2. *Let* $\boldsymbol{A} \in \boldsymbol{M}_n$ *be upper/lower triangular. Then*

$$\det(\boldsymbol{A}) = \prod_{i=1}^{n} a_{ii} = a_{11}a_{22}...a_{nn}$$

Theorem 4.3. *Let* $\boldsymbol{A} \in \boldsymbol{M}_n$. *Then* $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^T)$.

What effect does performing an elementary row operation have on the determinant?

Theorem 4.4. *Let* $\boldsymbol{A} \in \boldsymbol{M}_n$. *Let* $\boldsymbol{B}$ *be the matrix obtained by switching 2 rows of* $\boldsymbol{A}$. *Then* $\det(\boldsymbol{B}) = -\det(\boldsymbol{A})$.

Corollary 4.4.1. *Let* $\boldsymbol{A} \in \boldsymbol{M}_n$ *have two equal rows. Then* $\det(\boldsymbol{A}) = 0$.

Theorem 4.5. *Let* $\boldsymbol{A} \in \boldsymbol{M}_n$. *Let* $\boldsymbol{B}$ *be the matrix obtained by multiplying row* k *of* $\boldsymbol{A}$ *by* r. *Then* $\det(\boldsymbol{B}) = r(\det(\boldsymbol{A}))$.

Theorem 4.6. *Let $\boldsymbol{A} \in \boldsymbol{M}_n$. Let $\boldsymbol{B}$ be the matrix obtained by performing the row operation $R_j \to r * R_i + R_j$. Then* $\det(\boldsymbol{B}) = \det(\boldsymbol{A})$.

Theorem 4.7. $\boldsymbol{A}$ *invertible* $\iff \det(\boldsymbol{A}) \neq 0$

(In Summary): Effects of 3 "flavors" of elementary row operations

(1) Switching 2 rows $\to$ multiply by -1
(2) Multiplying a row by a nonzero scalar $r \to$ multiply by r
(3) Adding a multiple of one row to another $\to$ no effect

Determinants and their Geometry

- **Area of a parallelogram** determined by $[a, b]$ and $[c, d]$ is

$$\boldsymbol{A} = \left| \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right| = ad - bc$$

- **Volume of a parallelepiped** determined by $\vec{a} = [a_1, a_2, a_3], \vec{b} = [b_1, b_2, b_3]$, and $\vec{c} = [c_1, c_2, c_3]$ is

$$\boldsymbol{V} = \left| \det \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \right|$$

- **Volume of a n-box** determined by $\vec{a}_1, ...\vec{a}_n$ with $\vec{a}_i = [a_{i1}, a_{i2}, ...a_{in}]$ is

$$\boldsymbol{V} = \left| \det \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \right|$$

- **Cross Product** of $\vec{a} = [a_1, a_2, a_3]$ and $\vec{b} = [b_1, b_2, b_3]$:

$$\vec{a} \times \vec{b} = \det \begin{bmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = \det \begin{bmatrix} a_2 & a_3 \\ b_2 & b_3 \end{bmatrix} \vec{i} - \det \begin{bmatrix} a_1 & a_3 \\ b_1 & b_3 \end{bmatrix} \vec{j} + \det \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \vec{k}$$

$$= \|\vec{a}\| \|\vec{b}\| \sin(\theta) \vec{n}$$

- **Area of parallelogram** denoted by $\vec{a}, \vec{b} \in \mathbb{R}^3$ is $\boldsymbol{A} = \left\| \vec{a} \times \vec{b} \right\|$.

Notes on Cross Product

- $\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$
- $(\vec{a} \times \vec{b}) \times \vec{c} \neq \vec{a} \times (\vec{b} \times \vec{c})$
- $\mathbb{R}^3$ under cross product forms a structure called a **Lie Algebra**

Theorem 4.8. *Let* $\boldsymbol{A} \in \boldsymbol{M}_n$, *then* $\det(\boldsymbol{A}) = \det(\boldsymbol{A}^T)$.

Theorem 4.9. *Let* $\boldsymbol{A} \in \boldsymbol{M}_n, n \geq 2$. *Let* $\boldsymbol{B}$ *be a matrix obtained by swapping* 2 *rows of* $\boldsymbol{A}$. *Then* $\det(\boldsymbol{B}) = -\det(\boldsymbol{A})$.

Theorem 4.10. *Let* $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{M}_n$. $\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$.

Lemma 4.11. *Consider* $\boldsymbol{EA}$, *where* $\boldsymbol{E}, \boldsymbol{A} \in \boldsymbol{M}_n$ *and* $\boldsymbol{E}$ *is an elementary matrix. Then* $\det(\boldsymbol{EA}) = \det(\boldsymbol{E})\det(\boldsymbol{A})$

5 Chapter 5: Eigenvalues

Preliminary Definitions

Definition 5.1. *Let $\boldsymbol{A} \in \boldsymbol{M}_n$. A scalar λ is said to be an* ***eigenvalue*** *of $\boldsymbol{A}$ if there is a nonzero vector $\vec{x}$ such that $\boldsymbol{A}\vec{x} = \lambda\vec{x}$, The vector $\vec{x}$ is called an* ***eigenvector*** *of $\boldsymbol{A}$ that corresponds to λ.*

$\boldsymbol{A}\vec{x} = \lambda\vec{x}$ means $\vec{x}$ gets mapped to a scalar multiple of itself by $\boldsymbol{A}$ // $\vec{x}$ is scaled by a factor of λ by $\boldsymbol{A}$ // $\boldsymbol{A}\vec{x}$ is in span$(\vec{X})$.

Definition 5.2. *If λ is an eigenvector of $\boldsymbol{A}$, we call nullspace$(\boldsymbol{A} - \lambda\boldsymbol{I})$ the* ***eigenspace*** *corresponding to λ, denoted $\boldsymbol{E}_\lambda$.*

Definition 5.3. *If $\boldsymbol{A} \in \boldsymbol{M}_n, \det(\boldsymbol{A} - \lambda\boldsymbol{I})$ is a polynomial of degree n in the variable λ called the* ***characteristic polynomial*** *of $\boldsymbol{A}$.*

Theorem 5.1. *$\boldsymbol{A} \in \boldsymbol{M}_n$. Suppose $\vec{v}$ is an eigenvector $\boldsymbol{A}$ with eigenvalue λ. Then*
(1) $\vec{v}$ is an eigenvector of $\boldsymbol{A}^r$
(2) If $\boldsymbol{A}$ is invertible, $\vec{v}$ is an eigenvector of $\boldsymbol{A}^{-1}$ with eigenvalue $\frac{1}{\lambda}$.
(3) $\boldsymbol{E}_\lambda$ is a subspace of $\mathbb{R}^n$.

Theorem 5.2 (Extension of TFAE)**.** *$\boldsymbol{A}$ invertible $\iff$ no eigenvalue of $\boldsymbol{A}$ is* 0.

Definition 5.4. *Let $\boldsymbol{V}$ be a vector space and $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{V}$ be a linear transformation. A scalar λ is said to be an* ***eigenvalue*** *of $\boldsymbol{T}$ if there is a nonzero vector $\vec{v} \in \boldsymbol{V}$ such that $\boldsymbol{T}(\vec{v}) = \lambda\vec{v}$. $\vec{v}$ is called an* ***eigenvector*** *of $\boldsymbol{T}$ corresponding to λ.*

Theorem 5.3. *$\boldsymbol{A} \in \boldsymbol{M}_n$. Suppose $\lambda_1, \lambda_2, ...\lambda_n$ are distinct eigenvalues of $\boldsymbol{A}$. Let $\vec{x}_i$ be an eigenvector corresponding to λ_i. Then $\{\vec{x}_1, ..., \vec{x}_n\}$ are linearly independent.*

Diagonalization

Definition 5.5. *Let $\boldsymbol{P}, \boldsymbol{Q} \in \boldsymbol{M}_n$. We say that $\boldsymbol{P}$ is similar to $\boldsymbol{Q}$ if $\exists$ an invertible matrix $\boldsymbol{C} \in \boldsymbol{M}_n$ such that $\boldsymbol{C}^{-1}\boldsymbol{P}\boldsymbol{C} = \boldsymbol{Q}$.*
A note is that if $\boldsymbol{P}$ is similar to $\boldsymbol{Q} \implies \boldsymbol{Q}$ similar to $\boldsymbol{P}$. Additionally, similar matrices have the same rank, nullity, determinant, and have (potentially) different eigenvalues.

Definition 5.6. *An nxn matrix $\boldsymbol{A} : [a_{ij}]$ is* ***diagonal*** *if $a_{ij} = 0$ whenever $i \neq j$ (the entries not on the diagonal are* 0*).*

Definition 5.7. *A matrix $\boldsymbol{A} \in \boldsymbol{M}_n$ is* ***diagonalizable*** *if it is similar to a diagonal matrix; i.e. if $\exists C \in \boldsymbol{M}_n$ such that $\boldsymbol{C}^{-1}\boldsymbol{A}\boldsymbol{C} = \boldsymbol{D}$ where $\boldsymbol{D}$ is diagonal.*

Definition 5.8. *Let $\boldsymbol{A} \in \boldsymbol{M}_n$ and suppose $\hat{\lambda}$ is an eigenvalue of $\boldsymbol{A}$. The* ***algebraic multiplicity*** *of $\hat{\lambda}$ is the number of factors of the form $(\hat{\lambda} - \lambda)$ that appear in the factorization of the characteristic polynomial of $\boldsymbol{A}, p(\lambda) = det(\boldsymbol{A} - \lambda \boldsymbol{I})$.*

The ***geometric multiplicity*** *of $\hat{\lambda}$ is the dimension of the eigenspace $\boldsymbol{E}_{\hat{\lambda}}$ of $\boldsymbol{A}$.*

Theorem 5.4 (Jordan Canonical Form)**.** *Let $\boldsymbol{A} \in \boldsymbol{M}_n$. Then $\boldsymbol{A}$ is similar to a matrix $\boldsymbol{J}$ of the form*

$$\boldsymbol{J} = \begin{bmatrix} \lambda_1 & * & \cdots & \cdots & 0 \\ 0 & \lambda_2 & * & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \lambda_{n-1} & * \\ 0 & \cdots & \cdots & \cdots & \lambda_n \end{bmatrix}$$

where the $*$ entries are 0 or 1. $\boldsymbol{J}$ is called the ***Jordan Canonical Form*** *of $\boldsymbol{A}$. The diagonal entries of $\boldsymbol{J}$ are the eigenvalues of $\boldsymbol{A}$, which appear the same number of times as their algebraic multiplicity.*

Lemma 5.5. *Let $\boldsymbol{A} \in \boldsymbol{M}_n$. Let $\lambda_1, ..., \lambda_n$ be scalars (in $\mathbb{R}$ or $\mathbb{C}$) and $\vec{v}_1, ...\vec{v}_n$ be vectors in $\mathbb{R}^n$ (or $\mathbb{C}^n$). Suppose*

$$\boldsymbol{C} = \begin{bmatrix} | & | & \cdots & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & \cdots & | \end{bmatrix}, \boldsymbol{D} = \begin{bmatrix} \lambda_1 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \lambda_n \end{bmatrix}.$$

Then $\boldsymbol{AC} = \boldsymbol{CD} \iff \vec{v}_1$ is the ***eigenspace*** *$\boldsymbol{E}_{\lambda_i}$ of $\boldsymbol{A}$, (i.e. $\boldsymbol{A}\vec{v}_i = \lambda_i \vec{v}_i$).*

Theorem 5.6. *$\boldsymbol{A}$ is diagonalizable $\iff$ $\boldsymbol{A}$ has n linearly independent eigenvectors (i.e. there is a basis consisting of eigenvectors $\boldsymbol{A}$, called an* ***eigenbasis****).*

Corollary 5.6.1. *Suppose $\boldsymbol{A} \in \boldsymbol{M}_n$ has n distinct eigenvalues. Then $\boldsymbol{A}$ is diagonalizable (the converse is not true).*

Theorem 5.7. *If $\boldsymbol{A} \in \boldsymbol{M}_n(\mathbb{R})$ and $\boldsymbol{A}$ is symmetric (i.e. $\boldsymbol{A} = \boldsymbol{A}^T$), then $\boldsymbol{A}$ is diagonalizable. Moreover, there is a diagonalizing matrix $\boldsymbol{C}$ such that*
(1) $\boldsymbol{C} \in \boldsymbol{M}_n(\mathbb{R})$ (implying all eigenvalues of $\boldsymbol{A}$ are real numbers).
(2) The columns of $\boldsymbol{C}$ are unit vectors that are orthogonal to each other (columns of $\boldsymbol{C}$ form an ***orthonormal*** *basis of $\mathbb{R}^n$*

Theorem 5.8. *Let λ be an eigenvalue of a matrix $\boldsymbol{A} \in \boldsymbol{M}_n$. Then the algebraic multiplicity of λ is no less than the geometric multiplicity of λ.*

Theorem 5.9. *$\boldsymbol{A}$ is diagonalizable if and only if for every eigenvalue λ of $\boldsymbol{A}$, the algebraic multiplicity equals the geometric multiplicity.*

6 Chapter 6: Projections

Definition 6.1. *Let $\boldsymbol{W}$ be a subspace of a vector space $\boldsymbol{V}$ with inner product $\langle , \rangle$. The* ***orthogonal complement*** *of $\boldsymbol{W}$ in $\boldsymbol{V}$ is defined as $\boldsymbol{W}^\perp = \{\vec{v} \in \boldsymbol{V} \mid \langle \vec{v}, \vec{w} \rangle = 0, \forall \vec{w} \in \boldsymbol{W}\}$*

Theorem 6.1. *$\boldsymbol{W}^\perp$ is a subspace of $\boldsymbol{V}$.*

Theorem 6.2. *$\boldsymbol{W} \cap \boldsymbol{W}^\perp = \{\vec{0}\}$.*

Given a subspace $\boldsymbol{W}$ of $\mathbb{R}^n$, how do we find $\boldsymbol{W}^\perp$?
(1) Find a spanning set $\vec{v}_1, ..., \vec{v}_k$ for $\boldsymbol{W}$
(2) Construct a matrix $\boldsymbol{A}$ whose matrix $\boldsymbol{A}$ whose rows are $\vec{v}_1, ..., \vec{v}_k$.
(3) Solve $\boldsymbol{A}\vec{x} = \vec{0}$ (i.e. find the nullspace of $\boldsymbol{A}$). The nullspace of $\boldsymbol{A}$ is $\boldsymbol{W}^\perp$

Theorem 6.3. *Let $\boldsymbol{W}$ be a subspace of $\mathbb{R}^n$.*
(1) $\dim(\boldsymbol{W}^\perp) = n - \dim(W)$
(2) $(\boldsymbol{W}^\perp)^\perp = \boldsymbol{W}$
(3) Each vector $\vec{v} \in \mathbb{R}^n$ can be uniquely written $\vec{v} = \vec{x} + \vec{y} = \vec{v}_W + \vec{v}_{W^\perp}$ where $\vec{x} \in \boldsymbol{W}, \vec{y} \in \boldsymbol{W}^\perp$

Definition 6.2. *Let $\boldsymbol{W}$ be a subspace of $\mathbb{R}^n$ and let $\vec{v} \in \mathbb{R}^n$. The* ***projection of $\vec{v}$ onto $\boldsymbol{W}$*** *is the unique vector $\vec{x} \in \boldsymbol{W}$ such that $\vec{v} = \vec{x} + \vec{y}, \vec{y} \in \boldsymbol{W}^\perp$. We denote $\vec{x}$ by $\vec{v}_{\boldsymbol{W}}$ and $\vec{y}$ by $\vec{v}_{\boldsymbol{W}^\perp}$.*

How do we find $\vec{v}_W$?
(1) Find a basis for $\boldsymbol{W}(\vec{v}_1, ..., \vec{v}_k)$.
(2) Find a basis for $\boldsymbol{W}^\perp, (\vec{v}_{k+1}, ..., \vec{v}_n)$, (i.e. find basis for nullspace of $\boldsymbol{A}\vec{x} = \vec{0}$, where the rows of $\boldsymbol{A}$ are the basis for $\boldsymbol{W}$).
(3) Coordinatize $\vec{v}$ with respect to $\boldsymbol{B} = (\vec{v}_1, ..., \vec{v}_n)$, (i.e. solve $\boldsymbol{C}\vec{x} = \vec{v}$, where the columns of $\boldsymbol{C}$ are $\vec{v}_1, ..., \vec{v}_n$).
(4) $\vec{v}_{\boldsymbol{W}} = r_1\vec{v}_1 + ... + r_k\vec{v}_k$, using $\vec{v}_{\boldsymbol{B}} = [r_1, ..., r_n]$ from step 3 to find r_i.

Theorem 6.4. *Let $\boldsymbol{W}$ be a subspace of $\mathbb{R}^n, \vec{v} \in \mathbb{R}^n$, and $\vec{v}_{\boldsymbol{W}}$ be the projection of $\vec{v}$ onto $\boldsymbol{W}$. Then $\|\vec{v} - \vec{v}_{\boldsymbol{W}}\| \leq \|\vec{v} - \vec{w}\|$ for all $\vec{w} \in \boldsymbol{W}$ (i.e. the distance between $\vec{v}, \vec{v}_{\boldsymbol{W}}$ is less than or equal to the distance between $\vec{v}, \vec{w}$).*

Note: this theorem is used when we want to determine a vector in $\boldsymbol{W}$ that best approximates the solutuon to $\boldsymbol{A}\vec{x} = \vec{b}$: we simply project the solution onto $\boldsymbol{W}$.

Theorem 6.5 (Projection Formula)**.** *Another way to calculate the* ***projection of $\vec{v}$ onto*** *$sp(\vec{a})$ is*

$$\boxed{\vec{v}_{\boldsymbol{W}} = (\frac{\vec{v} \cdot \vec{a}}{\vec{a} \cdot \vec{a}})\vec{a}}$$

The Gram-Schmidt Process

Definition 6.3. *A subset of nonzero vectors* $\boldsymbol{S}$ *in* $\boldsymbol{V}$ *is* ***orthonormal*** *if*

$$\langle \vec{v}, \vec{w} \rangle = \begin{cases} 1 & \vec{v} = \vec{w} \\ 0 & \vec{v} \neq \vec{w} \end{cases}$$

An ***orthonormal set*** *is an orthogonal set in which all vectors have length (norm)* 1.

Bonus notation: $\{\vec{v}_1, \vec{v}_2, ...\vec{v}_k\}$ *is orthonormal if* $\langle \vec{v}_i, \vec{v}_j \rangle = \delta_{ij}$*, where the* ***Kronecker Delta****,*

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Theorem 6.6. *Let* $\boldsymbol{V}$ *be a finite dimensional inner product space. Then every nonzero subspace of* $\boldsymbol{V}$ *has an orthonormal basis.*

Constructive Proof: Algorithm called the **Gram-Schmidt Process**

- Start with a subspace $\boldsymbol{W}$ of $\boldsymbol{V}$ and a basis for $\boldsymbol{W} = \{\vec{a}_1, ..., \vec{a}_k\}$
- Use this to construct an orthogonal basis $\{\vec{v}_1, ..., \vec{v}_k\}$ for $\boldsymbol{W}$. To do this, use facts about projection from 6.1.
- Normalize the vectors to get an orthonormal basis $\{\vec{q}_1, ..., \vec{q}_k\}$ for $\boldsymbol{W}$. Note that then $\vec{q}_i = \frac{\vec{v}_i}{\|\vec{v}_i\|}$.

Orthogonal Matrices

Definition 6.4. *An nxn matrix* $\boldsymbol{A}$ *is orthogonal if* $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$.
Note:

- *Orthogonal matrices are invertible (*$\boldsymbol{A}^{-1} = \boldsymbol{A}^T$*).*
- *If* $\boldsymbol{A}$ *is orthogonal, the columns of* $\boldsymbol{A}$ *are orthogonal vectors of length 1/the columns of* $\boldsymbol{A}$ *form an orthonormal set.*
- $\boldsymbol{A}$ *is orthogonal* $\iff$ *columns of* $\boldsymbol{A}$ *are orthonormal.*

Theorem 6.7. $\boldsymbol{A} \in \boldsymbol{M}_n$*. The columns of* $\boldsymbol{A}$ *are orthonormal* $\iff$ *the rows of* $\boldsymbol{A}$ *are orthonormal.*

Theorem 6.8. $\boldsymbol{A} \in \boldsymbol{M}_n$. *TFAE:*
(1) $\boldsymbol{A}$ is orthogonal.
(2) The rows of $\boldsymbol{A}$ form an orthonormal basis for $\mathbb{R}^n$.
(3) The columns of $\boldsymbol{A}$ form an orthnormal basis for $\mathbb{R}^n$.
(4, Orthogonal matrices preserve the dot product) For all $\vec{x}, \vec{y} \in \mathbb{R}^n, (\boldsymbol{A}\vec{x}) \cdot (\boldsymbol{A}\vec{y}) = \vec{x} \cdot \vec{y}$

Corollary 6.8.1. *For $\boldsymbol{A}$, an orthogonal nxn matrix $\vec{x}, \vec{y} \in \mathbb{R}^n$. Then*
(1, Orthogonal matrices preserve length) $\|\boldsymbol{A}\vec{x}\| = \|\vec{x}\|$
(2, Orthogonal matrices preserve angles) The angle between $\vec{x}, \vec{y}$ is the same as the angle between $\boldsymbol{A}\vec{x}, \boldsymbol{A}\vec{y}$.

Definition 6.5. *For $\boldsymbol{V}$, a vector space with a defined inner product. We say a linear transformation $\boldsymbol{T} : \boldsymbol{V} \to \boldsymbol{V}$ is **orthogonal** if $\langle \boldsymbol{T}(\vec{v}), \boldsymbol{T}(\vec{w}) \rangle = \langle \vec{v}, \vec{w} \rangle$ for all $\vec{v}, \vec{w} \in \boldsymbol{V}$ (i.e. dot product is preserved).*

Theorem 6.9 (Fundamental Theorem of Real Symmetric Matrices)**.** *Every nxn real symmetric matrix $\boldsymbol{A}$ can be diagonalized by an **orthogonal matrix** i.e. not only can we find a basis for $\mathbb{R}^n$ consisting of eigenvectors of $\boldsymbol{A}$, but we can also find an **orthonormal basis of** $\mathbb{R}^n$ consisting of eigenvectors of $\boldsymbol{A}$.*

Theorem 6.10. *Let $\boldsymbol{A}, \boldsymbol{B}$ be orthogonal nxn matrices. Then $\boldsymbol{AB}$ is orthogonal (i.e. orthogonal amtrices are **closed** under multiplication).*

7 Change of Bases

How to translate from any ordered basis to any other ordered basis

How do we move from vector expressed in $\boldsymbol{B} : \vec{v}_{\boldsymbol{B}}$ to translation in $\boldsymbol{B'} : \vec{\boldsymbol{v}}_{\boldsymbol{B}}$?
Claim: $\vec{v}_{\boldsymbol{B'}} = (\boldsymbol{M}_{\boldsymbol{B'}})^{-1}(\boldsymbol{M}_{\boldsymbol{B}})\vec{v}_{\boldsymbol{B}}$

Definition 7.1. *Let $\boldsymbol{B}, \boldsymbol{B'}$ be ordered bases of $\mathbb{R}^n$. Let $\boldsymbol{M}_{\boldsymbol{B}}, \boldsymbol{M}_{\boldsymbol{B'}}$ be the n x n matrices whose columns are determined by $\boldsymbol{B}, \boldsymbol{B'}$, respectively. The* ***change of basis (aka, change of coordinate) matrix from B to B′*** *is defined as*

$$c_{\boldsymbol{B},\boldsymbol{B'}} = (\boldsymbol{M}_{\boldsymbol{B'}})^{-1}(\boldsymbol{M}_{\boldsymbol{B}}) \text{ and } c_{\boldsymbol{B},\boldsymbol{B'}}\vec{v}_{\boldsymbol{B}} = \vec{v}_{\boldsymbol{B'}}$$