MAT188: Linear Algebra

Arnav Patil

University of Toronto

Contents

1 Vectors, Limits, Sets, and Planes

Introduction

A vector is a representation of displacement – it gives a set of instructions on how to move from one point to the next. This is known as a **coordinate vector**. When the coordinate vector is taken from the origin, it becomes known as the **standard representation** of the vector. Each vector has n components, where n is a positive integer.

The collection of all vectors with n components forms $\mathbb{R}^n$, also known as the **n-dimensional Euclidean vector space**.

We use the head to tail rule for vector addition; that is, add up the individual components of each vector being summed to get the resulting vector.

Definition: Parallel Vectors

Two vectors are parallel if $\vec{v} = k\vec{w}$ for some $k \in \mathbb{R}^n$, or is one is a scalar multiple of the other.

We can define a matrix as an array of numbers. For example, we say that

$$M = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

is a 2×3 matrix. A column vector is just a matrix with one column and multiple rows, and a row vector is just a matrix with multiple columns and one row.

There are eight "axioms" of linear algebra that will guide us through this course, and allows us to ensure vector arithmetic works the way we want it to:

1. Vector addition is associative, or $(\vec{v} + \vec{w}) + \vec{u} = \vec{v} + (\vec{w} + \vec{u})$,
2. Vector addition is commutative, or $\vec{v} + \vec{w} = \vec{w} + \vec{v}$,
3. $\vec{v} + \vec{0} = \vec{v}$,
4. For each $\vec{v} \in \mathbb{R}^n$, there exists a vector $\vec{x}$ such that $\vec{v} + \vec{x} = \vec{0}$,
5. $k(\vec{v} + \vec{w}) = k\vec{v} + k\vec{w}$,
6. $(c + k)\vec{v} = c\vec{v} + k\vec{v}$,
7. $c(k\vec{v}) = (ck)\vec{v}$, and
8. $1\vec{v} = \vec{v}$.

Solidify

Definition: Dot Product

Let $\vec{v}$ and $\vec{w}$ be two vectors with components $v_1, v_2, ..., v_n$ and $w_1, w_2, ..., w_n$ respectively. The dot product of $\vec{v}$ and $\vec{w}$, denoted by $\vec{v} \cdot \vec{w}$, is given by: $v_1w_1, v_2w_2, ..., v_nw_n$.

Definition: Norm

Let $\vec{v}$ be a vector in $\mathbb{R}^n$. The norm of $\vec{v}$ is given by:

$$\sqrt{v_1^2 + v_2^2 + ... + v_n^2}$$

For vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$, the angle between $\vec{v}$ and $\vec{w}$ can be given by:

$$\arccos\left(\frac{\vec{v}\cdot\vec{w}}{||\vec{v}||||\vec{w}||}\right)$$

Definition: Perpendicular

Two vectors $\vec{v}$ and $\vec{w}$ are perpendicular or orthogonal if their dot product is 0, or if $\vec{v}\cdot\vec{w}=0$.

Now we'll discuss a little bit about sets.

Definition: Subset

We say a set A is a subset of a set B, and write $A \subseteq B$, if all the elements of A are also in B. In other words, $A \subseteq B$, if for every $a \in A$, $a \in B$.

Definition: Equality of Sets

We say sets A and B are equal if A is a subset of B and B is a subset of A. That is $A = B$ if $A \subseteq B$ and $B \subseteq A$.

Expand

We can define lines in set notation. Let's take an example of $y = -2x$. As a set, we can describe it as:

$$l = \{(x, y)|y = -2x\} = \{(t, -2t)|t \in \mathbb{R}\}$$

We can generalize this to vectors:

$$l = \{\vec{p} + t\vec{d}|t \in \mathbb{R}\}$$

We refer to $\vec{v} = \vec{p} + t\vec{d}$ as the **vector parametric form** of the line l.

Similarly, we can define a plane in $\mathbb{R}^n$ as well.

Definition: Plane

A plane in $\mathbb{R}^n$ has the form:

$$P = \{t\vec{m} + s\vec{n} + \vec{b}|s, t \in \mathbb{R}\}$$

where $\vec{m}$ and $\vec{n}$ may NOT be parallel.

2 Systems of Linear Equations

Introduce

Definition: Linear Equations

An equation of the form $a_1x_1 + a_2x_2 + ... + a_nx_n = b$ is known as a linear equation.

A collection of linear equations with the same variables is known as a **system of linear equations**. Values for each x_n that make all the equations true simultaneously are known as **solutions to the system**.

There are two matrices associated with a system of linear equations.

$$\begin{cases} 2x_1 + 4x_2 & = 8 \\ 4x_1 + 3x_2 & = 10 \\ x_1 - x_2 & = 2 \end{cases}$$

The coefficient matrix:

$$\begin{bmatrix} 2 & 4 \\ 4 & 3 \\ 1 & -1 \end{bmatrix}$$

and the augmented matrix:

$$\left[\begin{array}{cc|c} 2 & 4 & 8 \\ 4 & 3 & 10 \\ 1 & -1 & 2 \end{array}\right]$$

We use the process of **Gauss-Jordan elimination** to systematically solve systems of linear equations. There are three fundamental operations that do not change the general solution to the system:

1. Multiplying an equation by a scalar,
2. Adding a multiple of an equation to another, and
3. Switching the order of equations.

Definition: Row Echelon Form and Reduced Row Echelon Form

A matrix is in row echelon form if it satisfies the following properties:

1. All zero rows are at the bottom,
2. The leading entry in each row is to the right of the leading entry of the row above, and
3. All entries below a leading entry are zero.

A matrix is in reduced row echelon form if it is in REF and additionally:

1. All the leading entries are 1 (we call them "leading ones"), and
2. Each leading 1 is the only nonzero entry in its column.

Definition: System of Linear Equations

A **particular solution** is one single tuple $(c_1, c_2, ..., c_n)$ that makes all the equations true simultaneously. The set of all possible solutions to a system of linear equations is known as the **general solution**.

Solidify

For each matrix M, there are many REF forms of it, however, there is only one RREF form.

The first nonzero entry is called that row's **leading entry**, or **leading 1** if the entry is a 1. A **pivot position** is a location in a matrix that corresponds to a leading 1 in the RREF in that matrix.

If a variable's column is a pivot column, it is known as a **leading/basic/dependent variable**; if not, it's known as a **free variable**.

Here is a step-by-step process to solve systems of linear equations:

1. Write the system in standard form by ordering the variables into an augmented matrix.
2. Reduce the matrix into its RREF.
3. Identify free and basic variables.
4. Transform the RREF matrix into equations. The result is a much simpler system of linear equations.

5. Solve for the basic variables in term of the free variables.

6. Let the free variables run over all scalars.

7. We typically write the general solution in set builder notation.

A system of equations is **consistent** if there's at least one solution; otherwise, it is called **inconsistent**.

Theorem: Solution Type

A linear system is inconsistent only if the RREF matrix has the equation $0 = 1$, in other words, if the augmented column is a pivot column. Otherwise, if a linear system is consistent, then it has:

- Infinitely many solutions (there's at least one free variable), or
- Exactly one solution (all the variables are leading).

Definition: Rank

The rank of a matrix M is the number of leading 1s in RREF(M).

Expand

Some matrix terminology:

- If A is $n \times n$, then the matrix is a **square matrix**,
- A square matrix is **diagonal** if all entries not in the diagonal line are 0, that is, $a_{ij} = 0$ when $i \neq j$,
- A square matrix is **upper triangular** if all its entries below the main diagonal are zero, and **lower triangular** inversely,
- A **zero matrix** has all entries set to 0, and
- A square diagonal matrix is called the **identity matrix** if all its nonzero entries are set to 1.

Definition: Matrix-Vector Product

Let A be an $n \times m$ matrix with row vectors $\vec{w}_1, \vec{w}_2, ..., \vec{w}_n$ and $\vec{x} = x_1, x_2, ..., x_n$ be a vector in $\mathbb{R}^n$. We define:

$$A\vec{x} = \begin{bmatrix} -- & \vec{w}_1 & -- \\ -- & \vec{w}_2 & -- \\ -- & \vdots & -- \\ -- & \vec{w}_n & -- \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} \vec{w}_1 \cdot \vec{x} \\ \vec{w}_2 \cdot \vec{x} \\ \vdots \\ \vec{w}_n \cdot \vec{x} \end{bmatrix}$$

Theorem: Matrix Form of a Linear System

We can write the linear system with augmented matrix $[A|\vec{b}]$ in matrix form as $A\vec{x} = \vec{b}$.

Definition: Linear Combination

A linear combination of finite vectors $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ in $\mathbb{R}^n$ is an expression of the form:

$$c_1\vec{v}_1 + c_2\vec{v}_2 + ... + c_n\vec{v}_n$$

Expand

Theorem: Algebraic Rules for $A\vec{x}$

If A is an $n \times m$ matrix, $\vec{x}$ and $\vec{y}$ are vectors in $\mathbb{R}^n$, and k is a scalar, then:

- $A(\vec{x}+\vec{y}) = A\vec{x} + A\vec{y}$, and
- $A(k\vec{x}) = k(A\vec{x})$.

3 Linear Transformations

Introduce

We sometimes use more archaic words to describe linear algebra than calculus, this is due to linear algebra's deep connections to geometry. For example, what we call a 'function' in calculus, we may call a 'transformation' or 'mapping' in linear algebra.

The space from which the input to a transformation comes is called the **domain**. The **codomain** is the space in which the output vectors lie, while the set of vectors that make up the output is called the **range**.

Definition: Linear Transformation

A linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$ is a mapping T: $\mathbb{R}^m \to \mathbb{R}^n$ that satisfies:

1. $T(\vec{x}+\vec{y}) = T(\vec{x}) + T(\vec{y})$ for all vectors $\vec{x}, \vec{y} \in \mathbb{R}^m$, and
2. $T(k\vec{x}) = kT(\vec{x})$ for all vectors $\vec{x} \in \mathbb{R}^m$ and all scalars $k \in \mathbb{R}$.

Solidify

Definition: Standard Vectors

The standard vector $\vec{e}_i \in \mathbb{R}^m$ is defined as a unit vector with all zeros except for the 1 in the i^{th} entry.

Theorem: Linear Transformations and Matrices

If $T : \mathbb{R}^m \to \mathbb{R}^n$ is a linear transformation, there exists an $n \times m$ matrix A such that $T(\vec{x}) = A\vec{x}$.

The **standard matrix representation** of the linear transformation T is:

$$\begin{bmatrix} | & | & & | \\ \vec{T}(\vec{e}_1) & \vec{T}(\vec{e}_2) & \cdots & \vec{T}(\vec{e}_m) \\ | & | & & | \end{bmatrix}$$

Expand

The linear transformation that rotates vector $\vec{u}$ by θ radians is:

$$T(\vec{u}) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

We call $\vec{x}^{||} = \left(\frac{\vec{u}\cdot\vec{x}}{||u||^2}\right)$ the **projection of $\vec{x}$ onto $\vec{u}$.**

$$proj_{\vec{u}}(\vec{x}) = \left(\frac{\vec{u}\cdot\vec{x}}{||u||^2}\right)$$

Similarly, the reflection is given by:

$$ref_L(\vec{x}) = \vec{x}^{||} - \vec{x}^{\perp}$$

or

$$ref_L(\vec{x}) = 2proj_{\vec{u}}(\vec{x}) - \vec{x}$$

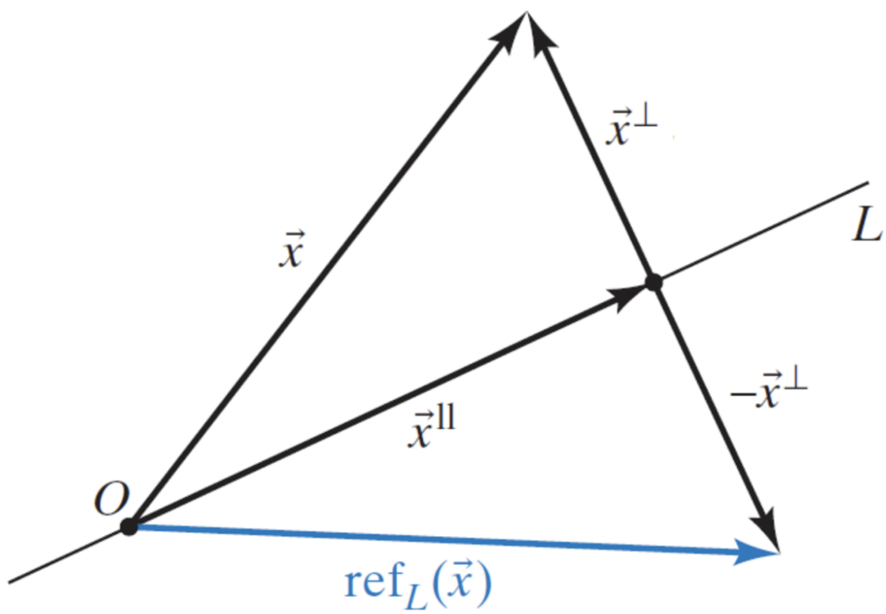

Figure 1: Projection and Reflection of a Vector

4 Composition and Inverse of Linear Transformations

Introduce

Not any two linear transformations can be composed. For $S \circ T$ to be a valid composition, the domain of S must be the codomain of T. The inverse, $T \circ S$, may not necessarily be defined. The composition of two linear transformations will also be a linear transformation as it is closed under vector addition and scalar multiplication.

Definition: Product of Matrices

Let $T : \mathbb{R}^m \to \mathbb{R}^n$, $T(\vec{x}) = A\vec{x}$ and $S : \mathbb{R}^n \to \mathbb{R}^p$, $S(\vec{y}) = B\vec{y}$ be linear transformations with associated standard matrices A and B respectively. Then BA is defined to be the unique matrix associated to the composition $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$.

Solidify

Theorem: The Columns of the Matrix Product

Let $A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_m \\ | & | & & | \end{bmatrix}$ be an $n \times m$ matrix and B is a $p \times n$ matrix. The BA is a $p \times m$ matrix given by:

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

Expand

Injective (or one-to-one) transformation – A transformation T is called injective if no two vectors in the domain get sent to the same vector in the same codomain.

Surjective (or onto) transformation – A transformation T is called surjective if every vector in the codomain has a vector that gets mapped to it, by T.

Bijectivity – A transformation is called bijective if it is both injective and surjective.

Theorem: Bijective Maps are Invertible

The mapping $T : X \to Y$ is invertible iff it is bijective.

Definition: Invertibility of Matrices

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

The product of two invertible matrices is also invertible. Furthermore, $(AB)^{-1} = B^{-1}A^{-1}$, meaning that the order of the matrix multiplication reverses when we take an inverse.

Theorem: Criterion for Invertibility

If A is an $n \times n$ matrix, and the RREF of $A = I_n$, then A is invertible.

Theorem: Inverting a Matrix

If A is an invertible $n \times n$ matrix, then A^{-1} can be computed by creating an augmented matrix $[A \,|\, I_n]$ (notice that we are augmenting by an $n \times n$ matrix, not just a single column), and performing elementary row operations on A until it reaches its RREF (which is I_n). The resulting augmented matrix will then be of the form $[I_n \,|\, A^{-1}]$.

5 Determinant

Introduce

Definition: 2×2 Determinant

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a 2×2 matrix. The determinant of A is the scalar $ad - bc$.

Theorem: Determinant and Invertibility of a 2×2 Matrix

1. A 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible iff $\det A = ad - bc \neq 0$
2. If A is invertible then $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

The absolute value of the determinant of A is the expansion factor, or the ratio by which T changes the area of any subset Ω in $\mathbb{R}^2$.

$$\frac{\textbf{area of } T(\Omega)}{\textbf{area of } \Omega} = |\det A|$$

Solidify

Definition: Cross Product in $\mathbb{R}^3$

The cross product $\vec{v} \times \vec{w}$ of two vectors $\vec{v}$ and $\vec{w}$ in $\mathbb{R}^3$ is the vector in $\mathbb{R}^3$ with the following three properties:

- $\vec{v} \times \vec{w}$ is orthogonal to both $\vec{v}$ and $\vec{w}$.
- $\|\vec{v} \times \vec{w}\| = \|\vec{v}\|\|\vec{w}\| \sin\theta$, where θ is the angle between $\vec{v}$ and $\vec{w}$, with $0 \leq \theta \leq \pi$.
- The direction of $\vec{v} \times \vec{w}$ follows the right-hand rule.

Definition: 3×3 Determinant

If $A = [\vec{v}_1\ \vec{v}_2\ \vec{v}_3]$ is a 3×3 matrix, then:

$$\det A = \vec{v}_3 \cdot (\vec{v}_1 \times \vec{v}_2)$$

Theorem: Invertibility and Determinant

A square matrix A is invertible if and only if $\det A \neq 0$.

Theorem: Determinant of the Transpose

If A is a square matrix then $\det A = \det A^T$.

Theorem: Determinants of Products and Powers

Let A, B be $n \times n$ matrices:

1. $\det AB = \det A \det B$
2. $\det(A^m) = (\det A)^m$

Theorem: Determinant of an Inverse

If A is an invertible matrix:

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

Expand

Theorem: Elementary Row Operations and Determinants

Suppose A is an $n \times n$ matrix:

1. If B is obtained from A by dividing a row of A by a scalar k, then $\det B = \frac{1}{k} \det A$.
2. If B is obtained from A by a row swap, then $\det B = -\det A$.
3. If B is obtained from A by adding a multiple of a row of A to another row, then $\det B = \det A$.

Definition: An Elementary Matrix

An elementary matrix is a matrix you get by applying an elementary row reduction step to the identity matrix.

Theorem: Cramer's Rule

Consider the linear system $A\vec{x} = \vec{b}$, where A is an invertible $n \times n$ matrix. The components x_i of the solution vector $\vec{x}$ are

$$x_i = \frac{\det A_{b,i}}{\det A}$$

where $A_{b,i}$ is the matrix obtained by replacing the i-th column of A by $\vec{b}$.

6 Subspaces, Span, Linear Independence, and Basis

Introduce

Definition: Subspace

A subspace W of $\mathbb{R}^n$ which contains $\vec{0}$ and is closed under vector addition and scalar multiplication.

Definition: Span

Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ be vectors in $\mathbb{R}^n$. The set of all linear combinations of $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ is called their span. That is,

$$\text{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\} = \{c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_m\vec{v}_m \mid c_1, c_2, \ldots, c_m \in \mathbb{R}\}$$

If $\text{span}(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m) = V$, then $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\}$ is called a spanning set for V, or is said to span V.

Theorem: Span is a Subspace

Given vectors $\vec{v}_1, \vec{v}_2, ..., \vec{v}_m$ in $\mathbb{R}^n$, $\text{span}(\vec{v}_1, \vec{v}_2, ..., \vec{v}_m)$ is a subspace.

Definition: Image

The image of $f : X \to Y$ is defined to be the set $\text{im}(f) = \{f(x) : x \in X\} = \{y \in Y \mid f(x) = y \text{ for some } x \in X\}$.

Theorem: Image is a Subspace

Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation given by $T(\vec{x}) = A\vec{x}$.

Then $\text{im}(T) = \text{Col}(A)$ is a subspace of $\mathbb{R}^n$.

Solidify

Definition: Kernel

The kernel of linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ is the set of all vectors in the domain such that $T(\vec{v}) = \vec{0}$.

$$\ker(T) = \{\vec{v} \in \mathbb{R}^m | T(\vec{v}) = \vec{0}\}$$

Theorem: Kernel is a Subspace

Let $T : \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation given by $T(\vec{x}) = A\vec{x}$.

Then $\text{Im}(T) = \text{Col}(A)$ is a subspace of $\mathbb{R}^n$.

Theorem:

For a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$, T is injective if and only if $\ker(T) = \{\vec{0}\}$, and T is surjective if and only if $\text{im}(T) = \mathbb{R}^n$.

Definition: Linear Relation

Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ be vectors in a subspace V of $\mathbb{R}^m$. A linear relation among $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is any equation of the form

$$c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}$$

where c_i are scalars.

Definition: Linearly Dependent

Let $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ be vectors in a subspace V of $\mathbb{R}^m$. We say $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ are linearly dependent if there are scalars $c_1, \ldots, c_n$ that are not all zero such that $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}$.

Definition: Linearly Independent

Let $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ are linearly independent if $c_1\vec{v}_1 + \cdots + c_n\vec{v}_n = \vec{0}$ has only one solution, where all the scalars are 0.

Expand

Let's summarize what we know about a set of vectors that's linearly independent:

1. The set $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ contains no redundant vector,
2. None of the $\vec{v}_i s$ can be written as a linear combination of the others,
3. None of the $\vec{v}_i s$ is in the span of the rest,
4. The only relation between $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ is the trivial relation,
5. $\text{Null}(A) = \vec{0}$,
6. Rank of A with m column vectors is m,
7. Linear map $T(\vec{x}) = A\vec{x}$ is injective, and
8. The kernel of the map $T(\vec{x}) = A\vec{x}$ is $\{\vec{0}\}$.

Definition: Basis

A basis of a subspace V of $\mathbb{R}^n$ is a linearly independent set of vectors in V that spans V.

Theorem:

Vectors $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ in V form a basis iff every vector $\vec{v}$ in V can be expressed as a linear combination $\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + ... + c_n\vec{v}_n$. The coefficients c are called the coordinates of $\vec{v}$ with respect to basis $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$.

7 Basis and Coordinates

Introduce

Theorem: Every Spanning Set is Larger or Equals Every Linearly Independent Set in V

Given a nonzero subject of $\mathbb{R}^n$, there are infinitely many bases of V. The theorem above guarantees that ALL bases of a subspace V have the same number of vectors. This is called the dimension of V.

The rank of matrix A is the dimension of $\text{Im}(A)$, and the nullity is the dimension of $\text{Ker}(A)$.
Let $T_A : \mathbb{R}^n \to \mathbb{R}^m$. The rank-nullity theorem states that:

$$\text{Rank}(A) + \text{Nullity}(A) = n$$

Solidify

Theorem: Bases and Unique Representation

Vectors $\vec{v}_1, \vec{v}_2, ..., \vec{v}_m$ in V only make up a basis for V if every vector $\vec{v}$ can be expressed as a linear combination of $\vec{v}_1, \vec{v}_2, ..., \vec{v}_m$.

Definition: Coordinates

Suppose $B = (\vec{v}_1, \dots, \vec{v}_n)$ is an ordered basis of a subspace V. The B-coordinates of $\vec{v} \in V$ are the unique scalars a_i such that

$$\vec{v} = a_1\vec{v}_1 + \cdots + a_n\vec{v}_n.$$

The B-coordinates are arranged into a column vector, denoted $[\vec{v}]_B$. That is,

$$[\vec{v}]_B = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Definition: Change-of-Coordinates Matrix

If $B = (\vec{b}_1, \dots, \vec{b}_n)$ and $C = (\vec{c}_1, \dots, \vec{c}_n)$ are two ordered bases of a subspace V, the change-of-coordinates matrix from B to C is the unique matrix S such that $S[\vec{v}]_B = [\vec{v}]_C$ for all $\vec{v} \in V$.

We usually use a subscript $B \rightarrow C$ $(S_{B\rightarrow C})$ to denote that S changes B-coordinates into C-coordinates.

Expand

Definition: Similar Matrices

Two $n \times n$ matrices are called similar if there exists an invertible matrix S such that $B = S^{-1}AS$.

8 Orthogonal Projection

Introduce

Definition: Orthogonal Set

A set of vectors is orthogonal if each vector's dot product with every other vector is 0.

Definition: Orthonormal Set

A set of vectors is orthonormal if it is orthogonal and the norm of each vector is 1.

Theorem: Orthonormal Sets are Linearly Independent

An orthonormal set is always linearly independent.

Solidify

Definition: Orthogonal Complement

If W is a subspace of $\mathbb{R}^n$, the orthogonal complement of W in $\mathbb{R}^n$ is the set

$$W^\perp = \{\vec{v} \in \mathbb{R}^n : \vec{v} \cdot \vec{w} = 0 \text{ for all } \vec{w} \in W\}$$

Theorem:

If W is a subspace for $\mathbb{R}^n$, then $W^\perp$ must also be a subspace of $\mathbb{R}^n$.

$$\dim W + \dim W^\perp = n$$

Definition: Orthogonal Projection

If W is a subspace of $\mathbb{R}^n$ and if $\vec{v} \in \mathbb{R}^n$, the orthogonal projection of $\vec{v}$ onto W is the unique vector $\vec{w} \in W$ such that $\vec{v} - \vec{w} \in W^\perp$. The orthogonal projection of $\vec{v}$ onto W is denoted $\text{proj}_W(\vec{v})$.

Expand

Theorem: Orthogonal Decomposition

For a vector $\vec{x}$ in subspace V, we can say $\vec{x} = \vec{x}^{||} + \vec{x}^\perp$ where $\vec{x}^{||}$ is in V and $\vec{x}^\perp$ is in $V^\perp$. Furthermore, this representation is unique for every $\vec{x}$ in V.

Theorem: Gram-Schmidt Method

Let V be a subspace of $\mathbb{R}^n$ and $B = \{\vec{b}_1, \ldots, \vec{b}_k\}$ be a basis for V. Then we can construct an orthonormal basis $U = \{\vec{u}_1, \ldots, \vec{u}_k\}$ for V where

$$\vec{u}_i = \frac{\vec{p}_i}{\|\vec{p}_i\|}$$

and

$$\vec{p}_1 = \vec{b}_1, \quad \vec{p}_i = \vec{b}_i - \text{Proj}_{V_{i-1}} \vec{b}_i,$$

$$V_{i-1} = \text{Span}\{\vec{u}_1, \ldots, \vec{u}_{i-1}\} = \text{Span}\{\vec{p}_1, \vec{p}_2, \ldots, \vec{p}_{i-1}\} = \text{Span}\{\vec{b}_1, \ldots, \vec{b}_{i-1}\}.$$

9 Orthogonal Transformations and Least Square

Introduce

A linear transformation that maps a vector from $\mathbb{R}^n$ to $\mathbb{R}^n$ is called **orthogonal** if the magnitude of $\vec{x}$ and $T(\vec{x})$ is the same, for all $\vec{x}$. Furthermore, the mapping $T(\vec{x}) = A\vec{x}$ is orthogonal when the columns of A are an orthonormal basis.

Theorem: Orthogonal Linear Transformations and Dot Product

A linear transformation is orthogonal iff $T(\vec{v}) \cdot T(\vec{w}) = \vec{v} \cdot \vec{w}$ for all $\vec{v}, \vec{w} \in \mathbb{R}^n$.

Theorem: Orthogonal Matrices

Consider an $n \times n$ matrix A. A is orthogonal iff $A^T A = I_n$ or $A^{-1} = A^T$.

Solidify

Theorem:

If A is an $n \times m$ matrix whose columns are linearly independent, then $A^T A$ is an invertible $m \times m$ matrix.

Theorem: Standard Matrix of Orthogonal Projection

Consider a subspace $V \subseteq \mathbb{R}^n$ with basis $B = \{\vec{v}_1, \ldots, \vec{v}_m\}$. Then if we take

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

the projection of $\vec{b}$ onto V is given by the following formula:

$$\text{Proj}_V(\vec{b}) = A(A^T A)^{-1} A^T \vec{b}.$$

10 Eigenvectors and Eigenvalues

Introduce

Working with diagonal matrices is almost always easier because it allows us to work with much more complicated operations.

Definition: Diagonalizable Matrices

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is diagonalizable if there is a basis B of $\mathbb{R}^n$ such that the B-matrix $[T]_B$ of T is diagonal. A matrix A is diagonalizable if the linear transformation defined by left multiplication by A is diagonalizable.

Definition: Eigenvectors & Eigenvalues

An eigenvector of a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is any non-zero vector $\vec{v} \in \mathbb{R}^n$ such that $T(\vec{v}) = \lambda \vec{v}$ for some scalar λ. The scalar λ is called the eigenvalue of T corresponding to the eigenvector $\vec{v}$.

Definition: Eigenbasis

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. An eigenbasis of $\mathbb{R}^n$ for T is a basis of $\mathbb{R}^n$ consisting of eigenvectors of T.

Theorem:

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is diagonalizable iff it has an eigenbasis.

Solidify

Theorem: Eigenvalues of a Matrix

A scalar λ is an eigenvalue of the $n \times n$ matrix A iff $\det(A - \lambda I_n) = 0$.

Definition: Characteristic Polynomial of a Matrix

The characteristic polynomial of an $n \times n$ matrix A is the degree n polynomial

$$f_A(x) = \det(A - xI_n).$$

Definition: Trace of a Matrix

The sum of the diagonal entries of a square matrix A is called the trace of A, denoted by $\mathrm{tr}(A)$.

Theorem: Characteristic Polynomial of a 2×2 Matrix

Let A be a 2×2 matrix. Then

$$\det(A - \lambda I_2) = \lambda^2 - (\mathrm{tr}(A))\lambda + \det A.$$

Definition: Algebraic Multiplicity

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation of a finite-dimensional vector space $\mathbb{R}^n$. The algebraic multiplicity of an eigenvalue λ is the largest power r such that $(x - \lambda)^r$ divides the characteristic polynomial of T.

Expand

Definition: Eigenspace

If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a linear transformation of the vector space $\mathbb{R}^n$ and if λ is an eigenvalue of T, then the subset

$$V_\lambda = \{\vec{v} \in \mathbb{R}^n \mid T(\vec{v}) = \lambda\vec{v}\}$$

consisting of all λ-eigenvectors (together with $\vec{0}$) is a subspace of V, called the λ-eigenspace, or the eigenspace corresponding to λ.

Definition: Geometric Multiplicity

The geometric multiplicity of λ is the dimension of the λ-eigenspace or, equivalently, the maximal size of a linearly independent set of eigenvectors with eigenvalue λ.

11 Diagonalization

Introduce

Theorem:

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation. Eigenvectors of T that correspond to distinct eigenvalues are linearly independent.

Theorem:

If $T : \mathbb{R}^n \to \mathbb{R}^n$ has n eigenvectors then T is diagonalizable.

Let's summarize what we know so far. Let $T(\vec{x}) = A\vec{x}$:

1. The linear transformation T is diagonalizable,
2. There exists an eigenbasis for T,
3. There exists a diagonal matrix D and an invertible matrix S such that $A = SDS^{-1}$,
4. The dimensions of the eigenspaces of T add up to n,
5. Geometric multiplicities of T adds up to n, and
6. Algebraic multiplicities of T add up to n and for every eigenvalue λ, $gemu(\lambda) = almu(\lambda)$

Solidify

Definition: Orthogonally Diagonalizable Map

We say a mapping T is orthogonally diagonalizable if T has an orthonormal eigenbasis.

Definition: Orthogonally Diagonalizable Matrix

A $n \times n$ matrix is orthogonally diagonalizable if there is an $n \times n$ orthogonal matrix Q such that $Q^T AQ$ is diagonal.

Theorem: The Spectral Theorem

The matrix A of a mapping T is orthogonally diagonalizable iff A is symmetric.

Expand

A **distribution vector** is a vector $\vec{x}$ in $\mathbb{R}^n$ where all components add up to 1, and all components are either positive or zero. A square matrix A is a **transition or stochastic matrix** if all of its columns are distribution vectors. Furthermore, if all entries are positive, or nonzero, then we call it **positive**.

12 Singular Value Decomposition

Singular Value Decomposition, or SVD, is a method used to decompose an $n \times m$ matrix into three matrices, $A = U\Sigma V^T$. U is an $n \times n$ is an orthogonal matrix, V is an $m \times m$ orthogonal matrix, and Σ is an $n \times m$

matrix $r = \text{Rank}(A)$ nonzero entries listed in decreasing order.

$$A = \begin{bmatrix} | & | & & | \\ \vec{u}_1 & \vec{u}_2 & \cdots & \vec{u}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ & 0 & & 0 \end{bmatrix} \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_m \\ | & | & & | \end{bmatrix}^T$$

We may also write this as:

$$A = \sigma_1 \vec{u}_1 \vec{v}_1^T + ... + \sigma_r \vec{u}_r \vec{v}_r^T$$

The values σ_i are called the singular values of A and they tell us a lot about the matrix.

We can use singular value decomposition to compress data; for example, we construct a single value decomposition for an RGB image, then set the smaller σ_i's, which are substantially smaller than σ_1, to 0. That way, we can make the vectors (and necessary operations) significantly less intense, while minimizing loss of quality.