# Contents


**1 Systems of Linear Equations and Matrices** **3**

1.1 Introduction to Systems of Linear Equations . . . . . 3
1.2 Gaussian Elimination . . . . . 4
1.3 Matrices and Matrix Operations . . . . . 4
1.4 Inverses and Properties of Matrices . . . . . 5
1.5 Elementary Matrices and Finding $A^{-1}$ . . . . . 7
1.6 More on Linear Systems and Invertible Matrices . . . . . 8
1.7 Diagonal, Triangular, and Symmetric Matrices . . . . . 9

**2 Determinants** **11**

2.1 Determinants by Cofactor Expansion . . . . . 11
2.2 Evaluating Determinants by Row Reduction . . . . . 12
2.3 Properties of Determinants . . . . . 12

**3 Euclidean Vector Spaces** **15**

3.1 Vectors in 2-Space, 3-Space and $n$-Space . . . . . 15
3.2 Norm, Dot Product, and Distance in $\mathbb{R}^{\mathbf{n}}$ . . . . . 16
3.3 Orthogonality . . . . . 18
3.4 Geometry of Linear Systems . . . . . 19

**4 General Vector Spaces** **21**

4.1 Real Vector Spaces . . . . . 21
4.2 Subspaces . . . . . 22
4.3 Linear Independence . . . . . 23
4.4 Coordinates and Bases . . . . . 23
4.5 Dimension . . . . . 24
4.6 Change of Basis . . . . . 24
4.7 Row Space, Column Space, Null Space . . . . . 26
4.8 Rank, Nullity, and Fundamental Matrix Spaces . . . . . 27
4.9 Matrix Transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$ . . . . . 27
4.10 Properties of Matrix Transformations . . . . . 28

**5 Eigenvalues and Eigenvectors** **29**

5.1 Eigenvalues and Eigenvectors . . . . . 29
5.2 Diagonalization . . . . . 30

**6 Inner Product Spaces** **33**
6.1 Inner Products . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 33
6.2 Angle and Orthogonality . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
6.3 Orthogonal and Orthonormal Sets . . . . . . . . . . . . . . . . . . . . . . . 35
6.4 Best Approximation & Least Squares . . . . . . . . . . . . . . . . . . . . . 37

**7 Diagonalization and Quadratic Forms** **39**
7.1 Orthogonal Matrices . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 39

# Chapter 1

# Systems of Linear Equations and Matrices

## 1.1 Introduction to Systems of Linear Equations

**Definition.** A *linear equation* can be written as

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

A *homogenous linear equation* has $b = 0$, so

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = 0$$

**Definition.** A *system of linear equations* or *linear system* is a finite set of linear equations.

A *solution* of a system of linear equations is a system sequence of $n$ numbers

$$x_1 = s_1, \dots, x_n = s_n$$

that makes each equation a true statement.

We often write solutions as $(s_1, s_2 \dots, s_n)$, called an *ordered n-tuple*.

Go over two and three variables and number of solutions. Cover *consistent* and *inconsistent*. Maybe *dependent*.

- $\begin{cases} x + y = 6 \\ x - y = 2 \end{cases}$
- $\begin{cases} \phantom{2}x - 2y = 4 \\ 2x - 4y = 8 \end{cases}$

Talk about parametric solutions $x = f(t)$ and $y = t$.

- $\begin{cases} 5x - 2y - \phantom{1}5z = 1 \\ 10x - 4y - 10z = 2 \\ 15x - 6y - 15z = 3 \end{cases}$

Talk row operations. Do a few examples with solutions.

### 1.1.1 Homework

#10, 15

## 1.2 Gaussian Elimination

Define *row echelon form* and *reduced row echelon form*.

**Ex 1.** •
$$\begin{cases} x_1 - x_2 + 2x_3 - x_4 = -1 \\ 2x_1 + x_2 - 2x_3 - 2x_4 = -2 \\ -x_1 + 2x_2 - 4x_3 + x_4 = 1 \\ 3x_1 - 3x_4 = -3 \end{cases}$$

### 1.2.1 Homework

#7,21,22,25

## 1.3 Matrices and Matrix Operations

**Definition.** A *matrix* is a rectangular array of numbers. The numbers in the array are called *entries*.

*Size* is the number of rows and columns.

*Scalars* are numbers.

Notation:

$$A = [a_{ij}]_{m\times n} = [a_{ij}]$$

Talk about row and column vectors.

$$\mathbf{a} = [a_1\ a_2\ \cdots\ a_n]$$

Define

- Square matrix
- Equality
- Addition and Subtraction
- Scalar Multiplication
- Multiplication
- Linear combinations

- System of linear equations as a matrix product
- Transpose
- Trace of a matrix (only for square matrices)

### 1.3.1 Homework

# 3

## 1.4 Inverses and Properties of Matrices

Properties:

1. $A + B = B + A$
2. $A + (B + C) = (A + B) + C$
3. $A(BC) = (AB)C$
4. $A(B + C) = AB + AC$
5. $(B + C)A = BA + CA$
6. $a(B + C) = aB + aC$
7. $(a + b)C = aC + bC$
8. $a(bC) = (ab)C$
9. $a(BC) = (aB)C$
10. $(-1) \cdot A = -A$

Show order matters in multiplication

**Definition.** A matrix whose entries are all zero is a *zero matrix*

Properties

1. $A + 0 = 0 + A = A$
2. $A - 0 = A$
3. $A - A = A + (-A) = 0$
4. $0A = 0$
5. If $cA = 0$, then $c = 0$ or $A = 0$.

**Ex 2.** $A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix}$.

Show $AB = AC$.

**Ex 3.** $M = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$ and $N = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$.

**Definition.** Define the *identity matrix*, $I_n$.

**Theorem.** *If $R$ is the reduced row echelon form of an $n \times n$ matrix, then either $R$ has at least one row of zeroes or $R = I_n$.*

**Definition.** If $A$ is a square matrix and if a matrix $B$ of the same size can be found such that $AB = BA = I$, then $A$ is *invertible* or *nonsingular* and $B$ is the inverse of $A$. If no such matrix can be found, then $A$ is *singular*.

**Ex 4.** $A = \begin{bmatrix} 4 & 7 \\ 3 & 5 \end{bmatrix}$, $B = \begin{bmatrix} -5 & 7 \\ 3 & -4 \end{bmatrix}$,

**Theorem.** *If $B$ and $C$ are inverses of $A$, then $B = C$.*

*Proof.* $BA = I$, so $(BA)C = IC = C$. But $(BA)C = B(AC) = BI = B$. so $B = C$. □

**Theorem.** *If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has an inverse, then*

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

**Theorem.** *If $A$ and $B$ are invertible matrices of the same size, then $AB$ is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.*

*Proof.* Just multiple them out. □

### 1.4.1 Powers of a Matrix

**Definition.**

1. $A^0 = I$
2. $A^n = AA \cdots A$ (n factors)
3. $A^{-n} = \left(A^{-1}\right)^n = A^{-1}A^{-1} \cdots A^{-1}$

**Theorem.**

1. *$A^r A^s = A^{r+s}$ for $r, s \geq 0$*
2. *$\left(A^r\right)^s = A^{rs}$ for $r, s \geq 0$*
3. *$\left(A^{-1}\right)^{-1} = A$*
4. *$A^{-n} = \left(A^{-1}\right)^n = \left(A^n\right)^{-1}$*

### 1.4.2 Matrix Polynomials

Go over an example.

### 1.4.3 Transposes

**Theorem.** 1. $\left(A^T\right)^T = A$

2. $(A+B)^T = A^T + B^T$
3. $(kA)^T = kA^T$
4. $(AB)^T = B^T A^T$
5. $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

*Proof.* $A^T(A^{-1})^T = (A^{-1}A)^T = I^T = I$

$(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$ □

### 1.4.4 Homework

# 4, #10, #18

## 1.5 Elementary Matrices and Finding $A^{-1}$

### 1.5.1 Elementary matrices

**Theorem.** *Every elementary matrix is invertible and the inverse is also an elementary matrix.*

### 1.5.2 Equivalence Theorem

**Theorem.** *If $A$ is an $n \times n$ matrix, then the following statements are equivalent.*

(a) *$A$ is nonsingular.*

(b) *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*

(c) *The reduced row echelon form of $A$ is $I_n$.*

(d) *$A$ is expressible as a product of elementary matrices.*

*Proof.* $a \to b$: Assume $A^{-1}$ exists and let $\mathbf{x}_0$ be any solution to $A\mathbf{x} = \mathbf{0}$. Then $A\mathbf{x}_0 = \mathbf{0}$. Multiply both sides by $A^{-1}$ to show that $\mathbf{x}_0 = \mathbf{0}$.

$b \to c$: The solution to $A\mathbf{x} = \mathbf{0}$ has to look like $x_1 = 0$, $x_2 = 0$ and so on. Thus the augmented matrix can be reduced to $[I_n|\mathbf{0}]$.

$c \to d$: Using the theorem that all row operations are just elementary matrices, we know that $E_k \cdots E_2 E_1 A = I_n$. But we also know that all elementary matrices are invertible (with elementary inverses), so $A = E_1^{-1} E_2^{-1} \cdots E_k^{-1} I_n = E_1^{-1} E_2^{-2} \cdots E_k^{-1}$.

$d \to a$: Use the theorem about multiplication and inverses. □

### 1.5.3 Inverting Matrices

If $A = E_1 E_2 \cdots E_k I_n$, then $A^{-1} = E_k^{-1} \cdots E_2^{-1} E_1^{-1} I_n$. The same row operations that reduce $A$ to $I_n$ will also transform $I_n$ into $A^{-1}$.

**Theorem.** *To find the inverse of a nonsingular matrix $A$, find the sequence of row operations that reduces $A$ to $I_n$ and perform that exact same sequence on $I_n$ to*

**Ex 5.** $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 5 & 6 & 0 \end{bmatrix}$

**Ex 6.** $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$

**Ex 7.** $A = \begin{bmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 4 & 1 & 4 \end{bmatrix}$

## 1.6 More on Linear Systems and Invertible Matrices

### 1.6.1 More About Solutions

**Theorem.** *A system of linear equations has zero, one, or infinitely many solutions. There are no other possiblities.*

*Proof.* If $A\mathbf{x} = \mathbf{b}$ is a system of linear equations, then it must have zero, one, or more than one solution. Assume that it has more than one solution. Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be two different solutions.

Let $\mathbf{x}_0 = \mathbf{x}_1 - \mathbf{x}_2$. Since $\mathbf{x}_1$ and $\mathbf{x}_2$ are different, $\mathbf{x}_0 \neq \mathbf{0}$.

$$A\mathbf{x}_0 = A(\mathbf{x}_1 - \mathbf{x}_2) = A\mathbf{x}_1 - A\mathbf{x}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

Now let $k$ be any scalar. Then

$$A(\mathbf{x}_1 + k\mathbf{x}_0) = A\mathbf{x}_1 + kA\mathbf{x}_0 = \mathbf{b} + k\mathbf{0} = \mathbf{b}$$

This means that $\mathbf{x}_1 + k\mathbf{x}_0$ is a solution for *any* scalar $k$. Thus there are infinitely many solutions $\square$

We can solve multiple systems at once by doing $[A|\mathbf{b}_0|\mathbf{b}_1| \cdots |\mathbf{b}_k]$.

### 1.6.2 Consistency

Determine which values of $\mathbf{b}$ make the following consistent.

**Ex 8.**

$$\begin{cases} x_1 - 2x_2 - x_3 = b_1 \\ -4x_1 + 5x_2 + 2x_3 = b_2 \\ -3x_1 + 3x_2 + x_3 = b_3 \end{cases}$$

### 1.6.3 Homework

#6, 15

## 1.7 Diagonal, Triangular, and Symmetric Matrices

### 1.7.1 Diagonal Matrices

**Definition.** A square matrix in which all the entries off the main diagonal are zero is a *diagonal matrix*.

Compute an inverse and do some powers.

### 1.7.2 Triangular Matrices

**Definition.** A square matrix in which all the entries above (or to the right of) the main diagonal are zero is *lower triangular*. If all the entries below (or to the left of) the main diagonal are zero is *upper triangular*.

**Theorem.**

1. *The transpose of a lower triangular matrix is upper triangular, and vice versa.*
2. *The product of lower triangular matrices is lower triangular. Same for upper.*
3. *A triangular matrix is invertible iff its diagonal entries are all nonzero.*
4. *The inverse of an invertible lower triangular matrix is lower triangular. Same for upper.*

### 1.7.3 Symmetric Matrices

**Definition.** A square matrix $A$ is *symmetric* if $A = A^T$.

**Theorem.** *If $A$ and $B$ are symmetric matrices with the same size, and if $k$ is any scalar, then*

1. *$A^T$ is symmetric.*
2. *$A + B$ and $A - B$ are symmetric.*
3. *$kA$ is symmetric.*
4. *$AB$ is symmetric iff $AB = BA$.*
5. *If $A$ is invertible, then $A^{-1}$ is also symmetric.*

### 1.7.4 Homework

# 32, 37

# Chapter 2

# Determinants

## 2.1 Determinants by Cofactor Expansion

### 2.1.1 Minors and Cofactors

A $2{\times}2$ matrix, $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, is invertible iff $ad - bc \neq 0$. The expression $ad - bc$ is the *determinant* of $A$. We write

$$\det(A) = ad - bc \text{ or } \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

**Definition.** If $A$ is a square matrix, then the *minor of entry* $a_{ij}$ is denoted by $M_{ij}$ and is defined to be the determinant of the submatrix that remains after the $i$th row and $j$th column are removed from $A$.

The number $(-1)^{i+j} M_{ij}$ is denoted by $C_{ij}$ and is called the *cofactor of entry* $a_{ij}$.

**Ex 9.** Write a 3x3 and find some minors.

**Theorem.** *If $A$ is an $n \times n$ matrix, then regardless of which row or column of $A$ is chosen, the number obtained by multiplying the entries in that row or column by the corresponding cofactors and adding the resulting products is always the same.*

### 2.1.2 The Determinant

**Definition.** If $A$ is an $n \times n$ matrix, then the number obtained from Theorem 2.1.1 is the *determinant of $A$*. The sums themselves are called *cofactor expansions of $A$*.

**Ex 10.** Do a basic 3x3 example. Do a 4x4 example where one column has a bunch of zeroes. Do an upper triangular matrix.

**Theorem.** *If $A$ is an $n \times n$ triangular matrix, then* $\det(A)$ *is the product of the entries on the diagonal.*

### 2.1.3 Homework

# 15, 23,33

## 2.2 Evaluating Determinants by Row Reduction

### 2.2.1 Basic Theorems

**Theorem.** *Let $A$ be a square matrix. If $A$ has a row or column of zeroes, then* $\det(A) = 0$.

**Theorem.** *Let $A$ be a square matrix. Then* $\det\left(A^T\right) = \det(A)$.

### 2.2.2 Elementary Row Operations

**Theorem.** *Let $A$ be an $n \times n$ matrix.*

1. *If $B$ is the matrix that results when a single row or column is multiplied by a scalar $k$, then* $\det(B) = k\det(A)$.
2. *If $B$ is the matrix that results when two rows or columns of $A$ are swapped, then* $\det(B) = -\det(A)$.
3. *If $B$ is the matrix that results when a multiple of one row of $A$ is added to another row or when a multiple of one column is added to another column, then* $\det(B) = \det(A)$.

Show this via some examples.

Show how this applies to elementary matrices.

Evaluate some by row reduction. Mention that using cofactor expansion to calculate the determinant of an $n \times n$ matrix takes $\sum_{k=1}^{n-1} \frac{n!}{k!}$ operations.

### 2.2.3 Homework

#10,14

## 2.3 Properties of Determinants

### 2.3.1 $\det(kA)$

**Theorem.** $\det(kA) = k^n \det(A)$

### 2.3.2 $\det(A + B)$

**Ex 11.**

$$A = \begin{bmatrix} 1 & 2 \\ 4 & -1 \end{bmatrix}, B = \begin{bmatrix} 3 & -1 \\ 0 & 1 \end{bmatrix}$$

**Ex 12.**

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, B = \begin{bmatrix} a_{11} & a_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

**Theorem.** *Let $A$, $B$, and $C$ be $n \times n$ matrices that differ only in the $r^{th}$ row. Assume that the $r^{th}$ row of $C$ can be obtained by adding corresponding entries of the $r^{th}$ rows of $A$ and $B$. Then*

$$\det(C) = \det(A) + \det(B)$$

Maybe skip this one.

### 2.3.3 $\det(AB)$

**Lemma.** *If $B$ is an $n \times n$ matrix and $E$ is an $n \times n$ elementary matrix, then* $\det(EB) = \det(E)\det(B)$.

*Proof.* If $E$ results from multiplying a row of $I_n$ by $k$, then $EB$ results by multipling that same row in $B$ by $k$. This means that

$$\det(EB) = k\det(B)$$

But we also know that $\det(E) = k\det(I_n) = k$, so

$$\det(EB) = \det(E)\det(B)$$

The other cases are the same. □

Mention writing matrices as the product of $E_i$. But what about noninvertible matrices?

**Theorem.** *A square matrix $A$ is invertible if and only if* $\det(A) \neq 0$.

*Proof.* Let $R$ be the reduced row echelon form of $A$ and let $E_1, E_2, \ldots, E_r$ be the elementary matrices that correspond to the row operations that produce $R$. So

$$R = E_r \cdots E_2 E_1 A$$

and so

$$\det(R) = \det(E_r) \cdots \det(E_2)\det(E_1)\det(A)$$

**Case 1:** Assume that $A$ is invertible. Then we know that $R = I$ and so $\det(R) = 1$. Thus, $\det(A) \neq 0$.
**Case 2:**Assume that $\det(A) \neq 0$. This means that $\det(R) \neq 0$. This also tells us that $R$ cannot have a row of zeroes. From a previous theorem, this means that $R = I$ and hence $A$ is invertible. □

Do an example.

**Theorem.** *If $A$ and $B$ are square matrices of the same size, then*

$$\det(AB) = \det(A)\det(B)$$

*Proof.* **Case 1:** Assume $A$ is not invertible. Then $AB$ is also not invertible. By the previous theorem, $\det(AB) = 0$ and since $\det(A) = 0$ we can say that $\det(AB) = \det(A)\det(B)$

**Case 2:** Assume that $A$ is invertible. Then we can write it as

$$A = E_1 E_2 \cdots E_r$$

so

$$AB = E_1 E_2 \cdots E_r B$$

Using the lemma, we can say that

$$\begin{aligned}\det(AB) &= \det\left(E_1\right)\det\left(E_2\right)\cdots\det\left(E_r\right)\det(B)\\ &= \det\left(E_1 \cdots E_r\right)\det(B)\\ &= \det(A)\det(B)\end{aligned}$$

□

**Theorem.** *If $A$ is invertible, then*

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

*Proof.* $A^{-1}A = I$, so $\det\left(A^{-1}A\right) = \det(I) = 1$. So $\det\left(A^{-1}\right)\det(A) = 1$. Since $A$ is invertible, $\det(A) \neq 0$ so we can divide. □

**Theorem** (Equivalent Statements)**.** *If $A$ is an $n \times n$ matrix, then the following statements are equivalent.*

*(a)* *$A$ is invertible.*

*(b)* $A\mathbf{x} = \mathbf{0}$ *has only the trivial solution.*

*(c)* *The reduced row echelon form of $A$ is $I$.*

*(d)* *$A$ can be expressed as a product of elementary matrices.*

*(e)* $A\mathbf{x} = \mathbf{b}$ *is consistent for every $n \times 1$ matrix* $\mathbf{b}$.

*(f)* $A\mathbf{x} = \mathbf{b}$ *has exactly one solution for every $n \times 1$ matrix* $\mathbf{b}$.

*(g)* $\det(A) \neq 0$.

### 2.3.4 Homework

# 9,10,15

# Chapter 3

# Euclidean Vector Spaces

## 3.1 Vectors in 2-Space, 3-Space and $n$-Space

**Definition.** Define 2-space and 3-space. Talk about vectors as arrows with *direction* and *length/magnitude*. These are *geometric vectors*. Talk about *initial point* and *terminal point*

Notation: $\mathbf{a}, \mathbf{b}, \mathbf{v}, \mathbf{w}, \mathbf{u}, \mathbf{x}$ or $\vec{x}$

**Definition.** Vectors with the same direction and length are *equivalent*. This is how we define equality, so $\mathbf{v} = \mathbf{w}$.

**Definition.** Parallelogram definition for addition. Then do triangle definition.

Either way we get:

$$\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$$

**Definition.** Define addition as translation.

**Definition.** Define *negative* of a vector and subtraction.

**Definition.** Define scalar multiplication. Talk about $|k|$ and direction. If $k = 0$ or $\mathbf{v} = \mathbf{0}$, then $k\mathbf{v} = \mathbf{0}$. $\mathbf{0}$ has no direction, but we may sometimes think of it as parallel to all vectors.

**Definition.** Vector addition is associative.

**Definition.** Define components in terms of vectors that start at the origin. Then talk about equality of vectors.

**Definition.** Define the vector between two points.

### 3.1.1 $n$-space

**Definition.** If $n$ is a positive integer, then an *ordered $n$-tuple* is a sequence of $n$ real numbers $(v_1, v_2, \dots, v_n)$. The set of all ordered $n$-tuples is called *$n$-space* and is denoted by $R^n$.

Give some examples of non-physics related $n$-tuples.

Define $\mathbf{v}$ notation and $\mathbf{0}$.

**Definition.** Define equality. Define component-wise operations.

**Theorem.** *Show commutativity, associativity, and scalar multiplication stuff.*

**Theorem.** $0\mathbf{v} = \mathbf{0}$, $k\mathbf{0} = \mathbf{0}$, $(-1)\mathbf{v} = -\mathbf{v}$

**Definition.** If $\mathbf{w}$ is a vector in $R^n$, then $\mathbf{w}$ is said to be a *linear combination* of the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ in $R^n$ if it can be expressed as

$$\mathbf{w} = k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_r\mathbf{v}_r$$

where $k_i$ are scalars. The scalars are *coefficients* of the linear combination.

### 3.1.2 Homework

#8, 14,26

## 3.2 Norm, Dot Product, and Distance in $\mathbb{R}^\mathbf{n}$

### 3.2.1 Norm

Talk about norm and magnitude in 2- and 3-space.

**Definition.** If $\mathbf{v} = (v_1, v_2, \dots, v_n)$ is a vector in $R^n$, then the *norm* of $\mathbf{v}$ (also called the *length* or *magnitude*) is denoted by $\|\mathbf{v}\|$ and is defined by the formula

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

**Theorem.** *If* $\mathbf{v}$ *is a vector in* $\mathbb{R}^n$ *and if* $k$ *is any scalar, then*

- $\|\mathbf{v}\| \geq 0$
- $\|\mathbf{v}\| = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$
- $\|k\mathbf{v}\| = |k|\|v\|$

**Definition.** A vector of norm 1 is a *unit vector*.

**Theorem.** *If* $\mathbf{v}$ *is any nonzero vector, then* $\mathbf{u} = \frac{1}{\|v\|}\mathbf{v}$ *is the unit vector in the same direction as* $\mathbf{v}$.

Normalize some vectors.

**Definition.** Define the standard unit vectors: $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$. Define the *standard unit vectors in* $\mathbb{R}^n$: $e_i$.

### 3.2.2 Distance

Define distance in terms of the norm.

### 3.2.3 Dot Product

Define the angle $\theta$ between two vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$. Note that $0 \leq \theta \leq \pi$.

**Definition.** If $\mathbf{u}$ and $\mathbf{v}$ are nonzero vectors in 2- or 3-space and if $\theta$ is the angle between $\mathbf{u}$ and $\mathbf{v}$, then the *dot product* (also called the *Euclidean inner product*) of $\mathbf{u}$ and $\mathbf{v}$ is denoted by $\mathbf{u} \cdot \mathbf{v}$ and is defined as

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}| \cos \theta$$

If $\mathbf{u} = \mathbf{0}$ or $\mathbf{v} = \mathbf{0}$, then $\mathbf{u} \cdot \mathbf{v} = 0$.

**Ex 13.** Find the angle of $(x, x, x)$ and $(0, 0, 1)$.

Let $\mathbf{u}$ and $\mathbf{v}$ be two vectors in 2-space with angle $\theta$. Let $\mathbf{u}$ and at $P$ and $\mathbf{v}$ at $Q$. Then using the law of cosines

$$|PQ|^2 = |\mathbf{u}|^2 + |\mathbf{v}|^2 - 2|\mathbf{u}||\mathbf{v}| \cos \theta$$

So

$$\begin{aligned}
|\mathbf{u}||\mathbf{v}| \cos \theta =& \frac{1}{2}\left(|\mathbf{u}|^2 + |\mathbf{v}|^2 - |PQ|^2\right) \\
=& \frac{1}{2}\left(|\mathbf{u}|^2 + |\mathbf{v}|^2 - |\mathbf{v} - \mathbf{u}|^2\right) \\
\mathbf{u} \cdot \mathbf{v} =& \frac{1}{2}\left(u_1^2 + u_2^2 + v_1^2 + v_2^2 - (v_1 - u_1)^2 - (v_2 - u_2)^2\right) \\
=& \frac{1}{2}\left(u_1^2 + u_2^2 + v_1^2 + v_2^2 - v_1^2 + 2v_1u_1 - u_1^2 - v_2^2 + 2v_2u_2 - u_2^2\right) \\
\mathbf{u} \cdot \mathbf{v} =& u_1v_1 + u_2v_2
\end{aligned}$$

**Definition.** If $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$ are vectors in $\mathbb{R}^n$, then the *dot product* or *inner product* of $\mathbf{u}$ and $\mathbf{v}$ is

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^{n} u_i v_i$$

**Theorem.** *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* $\mathbb{R}^n$ *and* $k$ *is any scalar, then*

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
2. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
3. $k(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v}$
4. $\mathbf{v} \cdot \mathbf{v} \geq 0$ *and* $\mathbf{v} \cdot \mathbf{v} = 0$ *if and only if* $\mathbf{v} = 0$.
5. $\mathbf{0} \cdot \mathbf{v} = 00$
6. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
7. $k(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (k\mathbf{v})$

### 3.2.4 Geometry

**Theorem** (Cauch-Schwarz Inequality)**.** *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* $\mathbb{R}^n$*, then*

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

**Theorem** (Triangle Inequality)**.** *If* $\mathbf{u}$*,* $\mathbf{v}$*, and* $\mathbf{w}$ *are vectors in* $\mathbb{R}^n$*, then:*

- $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$
- $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$

**Theorem** (Parallelogram Equation)**.** *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* $\mathbb{R}^n$*, then*

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2(|\mathbf{u}|^2 + |\mathbf{v}|^2)$$

**Theorem.** *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* $\mathbb{R}^n$*, then*

$$\mathbf{u} \cdot \mathbf{v} = \frac{1}{4}|\mathbf{u} + \mathbf{v}|^2 - \frac{1}{4}|\mathbf{u} - \mathbf{v}|^2$$

### 3.2.5 Matrices

If $\mathbf{u}$ and $\mathbf{v}$ are column matrices, then $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u}$.

If they're row matrices/vectors, then $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}\mathbf{v}^T = \mathbf{v}\mathbf{u}^T$.

### 3.2.6 Homework

# 12, 16, 24, 26

## 3.3 Orthogonality

Recall that $\theta = \arccos\left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|}\right)$

**Definition.** Two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are said to be *orthogonal* if $\mathbf{u} \cdot \mathbf{v} = 0$. The zero vector is orthogonal to every vector. A nonempty set of vectors in $\mathbb{R}^n$ is called an *orthogonal set* if all pairs of distinct vectors in the set are orthogonal. An orthogonal set of unit vectors is an *orthonormal set*.

### 3.3.1 Lines and Planes

Define them in terms of a point and a normal vector.

### 3.3.2 Projection

**Theorem.** *If* $\mathbf{u}$ *and* $\mathbf{a}$ *are vectors in* $\mathbb{R}^n$*, and if* $\mathbf{a} \neq 0$*, then* $\mathbf{u}$ *can be expressed in exactly one way of the form* $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$ *where* $\mathbf{w}_1$ *is a scalar multiple of* $\mathbf{a}$ *and* $\mathbf{w}_2$ *is orthogonal to* $\mathbf{a}$.

*Proof.* We know that $\mathbf{w}_1 = k\mathbf{a}$, where $k$ is a scalar. This means we want

$$\mathbf{u} = k\mathbf{a} + \mathbf{w}_2$$

To find $k$, we do

$$\begin{aligned}
\mathbf{u} \cdot \mathbf{a} =& \left(k\mathbf{a} + \mathbf{w}_2\right) \cdot \mathbf{a} \\
=& k(\mathbf{a} \cdot \mathbf{a}) + \mathbf{w}_2 \cdot \mathbf{a} \\
=& k\|\mathbf{a}\| \\
k =& \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}
\end{aligned}$$

This is the *only* possible value for $k$. Thus

$$\begin{aligned}
\mathbf{w}_2 =& \mathbf{u} - \mathbf{w}_1 \\
=& \mathbf{u} - k\mathbf{a} \\
=& \dots
\end{aligned}$$

□

**Definition.**
- $\operatorname{proj}_{\mathbf{a}} \mathbf{u} = \frac{\mathbf{u} \cdot \mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a}$ is the vector component of $\mathbf{u}$ along $\mathbf{a}$
- Do the orthogonal.

Do examples

### 3.3.3 Homework

# 22, 28

## 3.4 Geometry of Linear Systems

### 3.4.1 Lines

Let $L$ be a line that contains the point $\mathbf{x}_0$ and is parallel to $\mathbf{v}$.If $\mathbf{x}$ is any point on $L$, then

$$\mathbf{x} = \mathbf{x}_0 + t\mathbf{v}$$

. We let $t$ vary from $-\infty$ to $\infty$ and it is a *parameter*.

**Definition.** If $\mathbf{x}_0$ and $\mathbf{v}$ are vectors in $\mathbb{R}^n$, and if $\mathbf{v} \neq \mathbf{0}$, then the equation

$$\mathbf{x} = \mathbf{x}_0 + t\mathbf{v}$$

defines the line through $\mathbf{x}_0$ that is parallel to $\mathbf{v}$.

**Definition.** If $\mathbf{x}_0$, $\mathbf{v}_1$, and $\mathbf{v}_2$ are vectors in $\mathbb{R}^n$ and if $\mathbf{v}_1$ and $\mathbf{v}_2$ are not collinear, then the equation

$$\mathbf{x} = \mathbf{x}_0 + t_1\mathbf{v}_1 + t_2\mathbf{v}_2$$

defines the plane through $\mathbf{x}_0$ that is parallel to $\mathbf{v}_1$ and $\mathbf{v}_2$.

These are *vector forms* of a line and plane in $\mathbb{R}^n$.

**Ex 14.** $\mathbf{x}_0 = (-2, 3)$ and $\mathbf{v} = (1, 3)$.

Now do one about a plane.

**Ex 15.** Find the parametric equations of $3x - y - 2z = 7$

Do some 4d examples.

**Definition.** If $\mathbf{x}_0$ and $\mathbf{x}_1$ are distinct points in $\mathbb{R}^n$, then the line determined by these points is parallel to $\mathbf{v} = \mathbf{x}_1 - \mathbf{x}_0$, so

$$\mathbf{x} = \mathbf{x}_0 + t(\mathbf{x}_1 - \mathbf{x}_0)$$

or

$$\mathbf{x} = (1 - t)\mathbf{x}_0 + t\mathbf{x}_1$$

Do some examples.

Do line segment $(0 \leq t \leq 1)$

### 3.4.2 Linear Systems

Write a linear equation as a dot product. Look specifically at $= 0$.

**Theorem.** *If $A$ is an $m \times n$ matrix, then the solution set of the homogeneous system $A\mathbf{x} = \mathbf{0}$ consists of all vectors in $\mathbb{R}^n$ that are orthogonal to every row vector of $A$.*

**Ex 16.** $A = \begin{bmatrix} 3 & 2 & -1 \\ 6 & 4 & -2 \\ -3 & -2 & 1 \end{bmatrix}$ and $\mathbf{b} = (2, 4, -2)$. Do general solution of the homogeneous first. Find general solution to $A\mathbf{x} = \mathbf{b}$. Write it as homogeneous plus specific solution.

**Theorem.** *The general solution of a consistent linear system $A\mathbf{x} = \mathbf{b}$ can be obtained by adding any specific solution of $A\mathbf{x} = \mathbf{b}$ to the general solution to $A\mathbf{x} = \mathbf{0}$.*

# Chapter 4

# General Vector Spaces

## 4.1 Real Vector Spaces

**Definition.** Let $V$ be an arbitrary nonempty set of objects on which two operations are defined: addition and scalar multiplication. If the following axioms are satisfied by all objects $\mathbf{u}$ and $\mathbf{v}$ in $V$ and all scalars $k$ and $m$, then we call $V$ a *vector space* and elements of $V$ *vectors*.

1. If $\mathbf{u}$ and $\mathbf{v}$ are elements of $V$, then $\mathbf{u}+\mathbf{v}$ is in $V$.
2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$
3. $\mathbf{u}+(\mathbf{v}+\mathbf{w})=(\mathbf{u}+\mathbf{v})+\mathbf{w}$
4. There is an object $\mathbf{0}$ in $V$, called a *zero vector* for $V$, such that $\mathbf{0}+\mathbf{u}=\mathbf{u}$ for all $\mathbf{u}\in V$.
5. For each $\mathbf{u}$ in $V$, there is an element $-\mathbf{u}$ in $V$, called the *negative* of $\mathbf{u}$, such that $\mathbf{u}+(-\mathbf{u})=(-\mathbf{u})+\mathbf{u}=\mathbf{0}$.
6. If $k$ is any scalar and $\mathbf{u}$ is an element of $V$, then $k\mathbf{u}$ is in $V$.
7. $k(\mathbf{u}+\mathbf{v})=k\mathbf{u}+k\mathbf{v})$
8. $(k+m)\mathbf{u}=k\mathbf{u}+m\mathbf{u}$
9. $k(m\mathbf{u})=(km)(\mathbf{u})$
10. $1\mathbf{u}=\mathbf{u}$

**Ex 17.**

- $V=\{\mathbf{0}\}$
- $V=\mathbb{R}^n$
- $V=\mathbb{R}^\infty$
- Matrices
- Real-valued functions: $F(-\infty,\infty)$

- (not) $V = \mathbb{R}^2$, $\mathbf{u} + \mathbf{v} = \left(u_1 + v_1, u_2 + v_2\right)$, and $k\mathbf{u} = \left(0, ku_2\right)$
- The set of all invertible 2x2 matrices.
- $(1, x)$, $\mathbf{u} + \mathbf{v} = \left(1, u_1 + v_1\right)$, $k\mathbf{u} = \left(1, u_1\right)$

### 4.1.1 Homework

# 2, 4, 6

## 4.2 Subspaces

**Definition.** A (proper) subset $W$ of a vector space $V$ is called a *subspace* of $V$ if $W$ is itself a vector space under the addition and scalar multiplication of $V$.

Note: Only need to verify 1, 4, 5, and 6.

**Ex 18.**

- The zero subspace
- Lines through the origin in $\mathbb{R}^2$ and $\mathbb{R}^3$
- Planes through the origin in $\mathbb{R}^3$.
- (not) Upper half-plane
- Symmetric matrices.
- Continuous functions: $C(-\infty, \infty)$
- Polynomials of degree $n$. Of degree $\leq n$.

**Theorem.** *If $W_1, W_2, \ldots, W_r$ are subspaces of $V$, then the intersection of these spaces is also a subspace.*

**Definition.** If $\mathbf{w}$ is a vector in $V$, then $\mathbf{w}$ is said to be a *linear combination* of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r$ if we can write

$$\mathbf{w} = k_1\mathbf{v}_1 + \cdots + k_r\mathbf{v}_r$$

where $k_i$ are scalars. The scalars are *coefficients*

**Theorem.** *If $S = \left\{\mathbf{w}_1, \ldots, \mathbf{w}_r\right\}$ is a nonempty set of vectors in $V$, then*

1. *The set $W$ of all linear combinations of vectors in $S$ is a subspace of $V$.*
2. *The set $W$ in part* (1) *is the "smalles" subspace of $V$ that contains all of the vectors in $S$.*

**Definition.** The subspace of a vector space $V$ that is formed from all possible linear combinations of the vectors in a nonempty set $S$ is called the *span of $S$* and we say that the vectors of $S$ *span* that subspace.

### 4.2.1 Homework

#3, 9, 16

## 4.3 Linear Independence

### 4.3.1 Homework

# 2, 10, 12

## 4.4 Coordinates and Bases

Do an example of coordinates with a different basis in $\mathbb{R}^2$.

**Definition.** If $V$ is any vector space and $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a finite set of vectors in $V$, then $S$ is called a *basis* for $V$ if the following two conditions hold:

1. $S$ is linearly independent.
2. $S$ spans $V$.

**Ex 19.** Standard basis for $\mathbb{R}^n$ and $P_n$.

**Ex 20.** $\mathbf{u} = (2, 3)$ and $\mathbf{v} = (4, 5)$ for $\mathbb{R}^2$.

**Ex 21.** $\mathbf{v}_1 = (1, 2, 1)$, $\mathbf{v}_2 = (2, 9, 0)$, and $\mathbf{v}_3 = (3, 3, 4)$

**Ex 22.** $P_\infty$ has no finite basis. It is *infinite dimensional*

**Theorem.** *If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for a vector space $V$, then every vector $\mathbf{v}$ in $V$ can be expressed uniquely as $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$.*

**Definition.** If $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis for $V$, and

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

is the expression for $\mathbf{v}$ in terms of $S$, then the scalars $c_1, \ldots, c_n$ are called the *coordinates* of $\mathbf{v}$ relative to $S$. The vector $(c_1, \ldots, c_n)$ in $\mathbb{R}^n$ is the *coordinate vector of* $\mathbf{v}$ *relative to* $S$. It is denoted by

$$(\mathbf{v})_S = (c_1, \ldots, c_n)$$

**Ex 23.** Go back to previous examples.

### 4.4.1 Homework

# 5, 11

## 4.5 Dimension

**Theorem.** *Let $V$ be a finite-dimensional vector space and let $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ be any basis.*

1. *If a set has more than $n$ vectors, then it is linearly dependent.*
2. *If a set has fewer than $n$ vectors, then it does not span $V$.*

**Definition.** The *dimension* of a finite-dimensional vector space $V$ is denoted by $\dim(V)$ and is defined to be the number of vectors in a basis for $V$. The zero vector space is defined to have dimension 0.

**Ex 24.**

$$\dim\left(\mathbb{R}^n\right) = n$$
$$\dim\left(P_n\right) = n + 1$$
$$\dim\left(M_{m\times n}\right) = mn$$

**Ex 25.** Make up an infinite solution.

**Theorem** (Plus/Minus Theorem)**.** *Let $S$ be a nonempty set of vectors in $V$.*

1. *If $S$ is a linearly independent set, and if $\mathbf{v}$ is a vector in $V$ that is outside of* $\text{span}(S)$*, then the set $S \cup \{\mathbf{v}\}$ that results by inserting $\mathbf{v}$ into $S$ is still linearly independent.*
2. *If $\mathbf{v}$ is a vector in $S$ that is expressible as a linear combination of other vectors in $S$, and if $S\backslash\{\mathbf{v}\}$ is the set obtained by removing $\mathbf{v}$ from $S$, then $S$ and $S\backslash\{\mathbf{v}\}$ span the same space; that is,*

$$\text{span}(S) = \text{span}(S\backslash\mathbf{v})$$

**Theorem.** *Let $S$ be a finite set of vectors in a finite-dimensional vector space $V$.*

1. *If $S$ spans $V$ but is not a basis for $V$, then $S$ can be reduced to a basis for $V$ by removing appropriate vectors from $S$.*
2. *If $S$ is linearly independent but is not a basis for $V$, then $S$ can be enlarged to a basis for $V$ by inserting appropriate vectors into $S$.*

**Theorem.** *If $W$ is a subspace of a finite-dimensional vector space $V$, then:*

1. *$W$ is finite-dimensional*
2. $\dim(W) \leq \dim(V)$
3. *$W = V$ iff* $\dim(W) = \dim(V)$

## 4.6 Change of Basis

If $\mathbf{v}$ is a vector in a finite-dimensional vector space $V$, and if we change the basis for $V$ from $B$ to $B'$, how are the coordinate vectors $[\mathbf{v}]_B$ and $[\mathbf{v}]_{B'}$ related?

Let $B = \{\mathbf{u}_1, \mathbf{u}_2\}$ and $B' = \{\mathbf{u}_1', \mathbf{u}_2'\}$. Suppose

$$[\mathbf{u}_1']_B = \begin{bmatrix} a \\ b \end{bmatrix} \text{ and } [\mathbf{u}_2']_B = \begin{bmatrix} c \\ d \end{bmatrix}$$

in other words

$$\begin{aligned} \mathbf{u}_1' &= a\mathbf{u}_1 + b\mathbf{u}_2 \\ \mathbf{u}_2' &= c\mathbf{u}_1 + d\mathbf{u}_2 \end{aligned}$$

Now let $\mathbf{v}$ be any vector in $V$ and let

$$[\mathbf{v}]_{B'} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}$$

be the new coordinate vector. So

$$\mathbf{v} = k_1\mathbf{u}_1' + k_2\mathbf{u}_2'$$

Working our way backwards

$$\begin{aligned} \mathbf{v} &= k_1\mathbf{u}_1' + k_2\mathbf{u}_2' \\ &= k_1(a\mathbf{u}_1 + b\mathbf{u}_2) + k_2(c\mathbf{u}_1 + d\mathbf{u}_2) \\ &= (k_1a + k_2c)\mathbf{u}_1 + (k_1b + k_2d)\mathbf{u}_2 \end{aligned}$$

Thus, the old coordinate vector for $\mathbf{v}$ is

$$[\mathbf{v}]_B = \begin{bmatrix} k_1a + k_2c \\ k_1b + k_2d \end{bmatrix}$$

This means that

$$[\mathbf{v}]_B = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} a & c \\ b & d \end{bmatrix} [\mathbf{v}]_{B'}$$

If we change the basis for $V$ from $B$ to $B'$, then for each vector $\mathbf{v}$ in $V$, the old coordinate vector $[\mathbf{v}]_B$ is related to the new coordinate vector $[\mathbf{v}]_{B'}$ by the equation

$$[\mathbf{v}]_B = P[\mathbf{v}]_{B'}$$

where the columns of $P$ are the coordinate vectors of the new basis relative to the old basis.

**Definition.** $P$ is the *transition matrix* from $B'$ to $B$. We often write it as $P_{B' \to B}$

**Ex 26.** Let $B = \{(1,0), (0,1)\}$ and $B' = \{(1,1), (2,1)\}$. Find $P_{B' \to B}$ and $P_{B \to B'}$

Talk about $P_{B' \to B} P_{B \to B'} = P_{B \to B} = I$

**Theorem.** *If $P$ is the transition matrix from a basis $B'$ to a basis $B$ for a finite-dimensional vector space $V$, then $P$ is invertible and $P^{-1}$ is the transition matrix from $B$ to $B'$.*

$[\text{new basis}|\text{old basis}] \to [I|\text{transition from old to new}]$

Redo the previous example.

### 4.6.1 Homework

#2, 8, 16

## 4.7 Row Space, Column Space, Null Space

**Definition.** Define row and column vectors

**Definition.** If $A$ is an $m \times n$ matrix, then the subspace of $\mathbb{R}^n$ spanned by the row vectors of $A$ is called the *row space* of $A$. The subspace of $\mathbb{R}^m$ spanned by the column vectors of $A$ is called the *column space* of $A$. The solution space of the homogeneous system $A\mathbf{x} = \mathbf{0}$, which is a subspace of $\mathbb{R}^n$, is the *null space* of $M$.

**Theorem.** *A system of linear equations $A\mathbf{x} = \mathbf{b}$ is consistent iff $\mathbf{b}$ is in the column space of $A$.*

*Proof.* Straightforward □

**Ex 27.** $A = \begin{bmatrix} 3 & 1 & -1 \\ 2 & -1 & 2 \\ 1 & 1 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 6 & -2 & 2 \end{bmatrix}$. $\mathbf{x} = \begin{bmatrix} 1 & 2 & -2 \end{bmatrix}$

**Theorem.** *If $\mathbf{x}_0$ is any solution of a consistent linear system $A\mathbf{x} = \mathbf{b}$ and if $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is a basis for the null space of $A$, then every solution of $A\mathbf{x} = \mathbf{b}$ can be expressed as*

$$\mathbf{x} = \mathbf{x}_0 + c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \tag{4.1}$$

*Conversely, for all choices of scalars $c_1, \dots, c_k$, $\mathbf{x}$ is a solution of $A\mathbf{x} = \mathbf{b}$*

This equation (4.1) gives us the *general solution* to $A\mathbf{x} = \mathbf{b}$. $\mathbf{x}_0$ is a *particular solution* of $A\mathbf{x} = \mathbf{b}$ and the remaining bit is the *general solution* of $A\mathbf{x} = \mathbf{0}$. Talk about translations

**Theorem.** *Elementary row operations of a matrix do not change the null space.*

**Theorem.** *Elementary row operations do not change the row space of a matrix.*

NOT TRUE FOR COLUMN SPACE

**Theorem.** *If a matrix $R$ is in row echelon form, then the row vectors with the leading 1s form a basis for the row space of $R$, and the column vectors with the leading 1s of the row vectors form a basis for the column space of $R$.*

Even though we can find a basis of the row space with row ops, we can't do that for the column space. But, elementary row operations do not alter dependence relationships among the vectors.

Suppose that $\mathbf{w}_1, \dots, \mathbf{w}_k$ are linearly dependent column vectorsr of $A$. Thus there are scalars $c_1, \dots, c_k$ that are not all zero such that

$$c_1\mathbf{w}_1 + \cdots c_k\mathbf{w}_k = \mathbf{0}$$

When we do elementary row operations, these become $\mathbf{w}_1'$, etc. It turns out that

$$c_1\mathbf{w}_1' + \cdots c_k\mathbf{w}_k' = \mathbf{0}$$

**Theorem.** *If $A$ and $B$ are row-equivalent matrices, then*

1. *A given set of vectors of $A$ is linearly independent iff the corresponding columns of $B$ are linearly independent.*
2. *A given set of column vectors of $A$ form a basis of the column space iff the corresponding columns of $B$ do as well.*

### 4.7.1 Homework

# 6(a,d), 7c, 12b

## 4.8 Rank, Nullity, and Fundamental Matrix Spaces

**Theorem.** *The row space and column space of a matrix $A$ have the same dimension.*

**Definition.** The *rank* of a matrix $A$, denoted $\operatorname{rank}(A)$, is the dimension of the row/column space. The dimension of the null space of $A$ is the *nullity*.

### 4.8.1 Homework

# 2ac, 9

## 4.9 Matrix Transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$

Go over function, image, value, domain, codomain, range.

**Definition.** If $V$ and $W$ are vector spaces, and if $f$ is a function with domain $V$ and codomain $W$, we say that $f$ is a *transformation* from $V$ to $W$ or that $f$ *maps* $V$ to $W$.

$$f : V \to W$$

If $V = W$, $f$ is also called an *operator*.

Build one out of other functions. Talk about linear and thus $\mathbf{w} = T(\mathbf{x})$ becomes $\mathbf{w} = A\mathbf{x}$. This is a *matrix transformation/operator*. We denote it as $T_A$ and $\mathbf{x} \xrightarrow{T_A} \mathbf{w}$. $A$ is the *standard matrix*

Give a brief example.

The standard matrix of $T$ is $[T]$

**Theorem.**

1. $T_A(\mathbf{0}) = \mathbf{0}$

2. $T_A(k\mathbf{u}) = kT_A(\mathbf{u})$
3. $T_A(\mathbf{u} + \mathbf{v}) = T_A(\mathbf{u}) + T_A(\mathbf{v})$
4. $T_A(\mathbf{u} - \mathbf{v}) = T_A(\mathbf{u}) - T_A(\mathbf{v})$

Do reflections about axes in 2 and 3. Project onto axes and planes.

**Definition.** $R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

Do contractions and dilations.

Do compressions and expansions (one direction)

Shear $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$

Project onto the line with angle $\theta$, $P_\theta = \begin{bmatrix} \cos^2\theta & \frac{1}{2}\sin 2\theta \\ \frac{1}{2}\sin 2\theta & \sin^2\theta \end{bmatrix}$

### 4.9.1 Homework

# 10ac

## 4.10 Properties of Matrix Transformations

Define composition and show $T_B \circ T_A = T_{BA}$

**Theorem.** $T : \mathbb{R}^n \to \mathbb{R}^m$ *is a matrix transformation if and only if the following relationships hold for all vectors* $\mathbf{u}$ *and* $\mathbf{v}$ *in* $\mathbb{R}^n$ *and for every scalar* $k$.

- $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
- $T(k\mathbf{u}) = kT(\mathbf{u})$

Any such transformation is also called a *linear transformation*.

### 4.10.1 Homework

# Chapter 5

# Eigenvalues and Eigenvectors

## 5.1 Eigenvalues and Eigenvectors

### 5.1.1 Eigenvalues

**Definition.** If $A$ is an $n \times n$ matrix, then a nonzero vector $\mathbf{x}$ in $\mathbb{R}^n$ is called an *eigenvector* of $A$ (or of $T_A$) if $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$. That is,

$$A\mathbf{x} = \lambda\mathbf{x}$$

for some scalar $\lambda$. The scalar $\lambda$ is an *eigenvalue* of $A$ (or $T_A$) and $\mathbf{x}$ is said to be an eigenvector corresponding to $A$.

**Theorem.** *If $A$ is an $n \times n$ matrix, then $\lambda$ is an eigenvalue of $A$ iff it satisfies*

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

*This is the* ***characteristic equation*** *of $A$.*

**Ex 28.** $A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$

**Ex 29.** $A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}$

**Ex 30.** Do a big upper-triangular.

**Theorem.** *If $A$ is an $n \times n$ triangular matrix, then the eigenvalues of $A$ are the entries on the main diagonal.*

### 5.1.2 Eigenvalues

Eigenvectors satisfy $\left(\lambda I_n - A\right)\mathbf{x} = \mathbf{0}$, so these belong to the null space. This is the *eigenspace* of $A$ corresponding to $\lambda$.

**Ex 31.** Do previous examples.

**Theorem.** *If $k$ is a positive integer, $\lambda$ is an eigenvalue of $A$, and $\mathbf{x}$ is an eigenvector for $\lambda$, then $\lambda^k$ is an eigenvalue of $A^k$ and $\mathbf{x}$ is a corresponding eigenvector.*

**Ex 32.** Do the 2×2 from before.

**Ex 33.** A square matrix $A$ is invertible iff $\lambda = 0$ is not an eigenvalue of $A$.

### 5.1.3 Homework

# 3d, 4d, 6a, 7a

## 5.2 Diagonalization

1. Given an $n \times n$ matrix $A$, does there exist an invertible matrix $P$ such that $P^{-1}AP$ is diagonal.
2. Given an $n \times n$ matrix $A$, does $A$ have $n$ linearly independent eigenvectors?

**Definition.** If $A$ and $B$ are square matrices and there is an invertible matrix $P$ such that $B = P^{-1}AP$, then $A$ and $B$ are *similar matrices*.

A matrix $A$ is said to be *diagonalizable* if it is similar to a diagonal matrix.

The following properties are *invariant under similarity*.

- Determinant
- Invertibility
- Rank
- Nullity
- Trace
- Characteristic polynomial
- Eigenvalues
- Eigenspace dimension.

**Theorem.** *If $A$ is an $n \times n$ matrix, the following statements are equivalent.*

1. *$A$ is diagonalizable*
2. *$A$ has $n$ linearly independent eigenvectors.*

*Proof.* (1) → (2): Let $P = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_n]$ and let $\lambda_i$ be the entries of $D = P^{-1}AP$.

(2) → (1): Let $P$ be made up of the eigenvectors. Let $D$ be made up of the eigenvalues. □

1. Find $n$ linearly independent eigenvectors.

2. Let $P$ be made up of the eigenvectors.
3. The matrix $P^{-1}AP$ will be diagonal and have eigenvalues $\lambda_1, \dots, \lambda_n$ corresponding to $\mathbf{x}_1, \dots, \mathbf{x}_n$.

**Ex 34.** Do previous 3×3

**Ex 35.** $A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 1 & 3 & 3 \end{bmatrix}$

**Theorem.** *If an $n \times n$ matrix $A$ has $n$ distinct eigenvalues, then $A$ is diagonalizable.*

**Ex 36.** If $A =$ 3×3 from before, find $A^8$.

### 5.2.1 Homework

# 14, 23

# Chapter 6

# Inner Product Spaces

## 6.1 Inner Products

**Definition.** An *inner product* on a real vector space $V$ is a function that associates a real number () $< \mathbf{u}, \mathbf{v} >$ with each pair of vectors in $V$ in such a way that the following axioms are satisfied for all vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in $V$ and all scalars $k$.

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
3. $\langle k\mathbf{u}, \mathbf{v} \rangle = k\langle \mathbf{u}, \mathbf{v} \rangle$
4. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v} \rangle = 0$ iff $\mathbf{v} = \mathbf{0}$.

The dot product is the Euclidean inner product.

**Definition.** If $V$ is a real inner product space, then the *norm* of a vector $\mathbf{v}$ in $V$ is denoted by $\|\mathbf{v}\|$ and is defined by

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$$

and the *distance* between two vectors is

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

A vector of norm 1 is a *unit vector*

**Theorem.** *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in a real inner product space* $V$*, and if* $k$ *is a scalar, then*

1. $\|\mathbf{v}\| \geq 0$ *with equality iff* $\mathbf{v} = \mathbf{0}$
2. $\|k\mathbf{v}\| = |k|\|\mathbf{v}\|$
3. $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$
4. $d(\mathbf{u}, \mathbf{v}) \geq 0$ *with equality iff* $\mathbf{u} = \mathbf{v}$.

**Ex 37.** Do a weighted Euclidean inner product.

### 6.1.1 Inner Products Generated by Matrices

**Definition.** For $\mathbb{R}^n$, we can do $\langle \mathbf{u}, \mathbf{v} \rangle = A\mathbf{u} \cdot A\mathbf{v}$. If $\mathbf{u}$ and $\mathbf{v}$ are in column form, then $\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{v})^T A\mathbf{u} = \mathbf{v}^T A^T A\mathbf{u}$.

**Ex 38.** Show that weighted products come from diagonal matrices.

### 6.1.2 Other Inner Products

**Ex 39.** If $U$ and $V$ are $n \times n$ matrices, then $\langle U, V \rangle = \operatorname{tr}\left(U^T V\right)$

**Ex 40.** The standard inner product on $P_n$

If $\vec{p} = a_0 + a_x x + \cdots + a_n x^n$ and $\vec{q} = b_0 + b_1 x + \cdots + b_n x^n$, then $\langle \vec{p}, \vec{q} \rangle = a_0 b_0 + a_1 b_1 + \cdots + a_n b_n$.

**Ex 41.** Let $\vec{f} = f(x)$ and $\vec{g} = g(x)$ be two function on $C[a, b]$. Then

$$\left\langle \vec{f}, \vec{g} \right\rangle = \int_a^b f(x) g(x)\, dx$$

### 6.1.3 Homework

#24

## 6.2 Angle and Orthogonality

### 6.2.1 Cauchy-Schwarz

Recall:

$$\theta = \arccos\left(\frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}\right)$$

**Theorem** (Cauchy-Schwarz)**.** *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in an inner product space* $V$*, then*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| = \|\mathbf{u}\| \|\mathbf{v}\|$$

*Proof.* If $\mathbf{u} = \mathbf{0}$, then $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ and $\|\mathbf{u}\| = 0$, so we have equality. Assume that $\mathbf{u} \neq \mathbf{0}$. Let

$$a = \langle \mathbf{u}, \mathbf{u} \rangle, b = 2\langle \mathbf{u}, \mathbf{v} \rangle, c = \langle \mathbf{v}, \mathbf{v} \rangle$$

Since the inner product of a vector with itself is nonnegative,

$$\begin{aligned} 0 \leq& \langle t\mathbf{u} + \mathbf{v}, t\mathbf{u} + \mathbf{v} \rangle \\ =& \langle \mathbf{u}, \mathbf{u} \rangle t^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle t + \langle \mathbf{v}, \mathbf{v} \rangle \\ =& at^2 + bt + c \end{aligned}$$

Since $at^2 + bt + c \geq 0$, it has either no real roots or a repeated real root. That means $b^2 - 4ac \leq 0$. Writing this back in vector form, we get $4\langle \mathbf{u}, \mathbf{v}\rangle^2 - 4\langle \mathbf{u}, \mathbf{u}\rangle\langle \mathbf{v}, \mathbf{v}\rangle \leq 0$. Thus

$$\langle \mathbf{u}, \mathbf{v}\rangle^2 \leq \langle \mathbf{u}, \mathbf{u}\rangle\langle \mathbf{v}, \mathbf{v}\rangle$$

But $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0$ and $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$. That means we can take the square root of both sides and get

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$$

□

This means

**Definition.**

$$\theta = \arccos\left(\frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}\right)$$

### 6.2.2 Properties of length and distance

**Theorem.** *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in an inner product space* $V$, *then*

1. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$
2. $d(\mathbf{u}, \mathbf{v}) \leq d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$

**Definition.** Two vectors $\mathbf{u}$ and $\mathbf{v}$ in an inner product space are *orthogonal* if $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

**Theorem.** *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are orthogonal vectors in an inner product space, then*

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

**Definition.** If $W$ is a subspace of an inner product space $V$, then the set of all vectors in $V$ that are orthogonal to every vector in $W$ is called the *orthogonal complement* of $W$ and is denoted $W^\perp$

### 6.2.3 Homework

# 9, 12

## 6.3 Orthogonal and Orthonormal Sets

**Definition.** Define orthogonal and orthonormal.

**Theorem.** *If $S$ is an orthogonal set of nonzero vectors in an inner product space, then $S$ is independent.*

**Theorem.**

1. *If $S$ is an orthogonal set in an inner product space $V$ and* $\mathbf{u}$ *is any vector in $V$, then*

$$\mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1\rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \cdots$$

2. *If $S$ is an orthonormal set in an inner product space $V$ and $\mathbf{u}$ is any vector in $V$, then*

$$\mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots$$

**Theorem.** *If $W$ is a finite-dimensional subspace of an inner product space $V$, then every vector $\mathbf{u}$ in $V$ can be expressed in exactly one way as*

$$\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$$

*where $\mathbf{w}_1$ is in $W$ and $\mathbf{w}_2$ is in $W^{\perp}$.*

**Theorem.** *Let $W$ be a finite-dimensional subspace of an inner product space $V$.*

1. *If $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is an orthogonal basis for $W$ and $\mathbf{u}$ is any vector in $V$, then*

$$\operatorname{proj}_W \mathbf{u} = \frac{\langle \mathbf{u}, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 + \cdots$$

2. *If $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is an orthonormal basis for $W$ and $\mathbf{u}$ is any vector in $V$, then*

$$\operatorname{proj}_W \mathbf{u} = \langle \mathbf{u}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \cdots$$

### 6.3.1 Gram-Schmidt Process

**Theorem.** *Every nonzero finite-dimensional inner product space has an orthonormal basis.*

**Definition** (The Gram-Schmidt Process)**.** To convert a basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ into an orthogonal basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$, perform the following steps.

**Step 1:** Let $\mathbf{v}_1 = \mathbf{u}_1$.

**Step 2:** Let $\mathbf{v}_2 = \mathbf{u}_2 - \frac{\langle \mathbf{u}_2, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1$

**Step 3:** Let $\mathbf{v}_3 = \mathbf{u}_3 - \frac{\langle \mathbf{u}_3, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \frac{\langle \mathbf{u}_3, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2$

**Step 4:** Let $\mathbf{v}_4 = \mathbf{u}_4 - \frac{\langle \mathbf{u}_4, \mathbf{v}_1 \rangle}{\|\mathbf{v}_1\|^2} \mathbf{v}_1 - \frac{\langle \mathbf{u}_4, \mathbf{v}_2 \rangle}{\|\mathbf{v}_2\|^2} \mathbf{v}_2 - \frac{\langle \mathbf{u}_4, \mathbf{v}_3 \rangle}{\|\mathbf{v}_3\|^2} \mathbf{v}_3$

(continue for $r$ steps)

To convert to an orthonormal basis, just normalize the vectors.

**Ex 42.** $\mathbf{u}_1 = (1, 1, 1)$, $\mathbf{u}_2 = (-1, 1, 0)$, $\mathbf{u}_3 = (1, 2, 1)$

**Ex 43.** Legendre polynomials for $P_3$. Start with $\{1, x, x^2, x^3\}$. Get $\left\{1, x, x^2 - \frac{1}{3}, x^3 - \frac{3}{5}x\right\}$. Scale to get Legendre polynomials.

### 6.3.2 QR-Decomposition

If $A$ is an $m \times n$ matrix with linearly independent column vectors and if $Q$ is the matrix that results by applying the Gram-Schmidt process to the column vectors of $A$, what relationship is there between $A$ and $Q$.

Look at this in block form.

**Theorem.** *If $A$ is an $m \times n$ matrix with linearly independent column vectors, then $A$ can be factored as*

$$A = QR$$

*where $Q$ is an $m\times n$ matrix with orthonormal column vectors and $R$ is an $n\times n$ invertible upper-triangular matrix.*

**Ex 44.** $A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix}$

## 6.4 Best Approximation & Least Squares

Suppose that $A\mathbf{x} = \mathbf{b}$ is an inconsistent linear system of $m$ equations in $n$ unknowns in which we suspect the inconsistency is caused by measurement errors. We want to find $\mathbf{x}$ that comes as close as possible to being a solution. $\|\mathbf{b} - A\mathbf{x}\|$ is minimized wrt the Euclidean norm on $\mathbb{R}^m$.

Such an $\mathbf{x}$ is the *least-squares solution* of the system. Note: it's least squares because something that minimizes $\|\mathbf{b} - A\mathbf{x}\|$ also minimizes $\|\mathbf{b} - A\mathbf{x}\|^2$

### 6.4.1 Best Approximation

Suppose that $\mathbf{b}$ is some vector in $\mathbb{R}^n$ and we would like to approximate it be a vector $\mathbf{w}$ in some subspace $W$ of $\mathbb{R}^n$.

**Theorem.** *If $W$ is a finite-dimensional subspace of an inner product space $V$, and if $\mathbf{b}$ is a vector in $V$, then $\operatorname{proj}_W \mathbf{b}$ is the **best approximation** to $\mathbf{b}$ from $W$ in the sense that*

$$\left\|\mathbf{b} - \operatorname{proj}_W \mathbf{b}\right\| < \|\mathbf{b} - \mathbf{w}\|$$

*for every vector $\mathbf{w}$ in $W$ that is not $\operatorname{proj}_W \mathbf{b}$.*

*Proof.* Using a dirty trick, we can write $\mathbf{b} - \mathbf{w}$ as $\mathbf{b} - \operatorname{proj}_W \mathbf{b} + \operatorname{proj}_W \mathbf{b} - \mathbf{w}$.

But $\operatorname{proj}_W \mathbf{b} - \mathbf{w}$ is in $W$ and $\mathbf{b} - \operatorname{proj}_W \mathbf{b}$ is in $W^\perp$, so they are orthogonal. This means we can use the generalized Pythagorean theorem to get

$$\begin{aligned} \left\|\mathbf{b} - \operatorname{proj}_W \mathbf{b}\right\|^2 + \left\|\operatorname{proj}_W \mathbf{b} - \mathbf{w}\right\|^2 &= \|\mathbf{b} - \mathbf{w}\|^2 \\ \left\|\mathbf{b} - \operatorname{proj}_W \mathbf{b}\right\|^2 &< \|\mathbf{b} - \mathbf{w}\|^2 \\ \left\|\mathbf{b} - \operatorname{proj}_W \mathbf{b}\right\| &< \|\mathbf{b} - \mathbf{w}\| \end{aligned}$$

□

### 6.4.2 Least-Squares Solutions to Linear Systems

To find the least squares solution to $A\mathbf{x} = \mathbf{b}$, we project $\mathbf{b}$ onto the column space $W$ of $A$ and instead solve

$$A\mathbf{x} = \operatorname{proj}_W \mathbf{b}$$

$$\begin{aligned}
A\mathbf{x} =& \operatorname{proj}_W \mathbf{b} \\
\mathbf{b} - A\mathbf{x} =& \mathbf{b} - \operatorname{proj}_W \mathbf{b} \\
A^T\left(\mathbf{b} - A\mathbf{x}\right) =& A^T\left(\mathbf{b} - \operatorname{proj}_W \mathbf{b}\right) \\
A^T\left(\mathbf{b} - A\mathbf{x}\right) =& \mathbf{0} \\
A^T A\mathbf{x} =& A^T\mathbf{b}
\end{aligned}$$

**Definition.** $A^T A\mathbf{x} = A^T\mathbf{b}$ is the *normal equation* or *normal system* associated with $A\mathbf{x} = \mathbf{b}$.

**Theorem.** $\mathbf{x} = \left(A^T A\right)^{-1} A^T\mathbf{b}$

**Theorem.** $Q^T Q = I$

*Proof.* The columns are all orthonormal. □

**Theorem.** *If $A$ is an $m \times n$ matrix with linearly independent column vectors, and if $A = QR$ is the $QR$-decomposition of $A$, then the least squares solution is $\mathbf{x} = R^{-1}Q^T\mathbf{b}$.*

*Proof.*

$$\begin{aligned}
\mathbf{x} =& \left(A^T A\right)^{-1} A^T\mathbf{b} \\
=& \left((QR)^T QR\right)^{-1} (QR)^T\mathbf{b} \\
=& \left(R^T Q^T QR\right)^{-1} R^T Q^T\mathbf{b} \\
=& \left(R^T R\right)^{-1} R^T Q^T\mathbf{b} \\
=& R^{-1}\left(R^T\right)^{-1} R^T Q^T\mathbf{b} \\
=& R^{-1} Q^T\mathbf{b}
\end{aligned}$$

□

# Chapter 7

# Diagonalization and Quadratic Forms

## 7.1 Orthogonal Matrices

**Definition.** A square matrix $A$ is said to be *orthogonal* if $AA^T = A^T A = I$.

**Ex 45.** Rotation matrix $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

**Theorem.** *The following are equivalent for an $n \times n$ matrix $A$.*

1. *$A$ is orthogonal.*
2. *The row vectors of $A$ form an orthonormal set in $\mathbb{R}^n$ with the Euclidean inner product.*