

Math 33A - Linear Algebra and Applications (23W)

Instructor: Chengxi Wang

Textbook: Otto Bruecher - Linear Algebra with Applications (5th Edition - 2012)

Topics: Linear systems, matrices & matrix operations, linear transformations, linear independence, base & subspaces, orthogonality, determinants, eigenvalues & eigenvectors, diagonalization

Table of Contents:

- 1) Linear Systems - 89
 - i) Matrices of Linear Systems - 91
- 2) Linear Transformations - 95
- 3) Matrix Properties & Operations - 98
 - i) Matrix Products - 98
 - ii) Matrix Inverses / Invertibility - 99
- 4) Vector Spaces - 101
 - i) Image, Span, and Kernel - 101
 - ii) Bases and Linear Independence - 103
 - iii) Bases and Dimension - 104
 - (*) Coordinates and Similarity - 105
- 5) Orthogonality - 106
 - i) Gram-Schmidt and Orthogonal Matrices - 107
- 6) Topics on Matrices - 109
 - i) Determinants - 109
 - (*) Cramer's Rule - 111
 - ii) Eigenvectors and Eigenvalues - 112
 - (*) Matrix Diagonalization - 114
- (*) Additional Notes - 116
 - i) Linear Spaces and Isomorphisms - 116
 - ii) Least Squares - 117

Linear Equations

1/2/23
Lecture 1

Linear Equation

A linear equation is an equation taking the following form:

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

• x_1, x_2, \dots, x_n represent distinct variables

• a_1, a_2, \dots, a_n, b are constants

• a_1, a_2, \dots, a_n are the coefficients of a linear equation

A solution for a linear equation represents a set of values for variables x_1, \dots, x_n such that the linear equation evaluates as true.

A system of linear equations represents a set of linear equations; a solution to a system of linear equations must cause all of the contained linear equations to evaluate as true simultaneously.

A solution may or may not be unique (the only solution).

1/11/23
Lecture 2

Solutions to [systems of] linear equations can be written as ordered tuples,
e.g. (s_1, s_2, \dots, s_n) s.t. $x_1 = s_1, \dots, x_n = s_n$ is a solution.

Methods for solving linear equations:

- (1) Substitution
- (2) Elimination

Matrices

1/13/23
Lecture 3

Linear systems are equivalent if they share the same solutions

Linear systems can be rewritten as matrices, eg.:

$$\begin{cases} x_1 - x_2 + 2x_3 = 4 \\ 2x_1 - x_2 + x_3 = 4 \\ 3x_1 + x_2 - x_3 = 1 \end{cases} \rightarrow \begin{bmatrix} 1 & -1 & 2 \\ 2 & -1 & 1 \\ 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 1 \end{bmatrix}$$

↑
constants

Matrices

A matrix is a rectangular array of numbers, arranged in rows & columns.

An individual item in a matrix is known as an "entry".

- (*) All entries in a matrix must be filled
- (*) The size of a matrix is defined by the # of rows and columns - a matrix with m rows and n columns is called $m \times n$ (has dimensions $m \times n$)

- Entries in a matrix are named by their location - " a_{ij} " refers to an entry in matrix A , row i , col. j
- Matrices A and B are equal if they share the same size, have the same corresponding entries ($a_{ij} = b_{ij} \forall i, j$)

Definitions

Square matrix - a matrix with an equal number of rows and columns

Diagonal matrix - all entries above/below the diagonals are 0 ($\{a_{ij} \mid a_{ij} = 0 \text{ if } i \neq j\}$)

- Upper triangular matrix - all entries below the diagonals are 0 (converse is lower-triangular)

(*) Diagonal and x-triangular matrices are only defined for square matrices

Vector - a matrix with either only one row (row vector), or only one column (column vector)

$\mathbb{R}^{m \times n}$ - set of all matrices with dimension $m \times n$

Matrix Algebra (§2.1)

- The sum of two matrices [of the same size only] is defined as the individual addition of corresponding entries, i.e. $A+B = C = \{c_{ij} = a_{ij} + b_{ij}\}$
- The scalar multiple of a matrix is computed by multiplying each element by a scalar, i.e. $B = \lambda A = \{b_{ij} = \lambda \cdot a_{ij}\}$
- The dot product of two vectors is defined as the scalar obtained from multiplying all pairs of corresponding entries, i.e. $\vec{v} \cdot \vec{w}$; $\vec{v}, \vec{w} \in \mathbb{R}^n = \sum_{i=1}^n v_i \cdot w_i$
- A matrix-vector product is the vector obtained from taking the dot product between a vector and each of a matrix's row/columns (where each row/column is treated as an individual vector)
 - Given matrix $A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}$, it can be written as a collection of row vectors or column vectors.

$\left[\begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \dots, \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} \right]$
 $\left[\begin{bmatrix} a_{11} & \dots & a_{1n} \end{bmatrix}, \begin{bmatrix} a_{21} & \dots & a_{2n} \end{bmatrix}, \dots, \begin{bmatrix} a_{m1} & \dots & a_{mn} \end{bmatrix} \right]$

Solving Linear Systems

1/18/23
Lecture 4

Linear Combinations

Vector $\vec{b} \in \mathbb{R}^n$ is a linear combination of vectors $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n$ if $\exists x_1, \dots, x_n \in \mathbb{R}$ s.t. $\vec{b} = x_1 \vec{v}_1 + \dots + x_n \vec{v}_n$

e.g. $\begin{bmatrix} 2 \\ 3 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$\begin{matrix} x_1 & v_1 & x_2 & v_2 \\ \end{matrix}$

Writing Linear Systems

Given a series of linear equations represented in the form $A\vec{v} = \vec{b}$, where A represents the coefficients of variables in \vec{v} and \vec{b} represents the constants of each equation, we can write $A\vec{v} = \vec{b}$ in another form:

$$A\vec{v} = \vec{b} \rightarrow [A][\vec{v}] = [\vec{b}] \rightarrow [A | \vec{b}] \text{ (Augmented matrix)}$$

(*) To solve linear systems, we can express them in the form of an augmented matrix and convert the left-hand side to an identity matrix to determine the values of the variables x_1, \dots, x_n , i.e.: $[A | \vec{b}] \rightarrow [I^n | \begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix}]$ (*Convert w/ row operations)

Solving Linear Systems (cont.)

1/20/23
Lecture 5
+ Lecture 6
(1/23/23)

Reduced Row Echelon Form (rref)

A matrix A is said to be in reduced row echelon form (rref), if:

- For every row with nonzero entries, the first entry is a 1 (called a leading 1 / pivot)
- If a column contains a leading 1, all other entries in the column are 0
- If a row contains a leading 1, then each row above it has a leading 1 further to the left

Solving Linear Systems

- Solutions for any arbitrary system of equations can be found by reducing the associated [augmented] matrix to RREF via elementary row operations

(*) Every matrix can be reduced to a unique RREF matrix

e.g.: $A = \begin{bmatrix} 1 & -3 & 0 & -5 & -7 \\ 3 & -12 & -2 & -27 & -33 \\ 2 & 10 & 2 & 24 & 29 \\ -1 & 6 & 1 & 14 & 17 \end{bmatrix} \longrightarrow \text{rref}(A) = \begin{bmatrix} 1 & 0 & 0 & 1 & 6 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \longrightarrow \begin{cases} x_1 + x_4 = 6 \\ x_2 + 2x_4 = 0 \\ x_3 + 3x_4 = 1 \end{cases} \text{ (solutions)}$
 * (not actually)

- The rank $R(A)$ of a matrix A is defined as the number of leading 1s in $\text{rref}(A)$
 - The number of solutions of a linear system depends on its rank and rref form:
 - If $\text{rref}(A)$ contains a row $[0 \ 0 \ \dots \ 0 \ | \ 1]$, the system is inconsistent (has no solutions)
 - If the system is consistent, it can either have:
 - Infinitely many solutions, if there is at least one free (i.e. non-pivot) variable ($R(A) < m$)
 - A unique solution, if all variables are leading ($R(A) = m$)
- (*) Can only occur if there are at least as many equations as unknowns ($m \geq n$)

Types of Elementary Row Operations

(1) Scalar Multiplication: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{\lambda r_1} \begin{bmatrix} \lambda & 0 \\ 0 & 1 \end{bmatrix}$ (*) (for $\lambda \neq 0$)

(2) Row Addition/Subtraction: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{r_2 + \lambda r_1} \begin{bmatrix} 1 & 0 \\ \lambda & 1 \end{bmatrix}$

(3) Swapping Rows: $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow[r_1 = r_2]{r_2 = r_1} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

Misc. Notes: Matrices

2/5/23
Notes

Writing Linear Systems

$$\left. \begin{aligned} c_1x_1 + c_2x_2 + c_3x_3 &= b_1 \\ c_4x_1 + c_5x_2 + c_6x_3 &= b_2 \\ c_7x_1 + c_8x_2 + c_9x_3 &= b_3 \end{aligned} \right\} \rightarrow x_1 \begin{bmatrix} c_1 \\ c_4 \\ c_7 \end{bmatrix} + x_2 \begin{bmatrix} c_2 \\ c_5 \\ c_8 \end{bmatrix} + x_3 \begin{bmatrix} c_3 \\ c_6 \\ c_9 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \rightarrow \begin{matrix} (A) \\ \begin{bmatrix} c_1 & c_2 & c_3 \\ c_4 & c_5 & c_6 \\ c_7 & c_8 & c_9 \end{bmatrix} \end{matrix} \begin{matrix} (\vec{x}) \\ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \end{matrix} = \begin{matrix} (\vec{b}) \\ \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \end{matrix}$$
$$\rightarrow \boxed{A\vec{x} = \vec{b}} \quad (A = \text{coefficient matrix})$$

Vectors

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = (v_1, v_2) \neq [v_1, v_2] \quad \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{matrix} y \\ x \end{matrix} \begin{matrix} \nearrow \vec{v} \\ (v_1, v_2) \end{matrix} \quad (\text{"Standard representation"})$$

Matrix-Vector Multiplication

A matrix-vector can be thought of either in terms of the rows of the matrix, or in terms of the columns.

$$(1) A\vec{x} = \vec{b} \rightarrow \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \vec{x} = \begin{bmatrix} \text{row 1} \\ \text{row 2} \end{bmatrix} \vec{x} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \vec{x} = \begin{bmatrix} r_1 \cdot \vec{x} \\ r_2 \cdot \vec{x} \end{bmatrix} = \begin{bmatrix} \langle c_1, c_2 \rangle \cdot \langle x_1, x_2 \rangle \\ \langle c_3, c_4 \rangle \cdot \langle x_1, x_2 \rangle \end{bmatrix} = \vec{b} \quad (\text{Rows})$$

$$(2) A\vec{x} = \vec{b} \rightarrow [\text{col. 1} \quad \text{col. 2}] \vec{x} = [v_1 \quad v_2] \vec{x} = x_1 \vec{v}_1 + x_2 \vec{v}_2 = x_1 \begin{bmatrix} c_1 \\ c_3 \end{bmatrix} + x_2 \begin{bmatrix} c_2 \\ c_4 \end{bmatrix} = \begin{bmatrix} x_1c_1 + x_2c_2 \\ x_1c_3 + x_2c_4 \end{bmatrix} = \vec{b} \quad (\text{Columns})$$

(*) The column picture characterizes \vec{b} as a linear combination of the columns of A w/ coeffs. x_1, x_2

Row-Echelon Form

Conditions for Row-Echelon Form (REF):

- 1) The leading coefficient (leftmost nonzero entry) of any row is to the right of the leading coefficient of the row above it
- 2) Any rows of all zeros are at the bottom of the matrix

Conditions for RREF:

- 1) Matrix is in REF
- 2) All pivots [leading coefficients] are equal to 1
- 3) All entries above a pivot are 0

Intro to Linear Transformations

1/25/23
Lecture 7

Functions Review

A function $T: X \rightarrow Y$ that maps elements of the domain (X) to elements of the target space (Y), such that for all $x \in X \exists T(x) = y$ for some $y \in Y$.

$$T: X \rightarrow Y$$

(input) $x \rightarrow y$ (output)

Linear Transformations

A function $T: \mathbb{R}^m \rightarrow \mathbb{R}^n$ is called a linear transformation if \exists a matrix $A \in \mathbb{R}^{n \times m}$ such that, for all $\vec{x} \in \mathbb{R}^m$, the following expression is true:

$$\longrightarrow \boxed{T(\vec{x}) = A\vec{x}} \quad (T \text{ is represented by a matrix } A)$$

Linear Transformations

- A function $T: \mathbb{R}^m \rightarrow \mathbb{R}^n$ takes as input a vector of size m , and outputs a vector of size n , i.e. takes a vector as input and also returns a vector
- The expression $T(\vec{x}) = A\vec{x}$ is also written $\vec{y} = A\vec{x}$ ($\vec{y} = T(\vec{x})$)

(*) Standard Bases of a Vector Space

- Given a vector space \mathbb{R}^n , we can define a set of standard bases of $\mathbb{R}^n = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$, where each vector $\vec{e}_i \in \mathbb{R}^n$ takes the form $(0, 0, \dots, 0, 1, 0, \dots, 0)$ (contains all 0s except for a 1 as the i -th component)



Matrices of Linear Transformations

Given a linear transformation $T: \mathbb{R}^m \rightarrow \mathbb{R}^n$, $T(\vec{x}) = A\vec{x}$ for some matrix A , we can write:

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

Linear Transformations (cont.)

1/27/23
Lecture 8

Properties of Linear Transformations

A transformation $T: \mathbb{R}^m \rightarrow \mathbb{R}^n$ is linear if (and only if):

- 1) $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$ for all vectors $\vec{v}, \vec{w} \in \mathbb{R}^m$ (Vector Addition)
- 2) $T(k\vec{v}) = kT(\vec{v})$ for all vectors $\vec{v} \in \mathbb{R}^m$, scalars $k \in \mathbb{R}$ (Scalar Multiplication)

(*) A transformation can only be linear if it can be expressed as a matrix

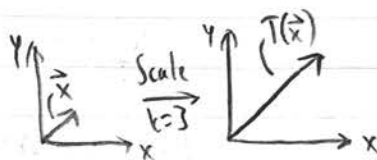
Linear Transformations in Geometry (§2.2)

- 1) Scaling - a scaling is a transformation that increases/decreases the size of an input by a certain factor k ; i.e. $\text{size}(T(\vec{x})) = k \cdot \text{size}(\vec{x})$

→ Scaling: $T(\vec{x}) = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} \vec{x}$ (Scaling by a factor of k)

(*) Notation:

- A scaling is a dilation/enlargement if $k \geq 1$
- A scaling is a contraction/shrinking if $0 < k < 1$



- 2) Orthogonal Projection - given a line $L \in \mathbb{R}^2$, any vector $\vec{x} \in \mathbb{R}^2$ can be written in the form $\vec{x} = \vec{x}'' + \vec{x}^\perp$, where \vec{x}'' is parallel to L (the projection of \vec{x} onto L) and \vec{x}^\perp is perpendicular to L . An orthogonal projection is the transformation $T(\vec{x}) = \text{proj}_L \vec{x}$ (projection of \vec{x} onto L) = \vec{x}'' .

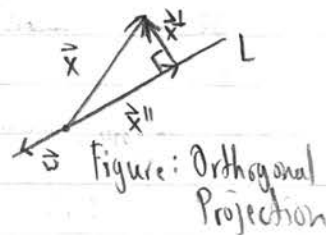
→ Given a vector \vec{w} parallel to L , we can write that:

$$\text{proj}_L(\vec{x}) = k\vec{w} \text{ (for some scalar } k) = \left(\frac{\vec{x} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \right) \vec{w},$$

which [in the case that $\vec{w} = \vec{u}$ (unit vector parallel to L)] becomes:

$$\text{proj}_L(\vec{x}) = (\vec{x} \cdot \vec{u}) \vec{u} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

(*) This equation only holds for line L that passes through the origin



→ Orthogonal Projection if $\text{proj}_L(\vec{x}) = \begin{bmatrix} u_1^2 & u_1 u_2 \\ u_1 u_2 & u_2^2 \end{bmatrix} \vec{x}$ (Orthogonal projection of \vec{x} onto L , where $\vec{u} \parallel L, \|\vec{u}\| = 1$)

Linear Transformations in Geometry

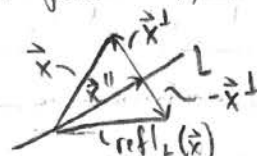
1/30/23

Lecture 9

Linear Transformations in Geometry (cont.)

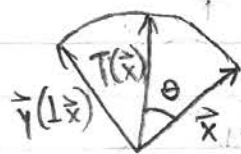
3) Reflection - Given a line L and vector $\vec{x} = \vec{x}'' + \vec{x}^\perp$ (in regards to L), the reflection of \vec{x} over L is the following:

$$\rightarrow \text{ref}_L(\vec{x}) = \vec{x}'' - \vec{x}^\perp = [2\text{proj}_L(\vec{x}) - \vec{x}]$$



$$\rightarrow \text{Reflection: } \text{ref}_L(\vec{x}) = \begin{bmatrix} 2u_1^2 - 1 & 2u_1u_2 \\ 2u_1u_2 & 2u_2^2 - 1 \end{bmatrix} \vec{x} \quad \left(\begin{array}{l} \text{Reflection of } \vec{x} \text{ over line } L, \\ \text{where } \vec{u} \text{ is a unit vector } \parallel L \end{array} \right)$$

4) Rotation - a rotation is a transformation that rotates a vector \vec{x} by a certain number of degrees (θ) about the origin.



We can characterize a rotation as a linear combination of \vec{x} and a vector perpendicular to \vec{x} , denoted \vec{y} ($\|\vec{y}\| = \|\vec{x}\|$).

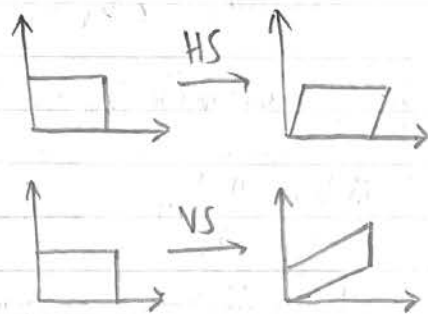
$$\rightarrow T(\vec{x}) = \cos(\theta) \cdot \vec{x} + \sin(\theta) \cdot \vec{y} = \cos(\theta) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \sin(\theta) \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$

$$\rightarrow \text{Rotation: } T(\vec{x}) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \vec{x} \quad \left(\begin{array}{l} \text{Rotation of } \vec{x} \text{ about the origin} \\ \text{by } \theta \text{ degrees} \end{array} \right)$$

5) Shears - a shear is a transformation that "slants" an input in either the y -direction (horizontal shear) or the x -direction (vertical shear).

$$\rightarrow \text{Horizontal Shear: } T(\vec{x}) = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \vec{x}$$

$$\rightarrow \text{Vertical Shear: } T(\vec{x}) = \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix} \vec{x}$$



Matrix Products

2/1/23
Lecture 10

Matrix Products

Notation: Given two functions $g(x)$ and $f(x)$, the composite of $f(x)$ and $g(x)$ is the function $z = g(f(x))$.

Matrix Products

→ Given two linear transformations $T_1(\vec{x}) = A\vec{x}$, $T_2(\vec{x}) = B\vec{x}$, we define the product of matrices A and B as the matrix of the composite of $T_1(\vec{x})$ and $T_2(\vec{x})$:

$$\longrightarrow T(\vec{x}) = T_2(T_1(\vec{x})) = B(A\vec{x}) = (BA)\vec{x} \quad (\text{Matrix product of A and B})$$

(*) The composite of two linear transformations is also a linear transformation

Note: Given two matrices $A \in \mathbb{R}^{n \times p}$ and $B \in \mathbb{R}^{q \times m}$, the product AB is only defined if $p = q$

→ The size of such a matrix AB would be $n \times m$.

Matrix Multiplication

Given matrices $B \in \mathbb{R}^{n \times p}$ and $A \in \mathbb{R}^{p \times m}$, $A = [\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m]$, the product BA is

$$\rightarrow BA = B \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_m \end{bmatrix} = \begin{bmatrix} B\vec{v}_1 & B\vec{v}_2 & \dots & B\vec{v}_m \end{bmatrix} \quad (\text{size } n \times m) \quad (\text{Product } BA \text{ of matrices A and B})$$

Properties of Matrix Multiplication

1) Matrix multiplication is noncommutative (i.e. $AB \neq BA$ (can be))

(*) If $AB = BA$ for some matrices A and B, we say A and B commute

2) Given matrix product BA , the (i, j) entry of BA is the dot product between the i^{th} row of B, j^{th} col. of A.

$$\rightarrow \text{i.e. } BA_{ij} = \sum_{k=1}^p b_{ik} a_{kj} = b_{i1} a_{1j} + b_{i2} a_{2j} + \dots + b_{ip} a_{pj} \quad (B \in \mathbb{R}^{n \times p}, A \in \mathbb{R}^{p \times m})$$

3) Given matrix $A \in \mathbb{R}^{n \times m}$, $AI_m = I_n A = A$ (product of a matrix A and the identity matrix, is A)

4) Matrix multiplication is associative, i.e. $(AB)C = A(BC)$

5) Matrix multiplication is distributive, i.e. $A(C+D) = AC+AD$

6) Scalar multiplication is commutative for matrix products, i.e. $(kA)B = A(kB) = k(AB)$

7) Block multiplication: Given a matrix partitioned into blocks, those blocks can be treated like entries when multiplying

$$\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} I_2 & I_2 \\ 0 & I_2 \end{bmatrix} \quad (\text{Block matrices})$$

Invertibility

2/3/23
Lecture 11

Transition Matrices

Definition: A vector $\vec{x} \in \mathbb{R}^n$ is a distribution vector if its components add up to 1, and all components ≥ 0 .

Definition: A matrix A is a transition matrix (stochastic matrix) if all of its columns are distribution vectors.

(*) A transition matrix is positive if all of its entries are positive (i.e. nonzero)

(*) A transition matrix A is regular (eventually positive) if A^m is positive (for some m)

→ Given regular transition matrix $A \in \mathbb{R}^{n \times n}$:

• There exists a unique equilibrium distribution vector \vec{x}_{equ} , s.t. $A\vec{x}_{\text{equ}} = \vec{x}_{\text{equ}}$.

(*) All elements of \vec{x}_{equ} are positive.

• Given any distribution vector $\vec{x} \in \mathbb{R}^n$, $\lim_{m \rightarrow \infty} (A^m \vec{x}) = \vec{x}_{\text{equ}}$ (the system will approach globally stable equilibrium distribution \vec{x}_{equ} in the long run)

$$\lim_{m \rightarrow \infty} A^m = \begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{bmatrix}$$

Matrix Inverses

Background: Given a function $f: X \rightarrow Y$, f is invertible if, for all $y \in Y$, \exists a unique solution $x \in X$ such that $f(x) = y$, in which case we can define inverse of f $f^{-1}: Y \rightarrow X$, $f^{-1}(y) = x$.

(*) Linear transformations, being functions, can also be invertible.

Invertible Matrices

A square matrix A is invertible if the associated linear transformation $\vec{y} = T(\vec{x}) = A\vec{x}$ is invertible, in which case we define the inverse transformation $\vec{x} = T^{-1}(\vec{y}) = A^{-1}\vec{y}$ (A^{-1} being the matrix of T^{-1}).

A square matrix $A \in \mathbb{R}^{n \times n}$ is invertible iff $\text{ref}(A) = I_n$, i.e. iff $\text{rank}(A) = n$.

(*) Given matrices $A, A^{-1} \in \mathbb{R}^{n \times n}$, $A^{-1}A = AA^{-1} = I_n$

(*) Invertibility and Linear Systems

Given matrix $A \in \mathbb{R}^{n \times n}$, linear system represented by $A\vec{x} = \vec{b}$:

• If A is invertible, the system has a unique solution $\vec{x} = A^{-1}\vec{b}$

• If A is not invertible, the system has either infinite solutions or no solution.

Invertibility (cont.)

2/6/23
Lecture 12

Finding Matrix Inverses

Given matrix $A \in \mathbb{R}^{n \times n}$, A^{-1} can be found by forming the $n \times 2n$ matrix $[A \mid I_n]$ and computing $\text{ref}[A \mid I_n]$.

1) If $\text{ref}[A \mid I_n] = [I_n \mid B]$ for some matrix B , A is invertible and $A^{-1} = B$.

2) If $\text{ref}[A \mid I_n] = [B \mid C]$ for some matrices B and C ($B \neq I_n$), A is not invertible.

Properties of Inverse Matrices

• Given invertible matrices $A, B \in \mathbb{R}^{n \times n}$, BA is also invertible $((BA)^{-1} = A^{-1}B^{-1})$

• Given two matrices $A, B \in \mathbb{R}^{n \times n}$ s.t. $BA = I_n$, A and B are both invertible ($A^{-1} = B$; $B^{-1} = A$)

Inverse of a 2×2 Matrix

Given 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, A is invertible iff $ad - bc \neq 0$.

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

(*) Determinant of a 2×2 Matrix

Determinant of a 2×2 Matrix

The quantity $ad - bc$ is called the determinant of 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, i.e.:

$$\rightarrow \det(A) = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad (\text{Determinant of } A = \begin{bmatrix} a & b \\ c & d \end{bmatrix})$$

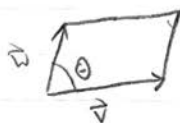
Notes

If we write $A = [\vec{v} \ \vec{w}]$ (2×2 matrix with columns \vec{v}, \vec{w}):

$$\rightarrow \det(A) = \|\vec{v}\| \|\vec{w}\| \sin(\theta) \quad (\text{where } \theta \text{ is the angle from } \vec{v} \text{ to } \vec{w} \text{ } (-\pi < \theta \leq \pi))$$

(*) $\det(A) = 0$ if $\vec{v} \parallel \vec{w}$ (\vec{v} and \vec{w} parallel)

Geometry: $|\det(A)|$ is the area of the parallelogram spanned by \vec{v} and \vec{w}



(Parallelogram spanned by \vec{v} and \vec{w})

Image, Span, and Kernel

2/8/23
Lecture 13

Definition: Image

Given a function $f: X \rightarrow Y$, the image of f is defined as all the values taken by the function in its target space:

$$\begin{aligned} \rightarrow \text{image}(f) &= \{f(x) \mid x \in X\} \\ &= \{y \mid y \in Y, \exists x \in X \text{ s.t. } f(x) = y\} \end{aligned}$$

$$(*) \text{im}(f) \subseteq \text{codomain}(f)$$

(*) The image of a linear transformation $T(\vec{x}) = A\vec{x}$ (denoted $\text{im}(T)$ or $\text{im}(A)$) is the span of the column vectors of A .

Definition: Span

Given a set of vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbb{R}^n$, the span of vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ is defined as the set of all linear combinations $c_1\vec{v}_1 + c_2\vec{v}_2 + \dots + c_m\vec{v}_m$ of the vectors $\vec{v}_1, \dots, \vec{v}_m$:

$$\text{span}(\vec{v}_1, \dots, \vec{v}_m) = \{c_1\vec{v}_1 + \dots + c_m\vec{v}_m \mid c_1, \dots, c_m \in \mathbb{R}\}$$

Images of Linear Transformations

Given linear transformation $T(\vec{x}) = A\vec{x}$, $A = [\vec{v}_1 \dots \vec{v}_m]$ -

$\rightarrow \text{im}(A) = \text{span}(\vec{v}_1, \dots, \vec{v}_m)$ has the following properties:

- $(\vec{0} \in \mathbb{R}^n) \in \text{im}(A)$
- $\text{im}(A)$ is closed under addition, i.e.:
 - (*) $\vec{v}_1, \vec{v}_2 \in \text{im}(A) \rightarrow (\vec{v}_1 + \vec{v}_2) \in \text{im}(A)$
- $\text{im}(A)$ is closed under scalar multiplication, i.e.:
 - (*) $\vec{v}_1 \in \text{im}(A) \rightarrow k\vec{v}_1 \in \text{im}(A) \quad (k \in \mathbb{R})$

Definition: Kernel

Given a linear transformation $T(\vec{x}) = A\vec{x}$, the kernel of $T(\vec{x})$ is the set of all zeros of $T(\vec{x})$, i.e. of all vectors \vec{x} such that $T(\vec{x}) = A\vec{x} = \vec{0}$ (set of solutions).

$$\begin{aligned} \rightarrow \text{kernel of } T &= \text{ker}(A) = \text{ker}(T) \\ &= \{\vec{x} \mid T(\vec{x}) = \vec{0}\} \end{aligned}$$

$$(*) \text{ker}(T) \subseteq \text{domain}(T)$$

(*) The kernel of $T(\vec{x}) = A\vec{x}$ can be seen as the set of solutions of the linear system $A\vec{x} = \vec{0}$

(*) The notion of the zeros of a function is not limited to linear algebra

Subspaces

2/10/23
Lecture 14

Properties of Kernel

Given a linear transformation T :

- 1) $\vec{0} \in \ker(T)$
- 2) The kernel is closed under addition
- 3) The kernel is closed under scalar multiplication.

Case: $\ker(A) = \{\vec{0}\}$

a) iff $\text{rank}(A) = m$ ($A \in \mathbb{R}^{n \times m}$)

b) $\ker(A) = \vec{0} \rightarrow m \leq n$

(*) $n < m \rightarrow \ker(A) \neq \{\vec{0}\}$

c) A square $\rightarrow \ker(A) = \vec{0}$ iff A invertible

Properties of Invertible Matrices

Given matrix $A \in \mathbb{R}^{n \times n}$, the following statements are equivalent:

- i. A is invertible
- ii. The system $A\vec{x} = \vec{b}$ has a unique solution $\vec{x} \forall \vec{b} \in \mathbb{R}^n$
- iii. $\text{ref}(A) = I_n$
- iv. $\text{rank}(A) = n$
- v. $\text{im}(A) = \mathbb{R}^n$
- vi. $\ker(A) = \{\vec{0}\}$

(*) equivalent - all true or all false

Subspaces

A subset W of the vector space \mathbb{R}^n is called a [linear] subspace of \mathbb{R}^n if it satisfies:

- 1) $\vec{0} \in W$
- 2) W is closed under addition (i.e. $\vec{w}_1, \vec{w}_2 \in W \rightarrow \vec{w}_1 + \vec{w}_2 \in W$)
- 3) W is closed under scalar multiplication (i.e. $\vec{w} \in W \rightarrow k\vec{w}, (k \in \mathbb{R}) \in W$)

Theorem: Given linear transformation $T(\vec{x}) = A\vec{x} : \mathbb{R}^m \rightarrow \mathbb{R}^n$:

- $\ker(A) = \ker(T) \subseteq \mathbb{R}^m$ (kernel is a subspace of the domain)
- $\text{im}(A) = \text{im}(T) \subseteq \mathbb{R}^n$ (image is a subspace of the codomain/target space)

Subspace Dimensions:

	Subspaces of \mathbb{R}^2	Subspaces of \mathbb{R}^3
3D	N/A	\mathbb{R}^3
2D	\mathbb{R}^2	Planes through $\vec{0}$
1D	Lines through $\vec{0}$	Lines through $\vec{0}$
0D	$\{\vec{0}\}$	$\{\vec{0}\}$

Bases and Linear Independence

2/13/23
Lecture 15

Bases and Linear Independence

Definition: A vector \vec{v}_i in the list $\vec{v}_1, \dots, \vec{v}_n$ is considered redundant if \vec{v}_i can be expressed as a linear combination of the vectors $\{\vec{v}_j \mid 1 \leq j \leq n, j \neq i\}$ [other vectors in the list]
(* "Redundant" is not an established linear algebra term)

Linear Independence

A set of vectors $\vec{v}_1, \dots, \vec{v}_n$ are called linearly independent if no vector \vec{v}_i in the set can be expressed as a linear combination of the other vectors in the set.

i.e. $\nexists \vec{v}_i$ s.t. $\vec{v}_i = c_1 \vec{v}_1 + c_2 \vec{v}_2 + \dots + c_k \vec{v}_k$ (for any constants c_1, \dots, c_k)

A set of vectors that does not fulfill the above (contains a vector that is expressible as a linear combination of the other vectors in the set) is called linearly dependent.

Bases

A set S of vectors $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n$ in a subspace V of \mathbb{R}^n is a basis of V if they span V (i.e. $V \subseteq \text{span}(S)$) and are linearly independent.

i.e. $\forall \vec{v} \in V \rightarrow \vec{v} = c_1 \vec{v}_1 + c_2 \vec{v}_2 + \dots + c_k \vec{v}_k$ (for some constants c_1, \dots, c_k)

Basis of an Image

Given a matrix A with column vectors $\vec{v}_1, \dots, \vec{v}_n$, a basis of the image $\text{im}(A)$ of A can be formed from the set of column vectors of A , omitting any redundant vectors.

Linear Relations

Given a set of vectors $\vec{v}_1, \dots, \vec{v}_n$, a [linear] relation among the vectors $\vec{v}_1, \dots, \vec{v}_n$ is an equation of the form:

$$c_1 \vec{v}_1 + c_2 \vec{v}_2 + \dots + c_n \vec{v}_n = \vec{0} \quad (c_1, \dots, c_n \in \mathbb{R})$$

Types of Linear Relations

- The trivial relation is the relation $c_1 = c_2 = \dots = c_n = 0$ (always exists)
- A nontrivial relation (if one exists) is a relation with at least one $c_i \neq 0$
(* \exists nontrivial $\rightarrow \vec{v}_1, \dots, \vec{v}_n$ linearly dependent)

Bases and Dimension

2/17/23
Lecture 16

Kernel and Relations

The vectors $\in \ker(A)$ correspond to linear relations among column vectors v_1, \dots, v_m of A , i.e.:

$$\vec{x} \in \ker(A) (A\vec{x} = \vec{0}) \rightarrow x_1 \vec{v}_1 + x_2 \vec{v}_2 + \dots + x_m \vec{v}_m = \vec{0}$$

Properties of Linear Independence

Given vectors $v_1, \dots, v_m \in \mathbb{R}^n$, the following statements are equivalent:

- i. Vectors v_1, \dots, v_m are linearly independent
- ii. No vector v_i is a linear combination of the other vectors in the list $\{v_j\}_{j \neq i}$
- iii. The only linear relation among the vectors v_1, \dots, v_m is the trivial relation

iv. $\ker [\vec{v}_1 \dots \vec{v}_m] = \{\vec{0}\}$

v. $\text{rank} [\vec{v}_1 \dots \vec{v}_m] = m$

(*) In a vector space \mathbb{R}^n , we can find at most n linearly independent vectors

Bases

Given a subspace V of \mathbb{R}^n , a set of vectors $\vec{v}_1, \dots, \vec{v}_m$ form a basis of V iff every vector $\vec{v} \in V$ can be expressed as a linear combination:

$$\vec{v} = c_1 \vec{v}_1 + c_2 \vec{v}_2 + \dots + c_m \vec{v}_m \quad (c_1, \dots, c_m \text{ being the coordinates of } \vec{v} \text{ with respect to basis } v_1, \dots, v_m)$$

Bases and Dimension

- Given a subspace $V \subseteq \mathbb{R}^n$, all bases of V will contain the same number of vectors.
- This number is called the dimension of V ($\dim(V)$)
- Given a subspace $V \subseteq \mathbb{R}^n$, $\dim(V) = m$:
 - We can find at most m linearly independent vectors in V
 - We need at least m vectors to span V
 - If m vectors in V are linearly independent, they form a basis of V ($\text{span } V$)

Rank and Dimension: $\dim(\text{im}(A)) = \text{rank}(A)$

Rank-Nullity, B-Coordinates, & Similarity

2/12/23
Lecture 17

Rank-Nullity Theorem

Given a matrix A , the nullity of A is the dimension of the kernel of A ($\dim(\ker(A))$).
For any matrix $A \in \mathbb{R}^{n \times m}$:

$$\text{nullity}(A) + \text{rank}(A) = \dim(\ker(A)) + \dim(\text{im}(A)) = m \quad \left(\begin{array}{l} \text{also written} \\ N(A) + R(A) = m \end{array} \right)$$

Properties of Invertible Matrices (cont.)

- vii. The column vectors of A are linearly independent.
- viii. The column vectors of A span \mathbb{R}^n .
- ix. The column vectors of A form a basis of \mathbb{R}^n .

(*) Bases of \mathbb{R}^n

A set of vectors $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n$ form a basis of \mathbb{R}^n iff the matrix $A = [\vec{v}_1 \dots \vec{v}_n]$ is invertible.

B-Coordinates

Given a basis $B = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m)$ of a subspace $V \subseteq \mathbb{R}^n$, any vector $\vec{x} \in V$ can be written uniquely as $\vec{x} = c_1 \vec{v}_1 + \dots + c_m \vec{v}_m$. The scalars c_1, \dots, c_m are called the B-coordinates of \vec{x} ; we then define the B-coordinate vector of \vec{x} , $[\vec{x}]_B = \langle c_1, \dots, c_m \rangle$. Thus, we write:

$$\vec{x} = S[\vec{x}]_B = \begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_m \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} \quad (S \in \mathbb{R}^{n \times m})$$

Linearity of Coordinates

Given basis B of subspace $V \subseteq \mathbb{R}^n$:

- i. $[\vec{x} + \vec{y}]_B = [\vec{x}]_B + [\vec{y}]_B \quad \forall \vec{x}, \vec{y} \in V$
- ii. $[k\vec{x}]_B = k[\vec{x}]_B \quad \forall \vec{x} \in V, k \in \mathbb{R}$

Matrices of Linear Transformations

Given $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$, basis $B = (\vec{v}_1, \dots, \vec{v}_n)$ of \mathbb{R}^n , \exists a unique B-matrix of T transforming $[\vec{x}]_B$ into $[T(\vec{x})]_B$:

$$T(\vec{x}) = B[\vec{x}]_B \rightarrow B = [T(\vec{v}_1)]_B \dots [T(\vec{v}_n)]_B$$

Similar Matrices

Given two $n \times n$ matrices A and B , we say A is similar to B if \exists invertible matrix S such that:

$$AS = SB \rightarrow B = S^{-1}AS$$

(*) Notes on Similarity

- Similarity is an equivalence relation
- Given $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$, basis $B = (\vec{v}_1, \dots, \vec{v}_n)$ of \mathbb{R}^n , standard matrix A and B-matrix B of T :

$$A = SBS^{-1} \quad (S = [\vec{v}_1 \dots \vec{v}_n])$$

Orthogonality & Projection

2/22/23
Lecture 17
(cont.)

Orthogonality & Vectors Review

- Vectors $\vec{v}, \vec{w} \in \mathbb{R}^n$ are perpendicular or orthogonal if $\vec{v} \cdot \vec{w} = 0$
- The length (or magnitude) of a vector \vec{v} is $\|\vec{v}\| = \sqrt{\vec{v} \cdot \vec{v}}$
- A vector is a unit vector if its length is 1

Orthonormal Vectors

A set of vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbb{R}^n$ are called orthonormal if they are all unit vectors and orthogonal to each other.

$$\vec{v}_i \cdot \vec{v}_j = \begin{cases} 1 & (i=j) \\ 0 & (i \neq j) \end{cases}$$

2/24/23
Lecture 18

Properties of Orthonormal Vectors

- Orthonormal vectors are linearly independent
- Orthonormal vectors $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n$ form an [orthonormal] basis of \mathbb{R}^n

(*) Cauchy-Schwarz Inequality

$$|\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|$$

A vector $\vec{x} \in \mathbb{R}^n$ is called orthogonal to a subspace $V \subseteq \mathbb{R}^n$ if \vec{x} is orthogonal to all vectors $\in V$.

→ Any vector $\vec{w} \in \mathbb{R}^n$ can be uniquely represented $\vec{w} = \vec{w}'' + \vec{w}^\perp$, where $\vec{w}'' \in V$ and $\vec{w}^\perp \perp V$.

Orthogonal Projection

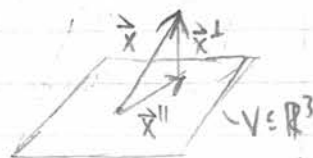
Given vector $\vec{x} \in \mathbb{R}^n$, subspace $V \subseteq \mathbb{R}^n$, the orthogonal projection of \vec{x} onto V is:

$$\text{proj}_V \vec{x} = \vec{x}'' \quad \begin{matrix} (*) \text{ If } V = \mathbb{R}^n \\ \rightarrow \text{proj}_V \vec{x} = \vec{x} \end{matrix}$$

(*) Orthogonal Projection

Orthogonal projection onto subspaces can be represented as a linear transformation, i.e. $T(\vec{x}) = \text{proj}_V \vec{x}$.

$$(*) \|\text{proj}_V \vec{x}\| \leq \|\vec{x}\|$$



Orthogonal Projection

Given vector $\vec{x} \in \mathbb{R}^n$, $V \subseteq \mathbb{R}^n$ with orthonormal basis $\vec{u}_1, \dots, \vec{u}_m$, then:

$$\text{proj}_V \vec{x} = \vec{x}'' = (\vec{u}_1 \cdot \vec{x}) \vec{u}_1 + \dots + (\vec{u}_m \cdot \vec{x}) \vec{u}_m \quad [= \vec{x} \text{ if } V = \mathbb{R}^n]$$

(*) Angle Between Vectors

$$\vec{x}, \vec{y} \rightarrow \theta = \cos^{-1} \left(\frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|} \right) \quad \begin{matrix} \Delta \theta \\ \vec{x} \end{matrix} \quad (0 \leq \theta \leq \pi)$$

Orthogonal Complement

Given subspace $V \subseteq \mathbb{R}^n$, the orthogonal complement V^\perp of V is the set of vectors $\vec{x} \in \mathbb{R}^n$ orthogonal to all vectors in V :

$$V^\perp = \{ \vec{x} \mid \vec{x} \in \mathbb{R}^n, \vec{x} \cdot \vec{v} = 0 \forall \vec{v} \in V \}$$

$$(*) V^\perp = \ker(\text{proj}_V \vec{x})$$

Properties of Orthogonal Complement

- $V^\perp \subseteq \mathbb{R}^n$
- $V \cap V^\perp = \{\vec{0}\}$
- $\dim(V) + \dim(V^\perp) = n$
- $(V^\perp)^\perp = V$

Gram-Schmidt & Orthogonal Transformations

2/17/23
Lecture 19

Gram-Schmidt (§5.2)

Given a basis $\vec{v}_1, \dots, \vec{v}_m$ of subspace $V \subseteq \mathbb{R}^n$, we can construct an orthonormal basis $\vec{u}_1, \dots, \vec{u}_m$ of V by expressing each vector $\vec{v}_j, j \geq 2$ as the sum of components parallel and perpendicular to the span of the preceding vectors $\vec{v}_1, \dots, \vec{v}_{j-1}$:

$$\vec{v}_j = \vec{v}_j'' + \vec{v}_j^\perp \quad (\text{relative to } \text{span}(\vec{v}_1, \dots, \vec{v}_{j-1}))$$

$$\rightarrow \vec{u}_1 = \frac{1}{\|\vec{v}_1\|} \vec{v}_1, \vec{u}_2 = \frac{1}{\|\vec{v}_2^\perp\|} \vec{v}_2^\perp, \dots, \vec{u}_m = \frac{1}{\|\vec{v}_m^\perp\|} \vec{v}_m^\perp \quad (\vec{u}_1, \dots, \vec{u}_m \text{ orthonormal basis of } V)$$

THM: Gram-Schmidt

$$\vec{v}_j^\perp = \vec{v}_j - (\vec{u}_1 \cdot \vec{v}_j) \vec{u}_1 - \dots - (\vec{u}_{j-1} \cdot \vec{v}_j) \vec{u}_{j-1} \quad [= \vec{v}_j - \vec{v}_j''] \quad (j=2, \dots, m)$$

(*) The Gram-Schmidt process represents a change of basis from $B = (\vec{v}_1, \dots, \vec{v}_m)$ to orthonormal $U = (\vec{u}_1, \dots, \vec{u}_m)$

$$\rightarrow \underbrace{\begin{bmatrix} \vec{v}_1 & \dots & \vec{v}_m \end{bmatrix}}_M = \underbrace{\begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_m \end{bmatrix}}_Q R \quad (\text{QR factorization of } M)$$

QR Factorization

Given an $n \times m$ matrix M with linearly independent columns $\vec{v}_1, \dots, \vec{v}_m$, \exists a unique representation $M = QR$, where Q is an $n \times m$ matrix with orthonormal columns $\vec{u}_1, \dots, \vec{u}_m$ and R is an upper triangular matrix with positive diagonal entries.

$$\therefore r_{11} = \|\vec{v}_1\|, r_{jj} = \|\vec{v}_j^\perp\|, r_{ij} = \vec{u}_i \cdot \vec{v}_j, (\text{for } i < j) \quad [\text{Gram-Schmidt coefficients}]$$

Orthogonal Transformations & Matrices (§5.3)

A linear transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is called orthogonal if it preserves the length of vectors, i.e. $\|T(\vec{x})\| = \|\vec{x}\|, \forall \vec{x} \in \mathbb{R}^n$

If $T(\vec{x}) = A\vec{x}$ is an orthogonal transformation, A is an orthogonal matrix.

Orthogonal Matrices

2/27/23
Lecture 19
(cont.)

Orthogonal Transformations

- Orthogonal transformations preserve orthogonality, i.e. $\vec{v} \perp \vec{w} \rightarrow T(\vec{v}) \perp T(\vec{w})$
- A linear transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is orthogonal iff the vectors $T(\vec{e}_1), T(\vec{e}_2), \dots, T(\vec{e}_n)$ form an orthonormal basis of \mathbb{R}^n (i.e. column vectors of A , $T(\vec{x}) = A\vec{x}$)
- An $n \times n$ matrix A is orthogonal iff its columns form an orthonormal basis of \mathbb{R}^n ★

- (*) Properties:
- 1) The product of two orthogonal matrices is orthogonal
 - 2) The inverse of an orthogonal matrix is orthogonal

Definition: Transpose

Given an $m \times n$ matrix A , the transpose A^T of A is the $n \times m$ matrix such that $A^T_{ij} = A_{ji}$ (and vice versa), i.e. the "reflection" of A over its diagonal.

- (*) Definitions:
- 1) A square matrix A is symmetric if $A^T = A$
 - 2) A square matrix A is skew-symmetric if $A^T = -A$

(1) A square matrix A is orthogonal iff $A^T A = I$ (i.e. $A^T = A^{-1}$)

Properties of Transpose

- 1) $(A+B)^T = A^T + B^T$
- 2) $(kA)^T = k(A^T)$
- 3) $(AB)^T = B^T A^T$
- 4) $\text{rank}(A^T) = \text{rank}(A)$
- 5) $(A^T)^T = (A^{-1})^T$

Properties of Orthogonal Matrices

Given $A \in \mathbb{R}^{n \times n}$, the following statements are equivalent:

- i. A is an orthogonal matrix.
- ii. $T(\vec{x}) = A\vec{x}$ preserves length, i.e. $\|A\vec{x}\| = \|\vec{x}\| \forall \vec{x}$.
- iii. The columns of A form an orthonormal basis of \mathbb{R}^n .
- iv. $A^T = A^{-1}$.
- vi. A preserves dot product, i.e. $(A\vec{x}) \cdot (A\vec{y}) = \vec{x} \cdot \vec{y} \forall \vec{x}, \vec{y}$.

Matrices of Orthogonal Projections

Given a subspace $V \subseteq \mathbb{R}^n$ with orthonormal basis $\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m$, the matrix P of the orthogonal projection of a vector $\vec{v} \in \mathbb{R}^n$ onto V :

$$P = QQ^T \quad (Q = [\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m]) \quad [\text{Orthogonal projection onto } V]$$

(*) The matrix P is symmetric ($P = P^T$)

Intro to Determinants

3/1/23

Lecture 20

Determinants (§6.1)

The determinant of a 3×3 matrix $A = [\vec{u} \ \vec{v} \ \vec{w}]$ is:

$$\det(A) = \det([\vec{u} \ \vec{v} \ \vec{w}]) = \vec{u} \cdot (\vec{v} \times \vec{w}) \quad [\text{Determinant of a } 3 \times 3]$$

THM A square matrix A is invertible iff $\det(A) \neq 0$.

$$(*) \det(A) = 0 \Leftrightarrow A \text{ not invertible}$$

(*) Linearity of Determinant

Given $L(\vec{x}) = \det(\vec{v}_1, \dots, \vec{v}_{i-1}, \vec{x}, \vec{v}_{i+1}, \dots, \vec{v}_n)$, $L(\vec{x})$ is a linear transformation [linearity of determinant in the i th row]. The determinant is also linear in all columns.

$$(*) \text{Reminder: } L(\vec{x}) \text{ linear} \rightarrow L(\vec{x} + \vec{y}) = L(\vec{x}) + L(\vec{y}); L(k\vec{x}) = kL(\vec{x})$$

Determinant of a Square Matrix

Given $n \times n$ matrix A , we define a pattern P as a selection of n items $\in A$, such that every row & column of A contains exactly one element $\in P$.

Definitions: • prod P denotes the product of all entries in a given pattern P .

• An inversion in a given pattern P is a pair of elements $\in P$, such that one is located above & to the right of the other in A .

• The sign/signature/signum of pattern P is defined as $\text{sgn } P = (-1)^{\# \text{ of inversions in } P}$.

THM
Determinant

$$\det(A) = \sum (\text{sgn } P)(\text{prod } P) \quad \forall \text{ patterns } P \text{ in } A.$$

(*) Determinants for Special Matrices

• The determinant of an upper triangular, lower triangular, or diagonal matrix A is the product of its diagonal entries.

• Given square matrices A, C [of potentially different size]:

$$\det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ B & C \end{pmatrix} = (\det A)(\det C) \quad [\text{Determinant of Block Matrices}]$$

$$(*) \det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = (\det A)(\det D) - (\det B)(\det C) \text{ does not always hold}$$

Determinant (cont.)

3/1/23
Lecture 20
(cont.)

Properties of Determinant (§6.2)

i. $\det(A^T) = \det(A)$ (square) [Transpose]

ii. $\det(AB) = \det(A)\det(B)$ [Multiplication]

iii. $\det(A^n) = \det(A)^n$ [Exponentiation]

(*) $\det(A^{-1}) = \det(A)^{-1}$ [Inverse]

(*) Row Operations & Determinant

i. $B = A$, multiplying a row by scalar k :

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

ii. $B = A$, swapping two rows of A :

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

iii. $B = A$, adding a [multiple of] row to another row:

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

3/3/23
Lecture 21

(*) Determinant can be calculated by reducing to ref via Gauss-Jordan, then multiplying the determinant of the ref matrix by scaling factors (as appropriate)

(*) Laplace Expansion

Definition: Given $n \times m$ matrix A , we denote the $(n-1) \times (m-1)$ matrix obtained from removing the i th row and j th column of A , A_{ij} . The determinant of A_{ij} is called a minor of A .

$$\begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1m} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{im} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nm} \end{bmatrix} = A \rightarrow A_{ij} = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1m} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{im} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nm} \end{bmatrix} \quad (A \in \mathbb{R}^{n \times m}) \quad [\text{Minors of } A]$$

→ Determinant (Laplace)

Given $n \times n$ matrix A :

$$\rightarrow \det(A) = \sum_{i=1}^n (-1)^{ij} (A_{ij}), \text{ for some } j \quad [\text{Expansion by Column}]$$

$$\rightarrow \det(A) = \sum_{j=1}^n (-1)^{ij} (A_{ij}), \text{ for some } i \quad [\text{Expansion by Row}]$$

Misc Properties

• If two matrices A and B similar, $\det(A) = \det(B)$.

• Given linear transformation w/ B -matrix B (for any basis B , incl. standard): $\det(T) = \det(B)$

(*) $T(\vec{x}) = A\vec{x} \rightarrow \det(T) = \det(A)$

Determinants and Geometry

3/3/23
Lecture 2
(cont.)

Determinants of Orthogonal Matrices

- A determinant of an orthogonal matrix is either 1 or -1 ($|\det A| = 1$)
- (*) $\det A = 1 \rightarrow A$ is a rotation matrix

Determinant and Vectors

Given matrix A w/ columns $\vec{v}_1, \dots, \vec{v}_n$:

$$|\det A| = \|\vec{v}_1\| \|\vec{v}_2^\perp\| \dots \|\vec{v}_n^\perp\| \quad \text{T.H.M.}$$

Determinants and Geometry (§6.3)

Given a set of vectors $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n$, we define the n -parallelpiped defined by $\vec{v}_1, \dots, \vec{v}_n$ as the set of all points $(x_1, \dots, x_n) \in \mathbb{R}^n = c_1 \vec{v}_1 + \dots + c_n \vec{v}_n$ ($0 \leq c_i \leq 1$). The n -volume of the n -parallelpiped is defined as:

$$V(\vec{v}_1, \dots, \vec{v}_n) = V(\vec{v}_1, \dots, \vec{v}_{n-1}) \|\vec{v}_n^\perp\|; \quad V(\vec{v}_1) = \|\vec{v}_1\| \quad [\text{M-Volume}]$$

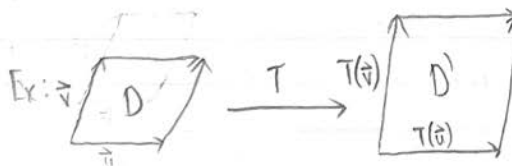
$$(*) A = [\vec{v}_1, \dots, \vec{v}_n] \rightarrow V(\vec{v}_1, \dots, \vec{v}_n) = \sqrt{\det(A^T A)} \quad \text{Ex: } V(\vec{v}_1, \vec{v}_2, \vec{v}_3) = \det([\vec{v}_1, \vec{v}_2, \vec{v}_3])$$

$$(*) n=n \rightarrow V(\vec{v}_1, \dots, \vec{v}_n) = |\det(A)|$$

(*) Determinants and Transformations

Given a linear transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $T(\vec{x}) = A\vec{x}$, $|\det A|$ is the expansion factor of T on n -parallelpiped:

$$V(A\vec{v}_1, \dots, A\vec{v}_n) = |\det A| V(\vec{v}_1, \dots, \vec{v}_n) \quad (\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n)$$

Ex:  $\rightarrow \text{area}(D') = |\det A| \cdot \text{area}(D)$

(*) Cramer's Rule

Given a linear system $A\vec{x} = \vec{b}$ ($A \in \mathbb{R}^{n \times n}$, invertible), the components x_i of the solution vector \vec{x} are:

$$x_i = \frac{\det(A_{\vec{b}, i})}{\det(A)}, \quad \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},$$

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

(*) Adjoint

Given invertible matrix $A \in \mathbb{R}^{n \times n}$, the classical adjoint $\text{adj}(A)$ of A is the $n \times n$ matrix, $\{\text{adj}(A)_{ij} = (-1)^{i+j} \det(A_{ji})\}$.

$$\rightarrow (*) A^{-1} = \frac{1}{\det(A)} \text{adj}(A)$$

Eigenvectors & Eigenvalues

3/6/23

Lecture 22

Diagonalization

Def.: A linear transformation $T(\vec{x}) = A\vec{x}$ is said to be diagonalizable if the matrix B of T with respect to some basis is diagonal, i.e. if A is similar to some diagonal matrix B [\exists invertible matrix S s.t. $S^{-1}AS = B$ for some diagonal matrix B].

Eigenvectors & Eigenvalues (§6.1)

Given a linear transformation $T(\vec{x}) = A\vec{x}$ ($T: \mathbb{R}^n \rightarrow \mathbb{R}^n$), a nonzero vector $\vec{v} \in \mathbb{R}^n$ is called an eigenvector of A (or T) if $A\vec{v} = \lambda\vec{v}$ for some scalar λ [eigenvalue associated w/ \vec{v}].

Notes on Eigenvectors

- A basis v_1, \dots, v_n of \mathbb{R}^n is called an eigenbasis of A if the vectors v_1, \dots, v_n are eigenvectors of A .
- A matrix A is diagonalizable iff \exists an eigenbasis of A , in which case:

$$v_1, \dots, v_n \text{ (eigenbasis of } A) \rightarrow S = \begin{bmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{bmatrix}, B = \begin{bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_n \end{bmatrix} \text{ s.t. } S^{-1}AS = B$$

- (*) iff matrices S and B diagonalize A , the columns of S form an eigenbasis of A .
- The only possible eigenvalues for an orthogonal matrix are 1 and -1 .
- A matrix A is invertible iff 0 is not an eigenvalue of A .

3/8/23

Lecture 23

THM

Characteristic Eq.

Given an $n \times n$ matrix A and a scalar λ , λ is an eigenvalue of A iff $\det(A - \lambda I_n) = 0$. (Characteristic/Secular Equation)

↳ (*) The eigenvalues of a triangular matrix are its diagonal entries

(*) Trace

Definition: Given a square matrix A , the trace $\text{tr}(A)$ of A is the sum of its diagonal entries.

Notes on Eigenvalues

3/8/23
Lecture 23
(cont.)

Reminder: λ eigenvalue of $A \rightarrow \det(A - \lambda I_n) = 0$ [Characteristic Equation]

Definition: Characteristic Polynomial (§7.2)

Given $n \times n$ matrix A , $\det(A - \lambda I_n)$ is a polynomial of degree n ; this polynomial is called the characteristic polynomial $f_A(\lambda)$ of A .

$$\det(A - \lambda I_n) = (-1)^n \lambda^n + (-1)^{n-1} (\text{tr} A) \lambda^{n-1} + \dots + \det A \quad [\text{Characteristic Polynomial}]$$

$$\hookrightarrow (*) A \in \mathbb{R}^{2 \times 2} \rightarrow f_A(\lambda) = \lambda^2 - (\text{tr} A)\lambda + \det A$$

\hookrightarrow The zeros of the characteristic are the eigenvalues of A

Notes on Eigenvalues

Definition: The algebraic multiplicity of an eigenvalue λ_0 of matrix A is defined as the degree of that root in the characteristic polynomial, i.e. $f_A(\lambda) = (\lambda_0 - \lambda)^k g(\lambda)$ [for some polynomial $g(\lambda) \rightarrow \text{altm}(\lambda_0) = k$]
[λ_0 is a root of multiplicity k of $f_A(\lambda)$]
 $g(\lambda_0) \neq 0$

- An $n \times n$ has at most n real eigenvalues
- (*) n odd \rightarrow has at least 1 real eigenvalue

(*) Properties of Eigenvalues

Given an $n \times n$ matrix A with n eigenvalues $\lambda_1, \dots, \lambda_n$ ($f_A(\lambda) = (\lambda_1 - \lambda) \dots (\lambda_n - \lambda)$):

$$\det(A) = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n; \quad \text{tr}(A) = \lambda_1 + \lambda_2 + \dots + \lambda_n \quad (**) \lambda_1, \dots, \lambda_n \text{ need not be distinct}$$

Definition: Eigenspaces (§7.3)

Given an eigenvalue λ of an $n \times n$ matrix A , the eigenspace E_λ associated with λ is defined as the kernel of the matrix $(A - \lambda I_n)$, i.e.:

$$E_\lambda = \ker(A - \lambda I_n) = \{ \vec{v} \mid A\vec{v} = \lambda \vec{v} \} \quad [\text{Eigenspace ass. w/ } \lambda]$$

\hookrightarrow Definition: The geometric multiplicity $\text{geomu}(\lambda)$ of λ is defined as the dimension of the eigenspace E_λ , i.e. $\text{geomu}(\lambda) = \text{nullity}(A - \lambda I_n) = n - \text{rank}(A - \lambda I_n)$ (**) $\text{geomu}(\lambda) \leq \text{altm}(\lambda)$

Diagonalization

3/12/23
Lecture 24

Notes on Eigenspaces

- An eigenspace E_λ is the set of all eigenvectors w/ associated eigenvalue λ
- The set of basis vectors $\vec{v}_1, \dots, \vec{v}_s$ across all eigenspaces of a matrix $A \in \mathbb{R}^{n \times n}$ is linearly independent
- $s = \sum \dim E_\lambda$ for all eigenvalues λ [sum of dimensions of eigenspaces] ($s \leq n$)
- (*) An $n \times n$ matrix A is diagonalizable iff $s = \sum \dim E_\lambda = n$ [$A \in \mathbb{R}^{n \times n}$]
- (*) If $A \in \mathbb{R}^{n \times n}$ has n eigenvalues, A is diagonalizable

(*) Eigenvalues and Similarity

- Given matrix A similar to B :
1. A and B have the same characteristic polynomial ($f_A(\lambda) = f_B(\lambda)$)
 2. $\text{rank } A = \text{rank } B$; $\text{nullity } A = \text{nullity } B$
 3. A and B have the same eigenvalues, with the same algebraic & geometric multiplicities (Eigenvectors may not be shared)
 4. $\det(A) = \det(B)$; $\text{tr}(A) = \text{tr}(B)$

Procedure for Diagonalization (*)

In order to diagonalize $n \times n$ matrix A (i.e. find invertible matrix S s.t. $S^{-1}AS = B$ [B diagonal]):

- 1) Find eigenvalues of A by solving the characteristic $f_A(\lambda) = \det(A - \lambda I_n) = 0$
- 2) For each eigenvalue λ , find a basis of the eigenspace $E_\lambda = \ker(A - \lambda I_n)$
- 3) A is diagonalizable iff the dimensions of the eigenspaces add up to n
- 4) Combining the basis vectors $\vec{v}_1, \dots, \vec{v}_n$ of the eigenspaces (from 2))
→ $S = [\vec{v}_1 \dots \vec{v}_n]$, $B = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ (λ_i being the eigenvalue associated with \vec{v}_i)

(*) Complex Numbers Review

- Expressed as $z = a + ib$ (a real; ib imaginary), $z = r(\cos \theta + i \sin \theta)$
→ $z(\alpha) \cdot z(\beta) = z(\alpha + \beta)$
- De Moivre's Formula: $(\cos \theta + i \sin \theta)^n = \cos(n\theta) + i \sin(n\theta)$

Orthogonal Diagonalization

3/10/23
Lecture 24
(cont.)

(*) Complex Eigenvalues (§7.5)

- Any polynomial $p(\lambda)$ with complex coefficients splits, i.e. $p(\lambda) = k(\lambda - \lambda_1) \dots (\lambda - \lambda_n)$ [k, λ_i complex]
- Has exactly n complex roots ((*) Counted with algebraic multiplicities - may not be distinct)
- A complex $n \times n$ matrix A has n complex eigenvalues (from the characteristic)
- Given real 2×2 matrix A with eigenvalues $a \pm ib$ ($b \neq 0$), eigenvector $\vec{v} + i\vec{w}$ associated with eigenvalue $a + ib$

$$\rightarrow S^T A S = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}; S = \begin{bmatrix} \vec{w} & \vec{v} \end{bmatrix}$$

(*) The relationship between eigenvalues and determinant/trace hold for complex eigenvalues

Orthogonal Diagonalization (§8.1)

Definition: A matrix A is orthogonally diagonalizable if there exists an orthogonal matrix S that diagonalizes A (i.e. $S^T A S$ [$= S^{-1} A S$] is diagonal)

$$S^T A S = B \text{ [} B \text{ diagonal]}; S \text{ orthogonal [Orthogonal Diagonalization]}$$

Spectral
Theorem

A matrix A is orthogonally diagonalizable iff A is symmetric (i.e. $A^T = A$).

- Given a symmetric matrix A : Let \vec{v}_1 and \vec{v}_2 be eigenvectors of A associated with eigenvalues λ_1 and λ_2 . If $\lambda_1 \neq \lambda_2$ (eigenvalues distinct) $\rightarrow \vec{v}_1 \cdot \vec{v}_2 = 0$ [\vec{v}_1, \vec{v}_2 orthogonal]
- A symmetric $n \times n$ matrix A has n real eigenvalues (counted w/ algebraic multiplicities)

3/13/23
Lecture 25

Procedure for Orthogonal Diagonalization (★)

Given a symmetric $n \times n$ matrix A :

- 1) Find the eigenvalues of A , and find a basis of each eigenspace
- 2) Use Gram-Schmidt to find an orthonormal basis of each eigenspace.
- 3) Form an orthonormal eigenbasis for A $\vec{v}_1, \dots, \vec{v}_n$ by concatenating the bases from 2):

$$S = [\vec{v}_1 \ \vec{v}_2 \ \dots \ \vec{v}_n] \text{ (} S \text{ orthogonal; } S^T A S \text{ diagonal)}$$

(*) Notes: Linear Spaces & Isomorphisms

4/2/23
Notes

Definition: Linear Spaces (§4.1)

A linear space is a set with an associated rule for addition ($f, g \in \text{linear space } V \rightarrow (f+g) \in V$) and scalar multiplication ($f \in V, k \in \mathbb{R} \rightarrow kf \in V$), s.t. $\forall f, g, h \in V$ and $c, k \in \mathbb{R}$:

i) $(f+g)+h = f+(g+h)$

v) $k(f+g) = kf + kg$

ii) $f+g = g+f$

vi) $(c+k)f = cf + kf$

iii) \exists neutral element $n = 0 \in V$ ($n+f = f \forall f$)

vii) $c(kf) = (ck)f$

iv) for each $f \in V, \exists g = -f \in V$ s.t. $f+g = 0$

vi) $1f = f$

↳ Def: A vector space is a linear space containing vectors (with the associated operations of vector addition and scalar multiplication).

↳ (*) A linear transformation can be considered a function that takes vector spaces as domain and codomain (+ other conditions: $T(f+g) = T(f) + T(g), T(kf) = kT(f)$).

Definition: Isomorphism (§4.2)

An isomorphism is an invertible linear transformation. Given two linear spaces V and W , we say V is isomorphic to W (and vice versa) if \exists an isomorphism $T: V \rightarrow W$.

(*) Isomorphisms

If $B = (f_1, f_2, \dots, f_n)$ is a basis of a linear space V , then the coordinate transformation $L_B(f) = [f]_B$ from V to \mathbb{R}^n is an isomorphism; in other words, any n -dimensional linear space V is isomorphic to, and can be transformed into, \mathbb{R}^n .

↳ (*) (if $n \neq \infty$)