MATH 2940 Study Guide

Matthew Mentis-Cort

January 24, 2023

Overview

1 Textbook Information

Lay, Lay, McDonald, “Linear Algebra with Applications”, 6th edition

2 Aids

Matrix Calculator

3 Sources

1. Lay, Lay, McDonald, “Linear Algebra with Applications”, 6th edition
2. Cornell Spring 2023 semester Canvas notes (Professor Frans Schalekamp)

4 Usage

This guide is meant to be a place to conveniently list most of the important information for the topic of linear algebra. It is also based on the Cornell course MATH 2940, as well as any other sources listed. This guide does not promise to cover all topics that one may be learning in a class they are participating in. It also does not promise to cover all the content from the Cornell course MATH 2940. This is best used either for review, or to read after reviewing the textbook on one’s own time and using this to review the main ideas.

5 Overview

5.1 Solving Linear Equations

- SoEs can have one of three solutions:
  1. exactly one solution
  2. no solutions
  3. infinitely many solutions
- linear equations can be quite long and cumbersome to write out, so we want a way to make the method for solving them more precise, and a shorter notation for writing them. Thus, **augmented** and **coefficient matrices**

  Example 5.1.1.

  $$\begin{aligned} x_1 - 2x_2 + x_3 &= 0 \\ x_2 - 8x_3 &= 8 \\ -4x_1 + 5x_2 + 9x_3 &= -9 \end{aligned}$$

  – augmented matrix (includes the right-hand sides of the equations):

  $$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -8 & 0 \\ -4 & 5 & 9 & -9 \end{bmatrix}$$

  – coefficient matrix (does not include right-hand sides of equations)

  $$\begin{bmatrix} 1 & -2 & 1 \\ 0 & 1 & -8 \\ -4 & 5 & 9 \end{bmatrix}$$

5.2 Matrices

Elementary Row operations (EROs)

1. **Replacement**

   Replacing row by sum of itself and multiple of another row
2. **Scaling**

   Multiply all entries in a row with a nonzero constant
3. **Interchange**

   Swap two rows

Solving SoEs with Matrices

- use Gauss-Jordan Elimination to get matrices in reduced echelon form and solve them
- once the **forward phase of G-J elimination** is done, the resulting matrix can be used to determine the solutions of the system:

- if there is a row such that $0 = b$ (where b is nonzero), i.e. a row that looks like $\begin{bmatrix} 0 & 0 & 0 & \dots & 0 & b \end{bmatrix}$, then the system is inconsistent. Otherwise, the system is consistent

- after **backward phase** is done, **parametric description** of solution of system can be determined
  - translate the matrix back into SoE. variables corresponding to pivot columns are **basic variables**, and the other variables are **free variables**.
  - Write the basic variables in terms of the free variables. This is known as the parametric description of the solution set.

Example 5.2.1. Reduce $\begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$ to reduced row echelon form.

In the below example, pivot positions are circled.

$$\begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix} \xrightarrow{①\leftrightarrow④} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

$$\xrightarrow{②\leftarrow②+1①} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & 6 & -6 \\ -2 & -3 & 0 & 3 & -1 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \xrightarrow{③\leftarrow③+2①} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & 6 & -6 \\ 0 & 5 & 10 & -15 & -15 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix}$$

$$\xrightarrow{③\leftarrow③+\text{-}(5/2)②} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & 6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & -3 & -6 & 4 & 9 \end{bmatrix} \xrightarrow{④\leftarrow④+\text{-}(3/2)②} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & 6 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -5 & 0 \end{bmatrix}$$

$$\xrightarrow{③\leftrightarrow④} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & 6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The matrix is now in echelon form. To reduce to reduced row echelon form, we need to use row operations to create 0s above the pivots. The following row operations achieve that.

$$\begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & 6 & -6 \\ 0 & 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{3 \leftarrow (-1/5) 3} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & -6 & -6 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{2 \leftarrow 2 + 6 3} \begin{bmatrix} 1 & 4 & 5 & -9 & -7 \\ 0 & 2 & 4 & 0 & -6 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{1 \leftarrow 1 + 9 3} \begin{bmatrix} 1 & 4 & 5 & 0 & -7 \\ 0 & 2 & 4 & 0 & -6 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

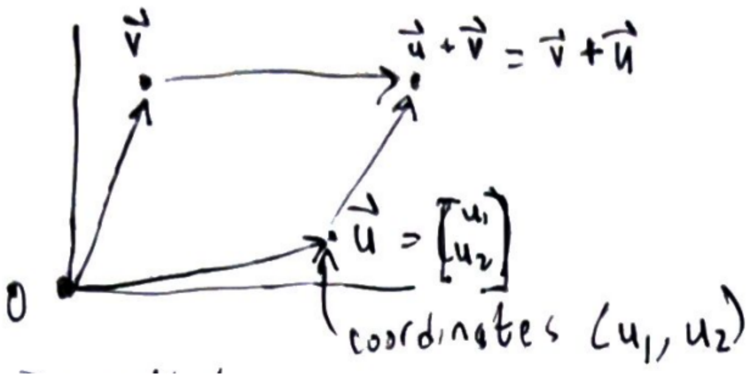
$$\xrightarrow{2 \leftarrow (1/2) 2} \begin{bmatrix} 1 & 4 & 5 & 0 & -7 \\ 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{1 \leftarrow 1 + -4 2} \begin{bmatrix} 1 & 0 & -3 & 0 & 5 \\ 0 & 1 & 2 & 0 & -3 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

5.3 Vectors

- solution to matrix can be written as a list of vectors $\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$
- vectors are equal when all entries are the same

Operations

- Addition
  - $\vec{u} + \vec{v}$ defined if $\vec{u}$ and $\vec{v}$ have the same number of entries
  - $\vec{u} + \vec{v} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix}$

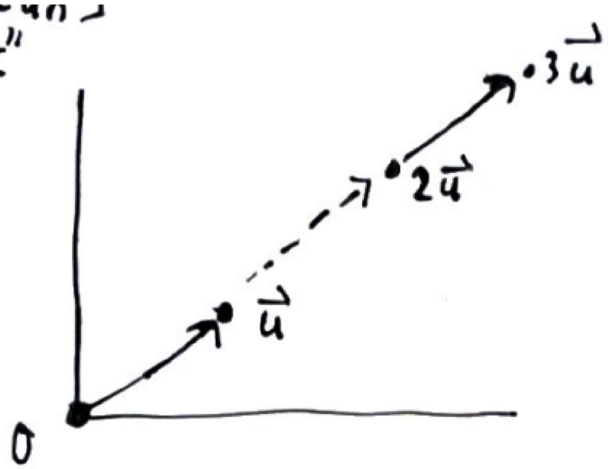


- Multiplication
  - scalar (real number) times vector only

$$- \; c\vec{u} = c \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} cu_1 \\ cu_2 \\ \vdots \\ cu_n \end{bmatrix}$$

c is called "scalar"

• geom. intuition



Properties

- zero vector $= \vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$
- $-\vec{u} = (-1)\vec{u}$
- $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
- $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
- $\vec{u} + \vec{0} = \vec{0} + \vec{u} = \vec{u}$
- $-\vec{u} + \vec{u} = \vec{0}$
- $(c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
- $(c + d)\vec{u} = c\vec{u} + d\vec{u}$
- $c(d\vec{u}) = (cd)\vec{u}$
- $1\vec{u} = \vec{u}$

Linear Combinations

- a **linear combination** is a vector written as the sum of other vectors
- the equation $\vec{y} = c_1\vec{v_1} + \ldots + c_p\vec{v_p}$
- $v_1, v_2, ..., v_p$ are vectors $c_1, c_2, ..., c_p$ are **weights**
- $\vec{y}$ is written as a linear combination of $v_1, v_2, ..., v_p$ with weights $c_1, c_2, ..., c_p$

Example 5.3.1. $\begin{bmatrix}4\\5\end{bmatrix}$ is a linear combination of $\begin{bmatrix}1\\1\end{bmatrix}$ and $\begin{bmatrix}2\\3\end{bmatrix}$ because $\begin{bmatrix}4\\5\end{bmatrix} = 2\begin{bmatrix}1\\1\end{bmatrix} + 1\begin{bmatrix}2\\3\end{bmatrix}$. The weights are 2 for $\begin{bmatrix}1\\1\end{bmatrix}$ and 1 for $\begin{bmatrix}2\\3\end{bmatrix}$.

5.4 Span

- because of our definition of linear combinations, we have the following relation:

  a vector equation
  $$x_1\vec{a_1} + x_2\vec{a_2} + \ldots + x_n\vec{a_n} = \vec{b}$$
  has the same solution set as the linear system with augmented matrix
  $$\begin{bmatrix}\vec{a_1} & \vec{a_2} & \ldots & \vec{a_n} & \vec{b}\end{bmatrix}$$

- solving this matrix is equivalent to finding weights in a linear combination
- the **span** of a set of vectors is set of all possible linear combinations of those vectors
- if an augmented matrix $\begin{bmatrix}\vec{a_1} & \vec{a_2} & \ldots & \vec{a_n} & \vec{b}\end{bmatrix}$ is consistent, then $\vec{b}$ is in the span of $\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$, or, in other words, $\vec{b} \in span\{\vec{a_1}, \vec{a_2}, \ldots, \vec{a_n}\}$
- the span of two vectors is a plane
- geometric intuition:

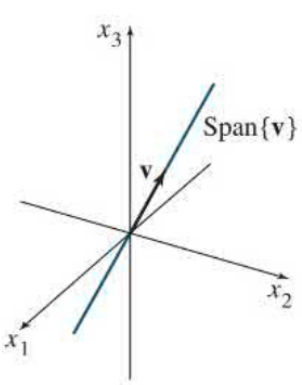


FIGURE 10 Span {**v**} as a line through the origin.

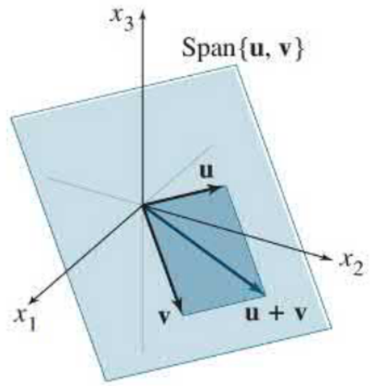


FIGURE 11 Span {**u**, **v**} as a plane through the origin.

5.5 The Matrix Equation $A\vec{x} = \vec{b}$

- fundamental idea of linear algebra is to view linear combo of vectors as product of matrix and vector (matrix vector multiplication)
- $A\vec{x}$ only defined when # columns of A = # entries of $\vec{x}$
- $A\vec{x}$ is a vector

Example 5.5.1.

$$\begin{aligned}\begin{bmatrix}1 & 2 & -1\\ 5 & 6 & 3\end{bmatrix}\begin{bmatrix}2\\7\\3\end{bmatrix} &= 2\begin{bmatrix}1\\5\end{bmatrix}+7\begin{bmatrix}2\\6\end{bmatrix}+3\begin{bmatrix}-1\\3\end{bmatrix}\\ &= \begin{bmatrix}2\\10\end{bmatrix}+\begin{bmatrix}14\\42\end{bmatrix}+\begin{bmatrix}-3\\9\end{bmatrix}\\ &= \begin{bmatrix}13\\61\end{bmatrix}\end{aligned}$$

- See Theorem 2 and Theorem 3
- use Gauss-Jordan Elimination see if system is consistent or not
- Properties of $A\vec{x}$ multiplication: Theorem 4

5.6 Homogenous and Nonhomogenous SoEs, and Parametric Vector Form

- **homogenous** SoEs are SoEs hat cna be written as $A\vec{x} = \vec{0}$
- homogenous SoEs are always consistent ($\vec{a} = \vec{0}$ is always solution, which is called the **trivial solution**)
- care about nontrivial solutions to the equation since we know that there is always a solution
- can write solution set of matrix equation as a vector using parametric description of solution set, known as **parametric vector form**
- to write in parametric vector form, see Parametric Vector Form

  Example 5.6.1. The vector form of the parametric description

  $$\begin{aligned}x_1 &= (4/3)x_3\\ x_2 &= 0\\ x_3 & \, is \, free\end{aligned}$$

  is

  $$\vec{x} = \begin{bmatrix}(4/3)x_3\\0\\x_3\end{bmatrix} = x_3\begin{bmatrix}4/3\\0\\1\end{bmatrix}$$

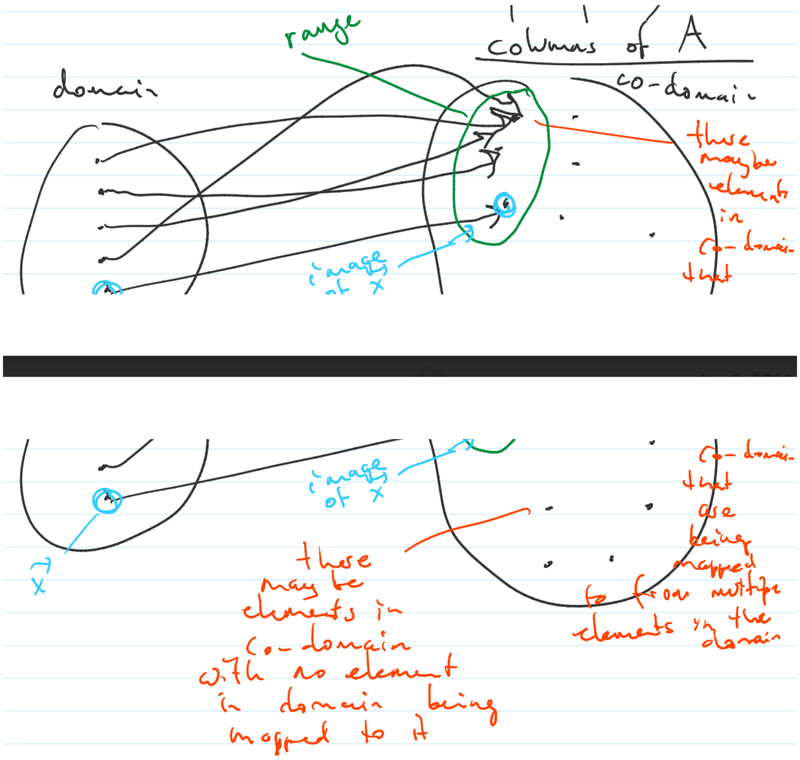
- Non-homogenous SoEs have same parametric vector form solution as $A\vec{x} = \vec{0}$, but with extra vector added that comes from $\vec{b}$ (see Theorem 5)

5.7 Linear Dependence

- see definitions for **linear independence and linear dependence**
- columns of a matrix A are linearly indep. $\Leftrightarrow A\vec{x} = \vec{0}$ has only trivial solution
- see Checking Linear Independence, Theorem 6, Theorem 7, and Theorem 8

5.8 Transformations

- a transformation T is a function that maps a vector $\vec{x}$ to a vector $A\vec{x}$
- $T(\vec{x}) = A\vec{x}$ for a matrix A. Notation: $T : \mathbb{R}^n \to \mathbb{R}^m$
- T has a **domain** and **co-domain**. For a $\vec{x}$, the **image** of $\vec{x}$ is $T(\vec{x}) = A\vec{x}$, and the **range** is all possible images under T.



- $\vec{x} \mapsto A\vec{x}$ means $\vec{x}$ maps to $A\vec{x}$
- Properties of linear transformations:
  1. $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$
  2. $T(c\vec{u}) = cT(\vec{u})$
- these properties lead to the following facts:
  1. $T(\vec{0}) = \vec{0}$ (though the $\vec{0}$ vectors can be in different spaces)
  2. $T(c\vec{u} + d\vec{v}) = cT(\vec{u}) + dT(\vec{v})$.
  3. $T(c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_p\vec{v}_p) = c_1T(\vec{v}_1) + c_2T(\vec{v}_2) + \ldots + c_pT(\vec{v}_p)$
- linear transformations can also be **onto** (all points in co-domain have at least one arrow pointing in) and/or **one-to-one** (no points in co-domain with more than 1 arrow pointing in)
- See Theorem 9, Theorem 10, Theorem 11

5.9 Key Matrices and Matrix Algebra

- think of matrices as structure with vectors as columns

$$
\begin{aligned}
A &= \begin{bmatrix} \vec{a_1} & \vec{a_2} & \ldots & \vec{a_n} \end{bmatrix} \\
&= \begin{bmatrix} a_{11} & a_{11} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}
\end{aligned}
$$

- **diagonal entries** are $a_{11}, a_{22}, a_{33}, \ldots$, which form **main diagonal** of A.
- **diagonal matrix** (only diagonal entries are nonzero) $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, **identity matrix** (diagonal matrix where diagonal entries are 1) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, and **zero matrix** (all entries 0) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ are also important to know about
- Addition:
  - $A + B$ only defined if A and B are the same size
  - result is sum of corresponding entries in A and B
- Scalar Multiplication:
  - $rA = \begin{bmatrix} r\vec{a_1} & r\vec{a_2} & \ldots r\vec{a_n} \end{bmatrix}$
- See Theorem 12 for properties of the operations
- Matrix Multiplication:
  - only defined if # rows of B = # columns of A
  - $AB = \begin{bmatrix} A\vec{b_1} & A\vec{b_2} & \ldots & A\vec{b_p} \end{bmatrix}$
  - if A is $m \times n$ matrix and B is $n \times p$ matrix, size of $AB = m \times p$
  - ORDER MATTERS!!!!! ($AB \neq BA$ generally, even if both multiplications are defined)
  - if $AB = 0$ one cannot conclude that $A = 0$ or $B = 0$
  - See Theorem 13 for properties of matrix multiplication
- Powers of a square matrix:
  - $A^k = A \times A \times A \times \ldots \times A$ (k times)
  - $A^0 = I$
- Transpose:
  - transpose of matrix is matrix where rows of original matrix become columns of the new matrix and vice versa

Example 5.9.1. $\begin{bmatrix} 1 & 7 & 9 \\ 6 & 4 & 2 \\ 3 & 4 & 3 \end{bmatrix}^T = \begin{bmatrix} 1 & 6 & 3 \\ 7 & 4 & 4 \\ 9 & 2 & 3 \end{bmatrix}$

– See Theorem 14 for properties of matrix transposes

5.10 Invertible Matrices

- invertibility only applies to square matrices, so matrices in this section are all square ($n \times n$)
- a matrix is invertible if there exists a matrix C such that $AC = I$ and $CA = I$. C is the inverse of A (denoted as A^{-1}) in this case.
- non-invertible matrices are sometimes called **singular matrices**, and invertible matrices are sometimes called **nonsingular matrices**.
- the matrix $\begin{bmatrix} 0 \end{bmatrix}$ is not invertible
- see Theorem 15 and Theorem 16 and Theorem 17
- To find inverse of matrix: see Finding the Inverse of an Invertible Matrix
- IMPORTANT THEOREM: Theorem 18
- a transformation T is invertible if there exists if there exists a transformation S such that
  i. $S(T(\vec{x})) = \vec{x}$ for all $\vec{x}$ in the domain of T
  ii. $T(S(\vec{b})) = \vec{b}$ for all $\vec{b}$ in the domain of S
- see Theorem 19

5.11 LU Factorization

- factorization that makes it easier to solve for $A\vec{x} = \vec{b}$ faster
- $LandU$ have a special structure:
  a. L is "lower triangular" with 1s on diagonal
  b. U is the echelon form of A (so "upper triangular", 0s below diagonal)

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{U} \tag{1}$$

- Method: see LU Factorization Algorithm and Using LU to solve $A\vec{x} = \vec{b}$

5.12 Vector Spaces and Subspaces

- a **vector space** is nonempty set, V, of objects (called vectors). Two operations can be performed on these vectors: addition, scalar multiplication
- 10 Vector axioms for vector space V, vectors $\vec{u}, \vec{v}, \vec{w}$, scalars $candd$:
  1. $\vec{u} + \vec{v} \in V$
  2. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
  3. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
  4. there is $\vec{0} \in V$ such that $\vec{u} + \vec{0} = \vec{u}$ for all $\vec{u} \in V$
  5. there is $\vec{-u} \in V$ such that $\vec{u} + (\vec{-u}) = \vec{0}$
  6. $c\vec{u} \in V$
  7. $c(\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
  8. $(c + d)\vec{u} = c\vec{u} + d\vec{u}$
  9. $c(d\vec{u}) = (cd)\vec{u}$
  10. $1\vec{u} = \vec{u}$
- from the axioms, we have
  1. $0\vec{u} = \vec{0}$
  2. $c\vec{0} = \vec{0}$
  3. $\vec{-u} = (-1)\vec{u}$
- polynomials of a degree n can be a vector space $\mathbb{P}^n$
- **subspace** of vector space V is a subset, H, of V that has three properties:
  1. H is a nonempty subset of $V \Leftrightarrow \vec{0} \in H$
  2. $\vec{u}, \vec{v} \in H$ implies $\vec{u} + \vec{v} \in H$ (closed under addition)
  3. $\vec{u} \in H$ and scalar c implies $c\vec{u} \in H$ (closed under multiplication)
- a subspace is vector space, vector space is subspace
- see Theorem 20

5.13 The Fundamental Spaces

Given an $m \times n$ matrix A, and a linear transformation T:

null space
: the solution set to $A\vec{x} = \vec{0}$. denoted by Nul$A$. Nul$A = \{\vec{x} : A\vec{x} = \vec{0}\}$ is the fancy math way of saying it. Nul$A$ is a vector space, because it is a subspace of $\mathbb{R}^n$.

column space
: the set of all linear combinations of the columns of A. denoted by Col$A$ If $A = \begin{bmatrix} \vec{a_1} & \dots & \vec{a_n} \end{bmatrix}$ then Col$A = \{\vec{b} : A\vec{x} = \vec{b} for some \vec{x}\} = \text{Span}\{a_1, \dots, a_n\}$. It is a subspace of $\mathbb{R}^m$.

row space
: the set of all linear combinations of the rows of A. denoted by Row$A$. Row$A$ = ColA^T.

kernel
: the set of all $\vec{u}$ such that $T(\vec{x}) = \vec{0}$. Kernel$(T) = \{\vec{x} : T(\vec{x}) = \vec{0}\} =$ Nul$A$

- see Solving for the Fundamental Spaces to solve for these mathematically

5.14 Bases, Bases for Fundamental Spaces, Change of Bases

- a **basis** is a linearly independent set of vectors that span a subspace
- If H is a subspace of V, a set of vectors $B \subseteq V$ is a **basis** for H means
  i. B is a linearly independent set
  ii. $H = \text{span}\{B\}$
- see Theorem 21, Theorem 22, Theorem 23, Theorem 24
- see Solving for the Bases of the Fundamental Spaces

5.15 Coordinate Systems

- see Theorem 25

  Given $\vec{x} = c_1\vec{b_1} + c_2\vec{b_2} + \ldots + c_n\vec{b_n}$, where $B = \{b_1, \ldots, b_n\}$:

- the **coordinates** of $\vec{x}$ relative to B are the weights $c_1, \ldots, c_n$ ($\begin{bmatrix}\vec{x}\end{bmatrix}_B = \begin{bmatrix}\vec{c_1} \\ \vec{c_2} \\ \vdots \\ \vec{c_n}\end{bmatrix}$)
- the **change of coordinates** matrix is a matrix such that when the coordinates of $\vec{x}$ relative to B is multiplied by it, $\vec{x}$ is produced. It is denoted P_B ($P_B = \begin{bmatrix}b_1 & b_2 & \ldots & b_n\end{bmatrix}$)
- $\vec{x} = P_B \begin{bmatrix}\vec{x}\end{bmatrix}_B$, $P_B^{-1}\vec{x} = \begin{bmatrix}\vec{x}\end{bmatrix}_B$
- P_B^{-1} is change-of-coordinates matrix from standard basis to B
- $\vec{x} \mapsto \begin{bmatrix}\vec{x}\end{bmatrix}_B$ is one-to-one and onto linear transformation
- $T(\vec{x}) = \begin{bmatrix}\vec{x}\end{bmatrix}_B$ is a linear transformation, one-to-one, onto
- an **isomorphism** is a linear transformation that is both one-to-one and onto.
- **change-of-coordinate-matrix** is a square matrix that converts from coordinates in one base to another. denoted $\underset{C \leftarrow B}{P}$ for bases B, C
- Given base $B = \{b_1, \ldots, b_n\}$ and $C = \{c_1, \ldots, c_n\}$ are bases of vector space V:

$$\begin{bmatrix}\vec{x}\end{bmatrix}_C = \underset{C \leftarrow B}{P} \begin{bmatrix}\vec{x}\end{bmatrix}_B$$

$$\underset{C\leftarrow B}{P} = \left[\left[\vec{b_1}\right]_C \quad \left[\vec{b_2}\right]_C \quad \cdots \quad \left[\vec{b_n}\right]_C\right]$$

- $(\underset{C\leftarrow B}{P})^{-1} = \underset{B\leftarrow C}{P}$
- $\left[\vec{c_1} \quad \vec{c_2} \;\vdots\; \vec{b_1} \quad \vec{b_2}\right] \sim \left[I \;\vdots\; \underset{C\leftarrow B}{P}\right]$

5.16 Dimensions

- **dimension**: number of vectors in a basis of a vector space V (provided there are a finite number of vectors). denoted $\dim(V)$.
- see Theorem 27, Theorem 26
- $\dim(\vec{0}) = 0$
- $\text{rank}A = \dim(\text{Col}A) = \dim(\text{Row}A) =$ number of pivot columns of $A =$ number of rows with a pivot in EF of A
- $\dim(\text{Nul}A) =$ nullity of A $= \#$ of non pivot columns in A.
- see Theorem 28, Theorem 18 (adds some new content with this knowledge), Theorem 29

5.17 Determinants

- the following only applies for **square** matrices
- **determinant** is the volume of a geometric object with edges as the rows of A
- $\det A = 0 \Leftrightarrow A$ is not invertible
- Properties of determinants:
  1. $\det I = 1$
  2. if A has 2 rows that are the same, $\det A = 0$
  3. If a matrix has a row of 0s, then $\det A = 0$
  4. replacement operations do not change determinant
  5. the determinant of a matrix B obtained by a row interchange on a matrix is $-\det A$
  6. scaling operation by a scalar k makes the determinant $k\det S$
  7. the determinant of a diagonal or triangular matrix is the product of its diagonal entries
  8. $\det A = \det A^T$
  9. $\det AB = (\det A)(\det B)$, $\det A^2 = (\det A)^2$
  10. $\det A^{-1} = \frac{1}{\det A}$

- $\det(\begin{bmatrix} a & b \\ c & d \end{bmatrix}) = ad - bc$
- to solve for determinant of matrix, see Finding the Determinant (Co-factor Expansion)

5.18 Entering the Eigenarena

- the following only applies to **square** matrices
- **eigenvalues** (denoted λ) are scalars that fulfill the $A\vec{x} = \lambda\vec{x}$ (basically saying that multiplying A by $\vec{x}$ gives a scaled version of $\vec{x}$), and **eigenvectors** are the $\vec{x}$ that correspond to the eigenvalue in the equation
- if given an eigenvalue (λ) of a matrix A, finding eigenvectors corresponding to λ corresponds to solving $(A - \lambda I)\vec{x} = \vec{0}$. the solution set to this system is called the **eigenspace** corresponding to λ (which is essentially $\text{Nul}A - \lambda I$)
- eigenvalues of triangular matrix are the values on its main diagonal
- see Theorem 30
- **characteristic equation** of A is $\det(A - \lambda I) = 0$. To find eigenvalues of A, solve this equation
- to find eigenspaces and eigenvalues, see Finding Eigenspaces and Eigenvalues
- if A has eigenvalues $\lambda_1, \dots, \lambda_n$, then A^{-1} has eigenvalues $\frac{1}{\lambda_1}, \dots \frac{1}{\lambda_n}$
- see Theorem 31
- A and B are **similar** means there exists an invertible matrix P such that $A = PBP^{-1}$
- see Theorem 32

5.18.1 Transformations in the Eigenarena

- λ and $\vec{x}$ are an eigenvalue and eigenvector of T means $T(\vec{x}) = \lambda\vec{x}$
- to find matrix for T relative to basis B:

$$M = \left[\left[T(\vec{b_1})\right]_B \quad \left[T(\vec{b_2})\right]_B \quad \dots \left[T(\vec{b_n})\right]_B \right]$$

5.19 Diagonalization (another factorization)

- the following only applies to **square** matrices
- a matrix A is **diagonalizable** if it is similar to a diagonal matrix D, which is the same thing as saying there exists an invertible matrix P such that $A = PDP^{-1}$
- if a matrix is diagonalizable, it is very easy to find A^k
- see Theorem 33, Theorem 34
- to diagonlize a matrix, see Diagonlizing Matrices
- $A^k = PD^kP^{-1}$

5.20 Power Method

- method to find approximate eigenvalues and eigenvectors using a computer
- to estimate a strictly dominant eigenvalue Using the Power Method

5.21 Complex Eigenvalues

- eigenvalues can unfortunately be complex numbers
- matrices with complex eigenvalues correspond to rotations in terms of transformations
- A^k will have real eigenvalues if A has real eigenvalues
- complex numbers have a real part and imaginary part. $\text{Re}(z)$ represents the real part of the complex number z, and $\text{Im}(z)$ is the imaginary part. $z = \text{Re}(z) + i\text{Im}(z)$
- **Properties of complex number algebra** (B is an $m \times n$ matrix, $\bar{B}$ is a matrix whose entries are the complex conjugates of entries in B, r is a complex number, $\vec{x}$ is a vector):
  - $\overline{r\vec{x}} = \overline{r}\overline{\vec{x}}$
  - $\overline{B\vec{x}} = \overline{B}\overline{\vec{x}}$
  - $\overline{BC} = \overline{BC}$
  - $\overline{rB} = \overline{r}\overline{B}$
- because of these properties, if $\lambda = \text{Re}(z) + i\text{Im}(z)$, then $\bar{\lambda} = \text{Re}(z) - i\text{Im}(z)$ is also an eigenvalue of A, where $\bar{\lambda}$ is the complex conjugate of A
- If A is real 2×2 matrix with complex eigenvalues $\lambda_1 = a + ib$ and $\lambda_2 = a - ib$ ($b \neq 0$) and corresponding eigenvector for λ_1 $\vec{v}$, then:

$$A = PCP^{-1}$$
$$C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$
$$P = \begin{bmatrix} \text{Re}(\vec{v}) & \text{Im}(\vec{v}) \end{bmatrix}$$

5.22 Dynamical Systems

- a **dynamical system** is a system that evolves over time
- is described by the difference equation $\vec{x_{k+1}} = A\vec{x_k}$
- a **trajectory** is $\vec{x_k}$ when it is graphed starting with different $\vec{x_0}$
- an **attractor** of a dynamical system is the point where all the trajectories in the system head to

5.23 Dot Products and Orthogonality

- **inner product/dot product** of two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$ is $\vec{u}^T\vec{v} = u_1v_1 + u_2v_2 + \ldots + u_nv_n = \vec{u} \cdot \vec{v}$

- Properties of dot product (for any $\vec{u}$, $\vec{v}, \vec{w} \in \mathbb{R}^n$, scalar c):
  i. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
  ii. $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$
  iii. $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot c\vec{v}$
  iv. $\vec{u} \cdot \vec{u} \geq 0$ and $\vec{u} \cdot \vec{u} = 0 \Leftrightarrow \vec{u} = \vec{0}$
  v. scaling by scalar c $\|c\vec{v}\| = c\|\vec{v}\|$
- **length/norm** of a vector is $\sqrt{\vec{x} \cdot \vec{x}}$, denoted $\|\vec{x}\|$
- **unit vector** is a vector with length 1 that points in the same direction of $\vec{v}$. The calculation for it is $\frac{1}{\|\vec{v}\|}$
- the **distance** between $\vec{u}$ and $\vec{v}$ is $\text{dist}(\vec{u}, \vec{v}) = \|\vec{u} - \vec{v}\| = \|\vec{v} - \vec{u}\|$
- orthogonal vectors are vectors that are perpendicular to each other (90° angle between each other). Abstractly, this means that for two vectors $\vec{u}$ and $\vec{v}$, $\vec{u} \cdot \vec{v} = 0$.
- see Theorem 35
- $\vec{u} \cdot \vec{v} = \|\vec{u}\|\|\vec{v}\| \cos\theta$
- Given W is a subspace of $\mathbb{R}^n$, $\vec{z} \in \mathbb{R}^n$ is orthogonal to W means $\vec{z}$ is orthogonal to every $\vec{w} \in W$, an **orthogonal complement** is the set of all vectors orthogonal to subspace W. denoted $W^{\perp}$. pronounced "W perp"
- $\vec{x} \in W^{\perp} \Leftrightarrow \vec{x}$ is orthogonal to $\vec{w_i}$ for all $i = 1, 2, \ldots, k$.
- $W^{\perp}$ is a subspace of $\mathbb{R}^n$
- see Theorem 36

5.24 Orthogonal Sets and Projections

- an **orthogonal set** is a set of vectors in which every vector is orthogonal to every vector besides itself (**pairwise orthogonal**). Defined mathematically, a set $\{u_1, \ldots, u_p\}$ is said to be an orthogonal set if $\vec{u_i} \cdot \vec{u_j} = 0$ whenever $i \neq j$
- see Theorem 37
- an **orthogonal basis** is an orthogonal set S that does not contain $\vec{0}$
- Given $W\{u_1, \ldots, u_p\}$ is orthogonal basis for subspace W:

  For each $\vec{y} \in W$, the weights in

  $$y = c_1\vec{u_1} + c_2\vec{u_2} + \ldots + c_p\vec{u_p}$$

  are given by

  $$c_j = \frac{\vec{y} \cdot \vec{u_1}}{\vec{u_j} \cdot \vec{u_j}} \; (j = 1, \ldots, p)$$

- $\vec{y} = \hat{y} + \vec{z}$ is a way of decomposing $\vec{y}$ into a sum of a vector that a multiple of another vector $\vec{u}$ and a vector orthogonal to $\vec{u}$. $\hat{y} = \alpha\vec{u}$ for some scalar α, and $\vec{z}$ is some vector orthogonal to $\vec{u}$

- **orthogonal projection** of $\vec{y}$ onto $\vec{u}$ is $\hat{y} = \text{proj}_L\vec{y} = \frac{\vec{y}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}$, where L is the subspace spanned by $\vec{u}$. The projection is onto L and not just $\vec{u}$ because even if $\vec{u}$ is scaled by a number $(c\vec{u})$, the projection will be the same
- an **orthonormal set** is an orthogonal set of unit vectors. any subset of an orthonormal set is also orthonormal. an **orthonormal basis** is an orthonormal set that spans a subspace
- see Theorem 38 and Theorem 39
- see Theorem 40, Theorem 41, and Theorem 42

5.25 Gram-Schmidt and QR-Factorization

- to create orthogonal bases, **Gram-Schmidt Process** was developed. It is the process of taking a basis an creating an orthogonal basis by taking the parts of the vectors that are orthogonal to the other vectors and using those as the vectors for the orthogonal basis.
- see Gram-Schmidt Process
- QR Factorization is a factorization of a **matrix with linearly independent columns** in which Q's columns are an orthonormal basis for Col$A$, and R is an upper triangular matrix that gives information about how the orthonormal basis was created
- see QR Factorization for how to perform this factorization

5.26 Least Squares Solution

- least squares solution is finding $\vec{x}$ that minimizes $\|A\vec{x} - \vec{b}\|$. this is the closest point to $\vec{b}$ in Col$A$. helpful when an SOE is inconsistent
- simplifies to just solving $A^TA\vec{x} = A^T\vec{b}$ (called the **normal equation**). Sometimes, the solution to the equation is denoted by $\hat{x}$.
- see Theorem 43 and Theorem 44

6 Important Theorems

Theorem 1. Each matrix is row equivalent to one and only one reduced echelon matrix.

Theorem 2. When A is $m \times n$ matrix with columns $a_1, \ldots, a_n$, and $\vec{b}$ in $\mathbb{R}^m$,

$$A\vec{x} = \vec{b}$$

has same solution set as

$$x_1\vec{a_1} + x_2\vec{a_2} + \ldots + x_n\vec{a_n} = \vec{b}$$

which has same solution set as

$$\begin{bmatrix} \vec{a_1} & \vec{a_2} & \ldots & \vec{a_n} & \vec{b} \end{bmatrix}$$

Theorem 3. Let A be an $m \times n$ matrix (coefficient matrix). Then the following statements are logically equivalent. That is, for a particular A, either they are all true statements or they are all false.

a. For each $\vec{b}$ in $\mathbb{R}^m$, the equation $A\vec{x} = \vec{b}$ has a solution.

b. Each $\vec{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.

c. The columns of A span $\mathbb{R}^m$.

d. A has a pivot position in every row.

Theorem 4. Let A be an $m \times n$ matrix, $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{R}^n$, and c is a scalar, then:

a. $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$

b. $A(c\vec{u}) = c(A\vec{u})$

Theorem 5. Suppose the equation $A\vec{x} = \vec{b}$ is consistent for some given $\vec{b}$, and let $\vec{p}$ be a solution. Then the solution set of $A\vec{x} = \vec{b}$ is the set of all vectors of the form $\vec{w} = \vec{p} + \vec{v_h}$, where $\vec{v_h}$ is any solution of the homogeneous equation $A\vec{x} = \vec{0}$.

Theorem 6. A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is linearly dependent if one of the vectors is a linear combination of the others. $(v_j \in span v_1, v_2, \ldots, v_{j-1}, v_{j+1}, \ldots, v_p)$ for some j.

Theorem 7. A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is linearly dependent if there are more vectors than entries in each vector. Any set $\{\vec{v}_1, \ldots, \vec{v}_p\}$ in $\mathbb{R}^n$ is linearly dependent if $p > n$.

Theorem 8. A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is linearly dependent if it contains the zero vector.

Theorem 9. Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ there exists a unique matrix A such that

$$T(\vec{x}) = A\vec{x}$$

namely

$$A = \begin{bmatrix} T(\vec{e_1}) & T(\vec{e_2}) & \ldots & T(\vec{e_n}) \end{bmatrix}$$

where

$$\vec{e_1} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \ \vec{e_2} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \ \vec{e_j} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \leftarrow j^{th} entry \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$\vec{e}_1, \ldots, \vec{e}_j$ are known as **standard basis vectors**. The matrix A is the **standard matrix for the linear transformation T**.

Theorem 10. Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, T is one-to-one exactly when $T(\vec{x}) = \vec{0}$ has only the trivial solution.

Theorem 11. Given a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ with standard matrix A:

a. T is onto $\Leftrightarrow$ span of columns of $A = \mathbb{R}^m$.

b. T is one-to-one $\Leftrightarrow$ $A\vec{x} = \vec{0}$ only has trivial solution, $\Leftrightarrow$ columns of A are linearly independent.

Theorem 12. Given matrices A, B, C of the same size, and r, s are scalars:

a. $A + B = B + A$

b. $(A + B) + C = A + (B + C)$

c. $A + 0 = A$

d. $r(A + B) = rA + rB$

e. $(r + s)A = rA + sA$

f. $r(sA) = (rs)A$

Theorem 13. Given A, B, C matrices such that sums and products below are defined:

a. $A(BC) = (AB)C$

b. $A(B + C) = AB + AC$

c. $(B + C)A = BA + CA$

d. $r(AB) = (rA)B = A(rB)$

e. $IA = A = AI$

Theorem 14. Given a matrix A:

a. $(A^T)^T = A$

b. $(A + B)^T = A^T + B^T$

c. $(rA)^T = rA^T$

d. $(AB)^T = B^T A^T$

Theorem 15. Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, if $ad - bc \neq 0$ then A is invertible and $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

Theorem 16. If A is invertible $n \times n$ matrix, then for every $\vec{b}$ in $\mathbb{R}^n$, $A\vec{x} = \vec{b}$ has the unique solution $\vec{x} = A^{-1}\vec{b}$.

Theorem 17. Given A, B invertible matrices, then

a. A^{-1} is also invertible and $(A^{-1})^{-1} = A$

b. AB is also invertible and $(AB)^{-1} = B^{-1}A^{-1}$

c. A^T is also invertible and $(A^T)^{-1} = (A^{-1})^T$

Theorem 18 (Invertible Matrix Theorem). If A is an $n \times n$ matrix, then the following statements are equivalent. That is, for a given A, the statements are either all true or all false.

a. A is an invertible matrix

b. $A\vec{x} = \vec{b}$ has unique solution for every $\vec{b} \in \mathbb{R}^n$.

c. A has n pivot positions, and a pivot position in every row and column. (because A is square)

d. A has reduced echelon form I

e. $A\vec{x} = \vec{0}$ only has trivial solution

f. the columns of A are linearly independent

g. columns of A span $\mathbb{R}^n$

h. A^T is invertible

i. the transformation $\vec{x} \mapsto A\vec{x}$ is onto

j. the transformation $\vec{x} \mapsto A\vec{x}$ is one-to-one

k. there exists an $n \times n$ matrix C such that $CA = I$

l. there exists an $n \times n$ matrix D such that $AD = I$

m. The columns of A form a basis of $\mathbb{R}^n$

n. $\text{Col}A = \mathbb{R}^n$

o. $\text{rank}A = n$

p. $\text{nullity}A = 0$

q. $\text{Nul}A = \{\vec{0}\}$

Theorem 19. A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ with standard matrix A is invertible $\Leftrightarrow$ A is an invertible matrix. In that case, $\vec{b} \mapsto A^{-1}\vec{b}$ is the inverse function of T.

Theorem 20. If $\vec{v}_1, \ldots, \vec{v}_p$ are in a vector space V, then $Span\{\vec{v}_1, \ldots, \vec{v}_p\}$ is a subspace of V.

Theorem 21. $\{v_1, \ldots, v_p\}$ $(p \geq 2)$ is linearly dependent $\Leftrightarrow$ one of the vectors is a linear combination of the others

Theorem 22 (The Spanning Set Theorem)**.** Given $H = \text{span}\{\vec{v}_1, \ldots, \vec{v}_p\}$ we can repeatedly find a vector that is a linear combination of the others and delete this vector from the generating set until we have a linearly independent set of vectors that still spans S.

Theorem 23. The pivot columns of matrix A form a basis for Col$A$.

Theorem 24. The row space of a matrix does not change after performing one row operation. (Row$A$ = Row$B$ if B can be obtained from A by performing 1 row operation)

Theorem 25 (Unique Representation Theorem)**.** Given $B = \{b_1, \ldots, b_n\}$ be a basis for vector space V. For any $\vec{x} \in V$ there is a unique set of scalars $c_1, \ldots, c_n$ such that $\vec{x} = c_1\vec{b_1} + c_2\vec{b_2} + \ldots + c_n\vec{b_n}$

Theorem 26. V vector space, with basis $B = \vec{b}_1, \ldots, \vec{b}_n$: Any set of vectors in V with more than n vectors ($n = \#$ vectors in B) is linearly dependent.

Theorem 27. Every basis for vector space V has same # of vectors.

Theorem 28 (Rank-Nullity Theorem)**.** rank$A$ + nullity of $A = n$

Theorem 29. Given vector space V, dim$V$ = p,

a. Any linearly independent set of size p is a basis for V.

b. Any set of p vectors that spans V is a basis for V.

Theorem 30. Given distinct eigenvalues $\lambda_1, \ldots, \lambda_p$ with eigenvectors $\vec{x}_1, \ldots, \vec{x}_p$, $\{\vec{x}_1, \ldots, \vec{x}_p\}$ is linearly independent.

Theorem 31. A is invertible $\Leftrightarrow$ 0 is not an eigenvalue of A.

Theorem 32. Similar matrices have same characteristic equation and same eigenvalues

Theorem 33. $n \times n$ matrix A is diagonlizable $\Leftarrow$ there are n eigenvectors of A that are linearly independent. P can be constructed using these eigenvectors as columns, and D can be constructured by putting the corresponding eigenvalues of A on the diagonal (in the same order as the eigenvectors). To visualize this:

$$A = \underbrace{\begin{bmatrix} \vec{p_1} & \vec{p_2} & \ldots & \vec{p_n} \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} \vec{\lambda_1} & 0 & \ldots & 0 \\ 0 & \vec{\lambda_2} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \ldots & 0 & \vec{\lambda_n} \end{bmatrix}}_{D} P^{-1}$$

where $\{p_1, \ldots, p_n\}$ is the set of eigenvectors, $\{\lambda_1, \ldots, \lambda_n\}$ is the set of eigenvalues, and λ_i is the eigenvalue corresponding to eigenvector $\vec{p_i}$.

Theorem 34. An $n \times n$ matrix with n eigenvectors is diagonlizable.

Theorem 35 (Pythagorean Theorem)**.** $\vec{u}$ and $\vec{v}$ are orthogonal $\Leftrightarrow$ $\|\vec{u} + \vec{v}\|^2 = \vec{u}^2 + \vec{v}^2$

Theorem 36. A is $m \times n$ matrix:

a. $(\text{Row}A)^\perp = \text{Nul}A$

b. $(\mathrm{Col}A)^{\perp} = \mathrm{Nul}A^T$

Theorem 37. An orthogonal set S is linearly independent if $\vec{0} \notin S$

Theorem 38. An $m \times n$ matrix U has orthonormal columns $\Leftrightarrow U^T U = I$.

Theorem 39. If U is matrix with orthonormal columns, and $\vec{x}, \vec{y} \in \mathbb{R}^n$, then

a. $\|U\vec{x}\| = \|\vec{x}\|$

b. $(U\vec{x}) \cdot (U\vec{y}) = \vec{x} \cdot \vec{y}$

c. $(U\vec{x}) \cdot (U\vec{y}) = 0 \Leftrightarrow x \cdot y = 0$

Theorem 40 (Orthogonal Decomposition Theorem)**.** Given a subspace W, and a vector $\vec{y}$ there is a unique way of writing

$$\vec{y} = \hat{y} + \vec{z}$$

In fact, if $\{u_1, \ldots, u_p\}$ is an orthogonal basis for W, then

$$\mathrm{proj}_W \vec{y} = \frac{\vec{y} \cdot \vec{u_1}}{\vec{u_1} \cdot \vec{u_1}} \vec{u_1} + \frac{\vec{y} \cdot \vec{u_2}}{\vec{u_2} \cdot \vec{u_2}} \vec{u_2} + \ldots + \frac{\vec{y} \cdot \vec{u_p}}{\vec{u_p} \cdot \vec{u_p}} \vec{u_p}$$

Additionally, $\mathrm{proj}_W \vec{y}$ is W, and $\vec{y} - \mathrm{proj}_W \vec{y}$ is in $W^{\perp}$. Also, if $\vec{y} \in W$, this is just writing $\vec{y}$ as a linear combination of the basis vectors.

Theorem 41 (Best Approximation Theorem)**.** $\mathrm{proj}_W \vec{y}$ is the vector closest to $\vec{y}$ in W, meaning $\mathrm{dist}(\vec{y}, \mathrm{proj}_W \vec{y}) < \mathrm{dist}(\vec{y}, \vec{w})$ (for all $\vec{w} \in W$, $\vec{w} \notin \mathrm{proj}_W \vec{y}$).

Theorem 42. If $\{u_1, \ldots, u_p\}$ orthonormal basis for W, then

$$\mathrm{proj}_W \vec{y} = QQ^T \vec{y}$$

where

$$Q = \begin{bmatrix} \vec{u_1} & \ldots & \vec{u_p} \end{bmatrix}$$

Theorem 43. The set of least-squares solutions $A\vec{x} = \vec{b}$ is equal to the solution set of $A^T A\vec{x} = A^T \vec{b}$.

Theorem 44. If A is $m \times n$ matrix, the following is either all true or all false:

a. The equation $A\vec{x} = \vec{b}$ has a unique least squares solution for each $\vec{b}$ in $\mathbb{R}^m$.

b. The columns of A are linearly independent.

c. The matrix $A^T A$ is invertible.

7 Toolbox

Make sure to check your work after using any of these methods!

7.1 Linear Equations

To solve systems of linear equations:

1. Variable Elimination
   - System of 2 variables
     (a) Eliminate one of the variables by adding or subtracting the equations together. This will give you can equation to solve for the other variable. Solve for this value.
     (b) Plug the value for the variable that was just solved for into either of the equations to solve for the value of the other variable.
   - System of 3 or more variables
     (a) Choose a variable in one equation to eliminate from the other equations. Eliminate this variable by adding or subtracting the equations together. Repeat the first step until you are able to solve for one variable.
     (b) Plug the value for the variable that was just solved for into either of the equations to solve for the value of the another variable. Repeat this step until all variables are solved for.
2. Substitution
   (a) Define one variable in terms of another, and plug it into all other equations. Use this equation to solve for this variable. If the system of equations has 3 or more equations, use all the other equations to solve for one variable in terms of another.

7.2 Gauss-Jordan Elimination

1. Create augmented matrix for the SoE.
2. Use first nonzero entry in the leftmost column as first pivot. If necessary, interchange rows so that there is a nonzero entry in the top entry of the first column.
3. Use EROs to create all 0s below the pivot position in the first column.
4. Ignore the row with the pivot position, and apply steps 2-3 to the submatrix that remains until there are no more nonzero rows to reduce.
   - after this step, one can determine the pivot positions and pivot columns of the matrix
5. Create zeroes above each pivot, starting with the rightmost pivot. If a pivot is not a 1, make it a 1 with a scaling operation.

7.3 Parametric Vector Form

1. Find the parametric description of the solution set for the SoE. Write all the basic variables in terms of the free variables.
2. Write the parametric description as a vector $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, substituting each variable for its actual value.
3. Write each free variable as the product of itself and a vector $\vec{v_n} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$, where each entry of $\vec{v_n}$ is the corresponding coefficient of the variable in the vector $\vec{x}$.
4. Write all the free variables and the vectors they are multiplied by as a sum expression.

7.4 Checking Linear Independence

Given a set of vectors $\{\vec{v}_1, \dots, \vec{v}_p\}$:

1. Put the vectors into a matrix A
2. Solve for $A\vec{x} = \vec{0}$. If there is only the trivial solution, the S is linearly independent. Otherwise, it is linearly dependent.

7.5 Finding the Inverse of an Invertible Matrix

To find the inverse of a matrix A:

1. create the matrix $M = \begin{bmatrix} A \mid I \end{bmatrix}$.
2. perform Gauss-Jordan Elimination on M. If the matrix is invertible, the reduced echelon form of M will be $\begin{bmatrix} I \mid A^{-1} \end{bmatrix}$.

In other words,

$$\begin{bmatrix} A \mid I \end{bmatrix} \xrightarrow{G.J} \begin{bmatrix} I \mid A^{-1} \end{bmatrix}$$

7.6 LU Factorization Algorithm

1. Reduce A to echelon form. As you are row reducing, when you get 0s underneath a column, place entries in the next column of L such that the same sequence of row operations will reduce L to the corresponding column of I. L will have as many columns as pivots of A. Note: Generally, the entries of L are the corresponding entries of A divided by the top entry of that column, with 1s along the main diagonal (see 7.6.1 below).
2. U is the echelon form of A.

Example 7.6.1.

$$\begin{bmatrix} 2 & 3 & 4 \\ 4 & 8 & 11 \\ 4 & 10 & 16 \end{bmatrix} \xrightarrow{\substack{\textcircled{3} \leftarrow \textcircled{3} + \text{-}2\textcircled{1} \\ \textcircled{2} \leftarrow \textcircled{2} + \text{-}2\textcircled{1}}} \begin{bmatrix} 2 & 3 & 4 \\ 0 & 2 & 3 \\ 0 & 4 & 8 \end{bmatrix}$$

We have created 0s under the first column. Now, we can create the first column of L. We ask ourselves, what column could be turned into $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ with the row operations $\substack{\textcircled{3} \leftarrow \textcircled{3} + \text{-}2\textcircled{1} \\ \textcircled{2} \leftarrow \textcircled{2} + \text{-}2\textcircled{1}}$? The answer is $\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$, so this is our first column of L. We repeat this process for the other pivots of A:

$$\begin{bmatrix} 2 & 3 & 4 \\ 0 & 2 & 3 \\ 0 & 4 & 8 \end{bmatrix} \xrightarrow{\textcircled{3} \leftarrow \textcircled{3} + \text{-}2\textcircled{2}} \begin{bmatrix} 2 & 3 & 4 \\ 0 & 2 & 3 \\ 0 & 0 & 2 \end{bmatrix}$$

We have created 0s under the second column. Now, we can create the second column of L. We ask ourselves, what column could be turned into $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ with the row operations $\textcircled{3} \leftarrow \textcircled{3} + \text{-}2\textcircled{2}$? The answer is $\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$, and this is our second column of L. For the last column, it is $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, because L is lower triangular. Therefore, $L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 2 & 1 \end{bmatrix}$.

You will notice from this example that **the entries of L are the corresponding entries of A divided by the top entry of that column, with 1s along the main diagonal**. This is generally the case, but you should always go through this whole algorithm to ensure you don't make mistakes.

7.7 Using LU to solve $A\vec{x} = \vec{b}$

$A\vec{x} = \vec{b}$ can be written as $L(U\vec{x}) = \vec{b}$ with LU factorization. Therefore, to solve $A\vec{x} = \vec{b}$:

1. Solve $L\vec{y} = \vec{b}$ for $\vec{y}$.
2. Solve $U\vec{x} = \vec{y}$ for $\vec{x}$.

7.8 Solving for the Fundamental Spaces

Given a matrix A:

To solve for **Nul**A:

1. Solve for $\vec{x}$ in parametric vector form in $A\vec{x} = \vec{0}$ (using G.J. elimination).
2. Nul$A$ is the Span of the set of vectors that the free variables are multiplied by.

Example 7.8.1. Given

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & 4 \end{bmatrix}$$

what is NulA?

First, reduce $A\vec{x} = \vec{0}$ to REF.

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 & 0 \\ 1 & -2 & 2 & 3 & -1 & 0 \\ 2 & -4 & 5 & 8 & 4 & 0 \end{bmatrix} \xrightarrow{G.J.} \begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Then, find the parametric vector form of the solution set.

$$\begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \Rightarrow \vec{x} = \begin{bmatrix} 2x_2 + x_4 - 3x_5 \\ x_2 \\ -2x_4 + 2x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix}$$

Now, we can conclude that $\text{Nul}A = \text{Span} \left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}$.

To solve for **Col**A:

1. If $A = \begin{bmatrix} \vec{a_1} & \dots & \vec{a_n} \end{bmatrix}$, then Col$A$ is Span$\{a_1, \dots, a_n\}$.

To solve for **Row**A:

1. If $A = \begin{bmatrix} \vec{r_1} \\ \vdots \\ \vec{r_n} \end{bmatrix}$, Row$A$ is Span$\{r_1, \dots, r_n\}$.

7.9 Solving for the Bases of the Fundamental Spaces

To solve for a basis of **Nul**A:

1. Solve for $\vec{x}$ in parametric vector form in $A\vec{x} = \vec{0}$ (using G.J. elimination).
2. Nul$A$ is the the set of vectors that the free variables are multiplied by.

To solve for a basis of **Col**A:

1. Use G.J. elimination to find the echelon form of A.
2. Identify the pivot columns of A
3. Col$A$ is the set of vectors that are in the pivot columns of the echelon form of A in the the original matrix A.

To solve for a basis of **Row**A:

1. Row$A$ is all the nonzero vectors in the original matrix A or in the echelon form of A.

7.10 Finding the Determinant (Co-factor Expansion)

Assuming the original matrix is A:

1. Find a row or column where there are all 0s except for one entry. If there is not already one, use row operations to make a row or column have all 0s except for one entry in that row or column. It does not matter which entry is you decide to make nonzero (so it does not need to be picked such that the reduced form of the matrix corresponds to echelon form). It's usually a good idea to pick a row or column with a lot of 0s already in it. We will denote the row that the nonzero entry is in with i, the column it is in with j, and the entry itself with m.
2. The determinant of your starting matrix will be $m(-1)^{i+j}B$, where B is the matrix obtained by deleting row i and column j from A.
3. Repeat step 1 until a number is generated from the equation.

For reference, the actual Co-factor expansion formula is

for using i-th row:

$$\det A = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}\det M_{ij}$$

or, for using j-th column:

$$\det A = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}\det M_{ij}$$

where M_{ij} is A with row i and column j deleted.

7.11 Finding Eigenspaces and Eigenvalues

To find an **eigenspace** corresponding to an eigenvalue:

1. Solve the equation $(A - \lambda I)\vec{x} = \vec{0}$. The set of vectors being multiplied by the free variables form a basis for the eigenspace corresponding to λ for A.

To find all **eigenvalues** of a matrix:

1. Solve the equation $\det(A - \lambda I) = 0$.

7.12 Diagonlizing Matrices

To diagonlize a matrix A:

1. Find the eigenvalues of A.
2. Find a set of linearly independent vectors equal to the number of columns (or rows, since A is square) of A (so basically just find eigenvectors corresponding to eigenvalues of A).

3. Construct P and D according to the structure outlined in Theorem 33:

$$A = \underbrace{\begin{bmatrix} \vec{p_1} & \vec{p_2} & \dots & \vec{p_n} \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} \vec{\lambda_1} & 0 & \dots & 0 \\ 0 & \vec{\lambda_2} & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & \vec{\lambda_n} \end{bmatrix}}_{D} P^{-1}$$

where $\{\vec{p}_1, \dots, \vec{p}_n\}$ is the set of eigenvectors, $\{\lambda_1, \dots, \lambda_n\}$ is the set of eigenvalues, and λ_i is the eigenvalue corresponding to eigenvector $\vec{p}_i$.

7.13 Using the Power Method

To estimate a strictly dominant eigenvalue for a given matrix A:

1. Select an initial vector $\vec{x_0}$, whose largest entry is 1.
2. For $k = 0, 1, \dots$:
   a. Compute $A\vec{x_k}$.
   b. Compute $\vec{x_{k+1}} = \frac{1}{\mu_k} A\vec{x_k}$ where μ_k is the entry in $A\vec{x_k}$ that has the largest absolute value.
3. As this process is applied to greater values of k, μ_k will approach the dominant eigenvalue, and $\vec{x_k}$ approaches a corresponding eigenvector.

7.14 Gram-Schmidt Process

Given a basis $\{x_1, \dots, x_p\}$, for a nonzero subspace W of $\mathbb{R}^n$, to create an orthogonal basis $\{v_1, \dots, v_p\}$ for W, use the following equalities:

$$\begin{aligned}
\vec{v_1} &= \vec{x_1} \\
\vec{v_2} &= \vec{x_2} - \frac{\vec{x_2} \cdot \vec{v_1}}{\vec{v_1} \cdot \vec{v_1}} \vec{v_1} \\
\vec{v_3} &= \vec{x_3} - \frac{\vec{x_3} \cdot \vec{v_1}}{\vec{v_1} \cdot \vec{v_1}} \vec{v_1} - \frac{\vec{x_3} \cdot \vec{v_2}}{\vec{v_2} \cdot \vec{v_2}} \vec{v_2} \\
&\vdots \\
\vec{v_p} &= \vec{x_p} - \frac{\vec{x_p} \cdot \vec{v_1}}{\vec{v_1} \cdot \vec{v_1}} \vec{v_1} - \frac{\vec{x_p} \cdot \vec{v_2}}{\vec{v_2} \cdot \vec{v_2}} \vec{v_2} - \dots - \frac{\vec{x_p} \cdot \vec{v_{p-1}}}{\vec{v_{p-1}} \cdot \vec{v_{p-1}}} \vec{v_{p-1}}
\end{aligned}$$

Explained in English, we take the first vector in the original basis and use it as the first vector in our orthogonal basis. Then, for every other vector in the original basis, we find the component of that vector orthogonal to each of the vectors in our orthogonal basis, and add this component to the orthogonal basis.

7.15 QR Factorization

Given an $m \times n$ matrix $A = \begin{bmatrix} \vec{x_1} & \dots & \vec{x_n} \end{bmatrix}$, where $\{x_1, \dots, x_n\}$ forms a basis for Col$A$ (linearly independent columns)

1. Use the Gram-Schmidt Process on $\{x_1, \ldots, x_n\}$ to form an orthogonal basis for Col$A$. We will call this orthogonal basis $\{v_1, \ldots, v_n\}$.
2. Normalize all the vectors (make them vectors of length 1 by dividing each of them by their lengths) in $\{v_1, \ldots, v_n\}$. We will call this orthonormal set $\{q_1, \ldots, q_n\}$.
3.

$$A = \underbrace{\begin{bmatrix} \vec{q_1} & \ldots & \vec{q_n} \end{bmatrix}}_{Q} \underbrace{Q^T A}_{R}$$

8 Terminology

8.1 Notation

$\sim$ a symbol used to represent row equivalence between two matrices.

$\blacksquare$ a symbol used to represent a nonzero number

* a symbol used to represent any number

$\left[\vec{x}\right]_B$ the coordinates of $\vec{x}$ relative to the basis B

P_B a change-of-coordinates matrix relative to basis B

$\mathbf{dim}(V)$ the dimension of the basis V.

$\underset{C\leftarrow B}{P}$ a change-of-coordinates matrix from basis B to basis C.

$\mathbf{det}A$ the determinant of the matrix A.

λ an eigenvalue corresponding to an eigenvector.

$\mathbf{Re}(z)$ the real part of the complex number z.

$\mathbf{Im}(z)$ the imaginary part of the complex number z.

$\|\vec{x}\|$ the length/norm of a vector $\vec{x}$.

$\mathbf{dist}(\vec{u}, \vec{v})$ the distance between the vectors $\vec{u}$ and $\vec{v}$.

$W^{\perp}$ the set of all vectors orthogonal to subspace W. Pronounced "W perp".

$\hat{y}$ the component of $\vec{y}$ that is a multiple of $\vec{u}$ when $\vec{y}$ is decomposed into the sum of a vector orthogonal to another vector $\vec{u}$ and a vector that is a multiple of $\vec{u}$.

$\hat{x}$ the solution to the normal equation $A^TA\vec{x} = A^T\vec{b}$.

8.2 Definitions

linear equation
a linear equation in vars $x_1, x_2, \ldots, x_n$ is an equation that can be written as

$$a_1x_1 + a_2x_2 + \ldots a_nb_n = b$$

where $a_1, a_2, \ldots, a_n$ and b are real (or complex) numbers.

system of linear equations (SoE) or linear system
one or more linear equations with the same variables.

consistent system
an SoE that has one, or infinitely many solutions

inconsistent system
an SoE that has no solution

augmented matrix
a representation of a system of equations that includes the right-hand sides of the equations.

coefficient matrix
: a representation of a system of equations that does not include the right-hand sides of the equations.

leading entry
: The first nonzero entry of a matrix.

basic vairable
: variables in an SoE that correspond to pivot columns in the REF of a matrix.

free vairable
: variables in an SoE that are not basic variables. These variables can take on any value and have the equations of the SoE hold true.

echelon form (EF)
: A matrix is in echelon form if it has the following properties:

1. All nonzero rows are above any rows of all zeroes
2. Each leading entry of a row is in a column to the right of the leading entry above it
3. all entries in a column below a leading entry are zeroes

$$\begin{bmatrix} \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

reduced echelon form (REF)
: A matrix is in reduced echelon form if it is in reduced echelon form and has the following additional properties:

1. The leading entry of each nonzero row is 1
2. Each leading 1 is the only nonzero entry in its column

$$\begin{bmatrix} 1 & * & * & 0 & * & 0 & * \\ 0 & 0 & 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

pivot position
: given a matrix in echelon or reduced form, a pivot position is a position with a leading number in a row.

pivot columns
: given a matrix in echelon or reduced form, a pivot column is a column with a pivot columns.

forward phase of G-J Elimination
: getting the matrix into echelon form (steps 1-4 of G-J elimination)

backward phase of G-J Elimination
getting the matrix into echelon form (step 5 of G-J elimination)

parametric description
a description of all the variables involved in a system with their associated values

linear combination
a description of a vector as a sum of vectors multiplied by scalars. The scalars the vectors are multiplied by are called **weights**.

span
the set of all possible linear combinations of a set of vectors (span$\{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_p}\}$ is the set of all possible linear combinations of $\{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_p}\}$)

homogenous system of linear equations
an SoE that can be written as $A\vec{x} = \vec{0}$.

trivial solution
a solution to the matrix equation $A\vec{x} = \vec{0}$ where $\vec{x} = \vec{0}$.

nontrivial solution
a solution to the matrix equation $A\vec{x} = \vec{0}$ where $\vec{x} \neq \vec{0}$.

parametric vector form
a way of representing the solution set to a matrix equation where $\vec{x}$ is a sum of vectors with its free variables as weights.

linear independence and dependence
Given a set of vectors $\{\vec{v}_1, \ldots, \vec{v}_p\}$, they are **linearly independent** if

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \ldots + x_p\vec{v}_p = \vec{0}$$

has only the trivial solution. If there is a nontrivial solution to the equation, for example, a

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_p\vec{v}_p = \vec{0}$$

where $\{c_1, \ldots, c_p\}$ are not all zero, the set is **linearly dependent**.

linear dependence relation
Given a set of vectors $\{\vec{v}_1, \ldots, \vec{v}_p\}$, a **dependence relation** among $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \ldots + c_p\vec{v}_p = \vec{0}$$

where $\{c_1, \ldots, c_p\}$ are not all zero.

matrix transformation
a function that maps a vector $\vec{x}$ to a vector $A\vec{x}$.

domain of matrix trnasformation
the set of objects being mapped. If a matrix A has m rows and n columns, the domain is $\mathbb{R}^n$.

co-domain of matrix trnasformation
the set of objects being mapped to. If a matrix A has m rows and n columns, the domain is $\mathbb{R}^m$.

image of matrix transformation
: for a particular $\vec{x} \in \mathbb{R}^n$, $T(\vec{x}) = A\vec{x}$ is the image of $\vec{x}$ under T.

range of a matrix transformation
: the set of all images under T.

onto
: a transformation is onto when its range is its co-domain. In other words, for every $\vec{b}$ in the co-domain, there exists at least one $\vec{x}$ in the domain such that $T(\vec{x}) = \vec{b}$.

one-to-one
: a transformation is one-to-one when each $\vec{b}$ in the co-domain is mapped to at most 1 $\vec{x}$ in the domain.

standard basis vectors
: the set of vectors $\{\vec{e}_1, \ldots, \vec{e}_j\}$, where the vectors have 0s in all other positions besides their j^{th} position, which is a 1.

standard matrix for a linear transformation
: the matrix

$$A = \begin{bmatrix} T(\vec{e_1}) & T(\vec{e_2}) & \ldots & T(\vec{e_n}) \end{bmatrix}$$

where $\vec{e}_1, \ldots, \vec{e}_n$ are standard basis vectors.

diagonal entries
: the entries on the diagonal that goes through the middle of a matrix.

main diagonal
: the diagonal on which the diagonal entries lie on a matrix.

diagonal matrix
: a square matrix where only its diagonal entries are nonzero.

identity matrix
: a special case of a diagonal matrix where its diagonal entries are all equal to 1 (denoted as I or I_n).

zero matrix
: a matrix whose entries are all 0 (denoted as 0 or $0_{m\times n}$).

invertible/nonsingular matrix
: a matrix is invertible when exists a matrix C such that $AC = I$ and $CA = I$. A non-invertible matrix is called a singular matrix. The inverse of tha matrix is denoted as A^{-1}, if the original matrix is A.

lower triangular matrix
: a matrix that only has nonzero numbers below its diagonal.

$$\begin{bmatrix} \blacksquare & 0 & 0 & 0 \\ * & \blacksquare & 0 & 0 \\ * & * & \blacksquare & 0 \\ * & * & * & \blacksquare \end{bmatrix}$$

unit lower triangular matrix

a special case of the lower triangular matrix that has 1s on its diagonal.

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

upper triangular matrix

a matrix that has only has nonzero numbers above its diagonal.

$$\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & \blacksquare \end{bmatrix}$$

unit upper triangular matrix

special case of the upper triangular matrix that has 1s on its diagonal.

$$\begin{bmatrix} 1 & * & * & * & * \\ 0 & 1 & * & * & * \\ 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

triangular matrix

a matrix that is either lower or upper triangular.

vector space

a vector space is a nonempty set of objects (called vectors).

subspace

subspace, H, is a vector space that contains the zero vector of the vector space it is a subspace of, and is closed under addition ($\vec{u}, \vec{v} \in H$ implies $\vec{u} + \vec{v} \in H$) and scalar multiplication ($\vec{u} \in H$ and scalar c implies $c\vec{u} \in H$).

null space (Nul$A$)

a vector space that is a subspace of $\mathbb{R}^n$ that is the solution set to $A\vec{x} = \vec{0}$ for an $m \times n$ matrix A.

column space (Col$A$)

a vector space that is a subspace of $\mathbb{R}^m$ that is the set of all linear combinations of the columns of an $m \times n$ matrix A.

row space

a vector space that is the set of all linear combinations of the rows of A. denoted by Row$A$. Row$A$ = ColA^T.

kernel

the set of all $\vec{u}$ such that $T(\vec{x}) = \vec{0}$.

basis

If H is a subspace of V, a set of vectors $B \subseteq V$ is a **basis** for H means

i. B is a linearly independent set

ii. $H = \text{span}\{B\}$

coordinates relative to a basis
given $\vec{x} = c_1\vec{b_1} + c_2\vec{b_2} + \ldots + c_n\vec{b_n}$, where $B = \{b_1, \ldots, b_n\}$: the **coordinates relative to a basis** of $\vec{x}$ relative to B are the weights $c_1, \ldots, c_n$

change-of-coordinates matrix
a square matrix such that the result of multiplying it and the coordinates of a vector relative to a basis is the vector.

isomorphism
a linear transformation that is both one-to-one and onto.

dimension of a basis
the number of vectors in a basis of a vector space V (provided there are a finite number of vectors).

change-of-coordinate-matrix
is a square matrix that converts from coordinates in one base to another.

determinant
the volume of a geometric object with edges as the rows of a matrix A.

eigenvalues and eigenvectors
eigenvalues are scalars that fulfill $A\vec{x} = \lambda\vec{x}$, and eigenvectors are the nonzero vectors $\vec{x}$ that correspond to the eigenvalue in the equation.

eigenspace corresponding to λ
the solution set to $(A - \lambda I)\vec{x} = \vec{0}$ for a matrix A (essentially $\text{Nul}A - \lambda I$).

characteristic equation
the characteristic equation of a matrix A is $\det(A - \lambda I) = 0$.

similar
two matrices A and B are similar if there exists an invertible matrix P such that $A = PBP^{-1}$.

diagonalizable
a matrix A is diagonalizable if it is similar to a diagonal matrix D (basically $A = PDP^{-1}$, where P is an invertible matrix).

dynamical system
is a system that evolves over time

trajectory
$\vec{x_k}$ when it is graphed starting with different $\vec{x_0}$ in a dynamical system.

attractor
of a dynamical system is the point where all the trajectories in a dynamical system head to

inner/dot product
inner product/dot product of two vectors $\vec{u}$ and $\vec{v}$ in $\mathbb{R}^n$ is $\vec{u}^T\vec{v} = u_1v_1 + u_2v_2 + \ldots + u_nv_n = \vec{u} \cdot \vec{v}$

length/norm
: the length of a vector, calculated with $\sqrt{\vec{x}\cdot\vec{x}}$, denoted $\|\vec{x}\|$

unit vector
: is a vector with length 1 that points in the same direction of $\vec{v}$. The calculation for it is $\frac{1}{\|\vec{v}\|}$

distance
: the distance between $\vec{u}$ and $\vec{v}$ is $\|\vec{u}-\vec{v}\|$.

orthogonality
: if two objects are orthogonal, they are perpendicular to each other (90° angle between each other). Abstractly, this means that for two vectors $\vec{u}$ and $\vec{v}$, $\vec{u}\cdot\vec{v}=0$.

orthogonal complement
: the set of all vectors orthogonal to a subspace W.

pairwise orthogonal
: the condition of being orthogonal to every other vector besides itself 3in a set.

orthogonal set
: a set of vectors in which every vector is orthogonal to every vector besides itself. Defined mathematically, a set $\{u_1,\ldots,u_p\}$ is said to be an orthogonal set if $\vec{u_i}\cdot\vec{u_j}=0$ whenever $i\neq j$.

orthogonal basis
: an orthogonal set S that does not contain the zero vector.

orthogonal projection
: an orthogonal projection of $\vec{y}$ onto $\vec{u}$ is $\hat{y}=\text{proj}_L\vec{y}=\frac{\vec{y}\cdot\vec{u}}{\vec{u}\cdot\vec{u}}\vec{u}$, where L is the subspace spanned by $\vec{u}$.

orthonormal set
: an orthogonal set of vectors that have length 1.

orthonormal basis
: is an orthonormal set that spans a subspace.

Gram-Schmidt Process
: the process of generating an orthogonal basis from a basis of a subspace.

normal equation
: the equation $A^TA\vec{x}=A^T\vec{b}$, whose solutions the least square solutions.