MATH 1553/1554
LINEAR ALGEBRA

Saptak Das

Fall 2024

Contents

1 Overview

- Heavily conceptual.
  - Careful choice of terms.
- Steep Difficulty Change.

Linear algebra is everywhere:

- Your Computer
- Data Compression
- Quantum Physics
- Population Studies (Epidemiology)
- Google Page Rank and Singular Value Decomposition

2 Systems of Linear Equations

Definition 2.1. Multidimensional Vector Spaces

- $\Re$: All Real Numbers
- $\Re^2$: All (x, y) such that x and y are real.
- $\Re^3$: All (x, y, z) such that x, y, and z are real.
- $\Re^n$: All $(x_1, x_2, \ldots, x_n)$ such that $x_1, x_2, \ldots, x_n$ are real.

Consider $S(x, y, 0)$, the x-y plane in $\Re^3$.

- **Question: Is S equal to $\Re^3$?**
- *Answer: No. $\Re^2$ has two coordinates. $\Re^3$ has three coordinates.*

2.1 Lines and Planes in $\Re^n$

Ex. 1: $x + y = 1$ in $\Re^2$
Parameterization: $x, 1 - x$

Ex. 2: $3x - y + z = 3$ in $\Re^3$
Parameterization: $(x, y, 3 - 3x + y)$

Ex. 3: $x - 2y + z - 5w = 7$ in $\Re^4$ aka "3-plane" in Re^4

2.2 Linear Equations

Examples of Linear Equations:

$$x + 2y + 5z = 3 \tag{1}$$
$$x = 5 - w \tag{2}$$
$$2x_1 - 3x_3 = \pi + x_2 - ln(n)x_4 \tag{3}$$

Examples of Nonlinear Equations:

$$\sqrt{x} \tag{4}$$
$$x^2 \tag{5}$$
$$sin(x) \tag{6}$$
$$ln(y) \tag{7}$$
$$xy = 5 \tag{8}$$

2.2.1 Solution of Linear Equation

The solution is the intersection of lines, planes, etc.

Definition 2.2. A linear equation system is considered to be **consistent** if it has at least one solution. Otherwise, the system is considered to be **inconsistent**.

Ex. 1: Consistent System. Solution is $(2, 0, 3)$**.**

$$x - y + z = 5 \tag{9}$$
$$-x + y + 2z = 4 \tag{10}$$

Ex. 2: Inconsistent Linear Equations in (x, y)

$$x + y = 0 \tag{11}$$
$$x + y = 1 \tag{12}$$

Ex. 3: Consistent Linear Equations in (x, y). Solution is at $(3, 1)$.

$$x - y = 2 \tag{13}$$
$$x + y = 4 \tag{14}$$

Ex. 4: Consistent Linear Equations in (x, y). Infinite number of solutions.

$$x + y = 0 \tag{15}$$
$$3x + 3y = 0 \tag{16}$$

2.3 Augmented Matrices

Ex. 1:

$$x - y = 10 \tag{17}$$
$$3x + 6y = 29 \tag{18}$$

Augemented Matrix Form:

$$\left(\begin{array}{cc|c} 1 & -1 & 10 \\ 3 & 6 & 29 \end{array}\right) \tag{19}$$

Solving Process using Row Reduction (Gaussian Elimination):

1. $R_2 - 3R_1 \to R_2$
2. $\frac{1}{9}R_2 \to R_2$

$$\left(\begin{array}{cc|c} 1 & -1 & 10 \\ 0 & 9 & -1 \end{array}\right) \tag{20}$$
$$\left(\begin{array}{cc|c} 1 & -1 & 10 \\ 0 & 1 & -\frac{1}{9} \end{array}\right) \tag{21}$$

Back-substitution after Gaussian Elimination:

$$x - y = 10 \tag{22}$$
$$y = -\frac{1}{9} \tag{23}$$

$\therefore x = \frac{89}{9}$ and $y = -\frac{1}{9}$

OR: Gauss-Jordan Elimination:

$$\left(\begin{array}{cc|c} 1 & 0 & \frac{89}{9} \\ 0 & 1 & -\frac{1}{9} \end{array}\right) \tag{24}$$

$\therefore x = \frac{89}{9}$ and $y = -\frac{1}{9}$

Ex. 2:

$$x + 2y + 3z = 6 \tag{25}$$
$$2x - 3y + 2z = 14 \tag{26}$$
$$3x + y - z = -2 \tag{27}$$

Augemented Matrix Form:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 2 & -3 & 2 & 14 \\ 3 & 1 & -1 & -2 \end{array}\right) \tag{28}$$

Solving Process using Row Reduction:

1. $R_2 - 2R_1 \to R_2$
2. $R_3 - 3R_1 \to R_3$
3. $R_2 \leftrightarrow R_3$
4. $-\frac{1}{5}R_2 \to R_2$
5. $R3 + 7R_2 \to R_3$
6. $\frac{1}{10}R_3 \to R_3$

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 3 & 1 & -1 & -2 \end{array}\right) \tag{29}$$
$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -7 & -4 & 2 \\ 0 & -5 & -10 & -20 \end{array}\right) \tag{30}$$
$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -5 & -10 & -20 \\ 0 & -7 & -4 & 2 \end{array}\right) \tag{31}$$
$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & -7 & -4 & 2 \end{array}\right) \tag{32}$$
$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 10 & 30 \end{array}\right) \tag{33}$$
$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 2 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right) \tag{34}$$

Back-substitution after Gaussian Elimination:

$$
\begin{aligned}
x + 2y + 3z &= 6 && (35)\\
y + 2z &= 4 && (36)\\
z &= 3 && (37)
\end{aligned}
$$

$\therefore x = 1,\ y = -2,\ z = 3$

1. $R_2 - 2R_3 \rightarrow R_2$
2. $R_1 - 3R_3 \rightarrow R_1$
3. $R_1 - 2R_2 \rightarrow R_1$

OR: Gauss-Jordan Elimination:

$$
\left(\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right) \tag{38}
$$

$$
\left(\begin{array}{ccc|c} 1 & 2 & 0 & -3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right) \tag{39}
$$

$$
\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \end{array}\right) \tag{40}
$$

$\therefore x = 1,\ y = -2,\ z = 3$

Definition 2.3. Row-Echelon Form:

1. All zero rows (if any) are at the bottom.
2. Each first non-zero entry (aka leading entry or pivot point) in a row is to the right of the first non-zero entry in the above row.
3. Below any leading entry, all entries are zero.

$$
\left(\begin{array}{ccc|c} a_1 & a_2 & a_3 & a_4 \\ 0 & a_5 & a_6 & a_7 \\ 0 & 0 & a_8 & a_9 \end{array}\right)
$$

Definition 2.4. Reduced Row-Echelon:

1. All row-echelon form requirements must be satisfied.
2. Each leading entry is 1.
3. Each pivot point is the **only** non-zero entry in its column.

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & a_1 \\ 0 & 1 & 0 & a_2 \\ 0 & 0 & 1 & a_3 \end{array}\right)$$

Ex. 3:

$$2x_1 - 3x_2 + x_3 + x_4 = 6 \tag{41}$$
$$x_1 - x_2 - x_4 = 4 \tag{42}$$
$$x_2 - x_3 = 2 \tag{43}$$

Augemented Matrix Form:

$$\left(\begin{array}{cccc|c} 2 & -3 & 1 & 1 & 6 \\ 1 & -1 & 0 & -1 & 4 \\ 0 & 1 & -1 & 0 & 2 \end{array}\right) \tag{44}$$

Solving Process using Row Reduction:

1. $R_1 \leftrightarrow R_2$
2. $R_2 - 2R_1 \to R_2$
3. $-R_2 \to R_2$
4. $R_3 - R_2 \to R_3$
5. $\frac{1}{3}R_3 \to R_3$
6. $3R_3 + R_2 \to R_2$
7. $R_3 + R_1 \to R_1$
8. $R_2 + R_1 \to R_1$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & -1 & 4 \\ 2 & -3 & 1 & 1 & 6 \\ 0 & 1 & -1 & 0 & 2 \end{array}\right) \tag{45}$$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & -1 & 4 \\ 0 & -1 & 1 & 3 & -2 \\ 0 & 1 & -1 & 0 & 2 \end{array}\right) \tag{46}$$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & -1 & 4 \\ 0 & -1 & 1 & 3 & -2 \\ 0 & 1 & -1 & 0 & 2 \end{array}\right) \tag{47}$$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & -1 & 4 \\ 0 & 1 & -1 & -3 & 2 \\ 0 & 1 & -1 & 0 & 2 \end{array}\right) \tag{48}$$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & -1 & 4 \\ 0 & 1 & -1 & -3 & 2 \\ 0 & 0 & 0 & 3 & 0 \end{array}\right) \tag{49}$$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & -1 & 4 \\ 0 & 1 & -1 & -3 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right) \tag{50}$$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & -1 & 4 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right) \tag{51}$$

$$\left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 4 \\ 0 & 1 & -1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right) \tag{52}$$

$$\tag{53}$$

2.4 Parametric Forms

Write the solution set to $x + y + z = 10$.

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 10 \end{array}\right) \tag{54}$$

The first one is a pivot. The other ones are free variables.

$$x = 10 - y - z \tag{55}$$

$$y = y \tag{56}$$

$$z = z \tag{57}$$

Geometrically, the system is a plane in $\Re^3$.

3 Matrix Equation

3.1 Linear Combinations and Span

Definition 3.1. Linear Combinations
A is a $m \times n$ matrix with columns $\vec{a_1}, \ldots, \vec{a_n}$ and $x \in \Re^n$, then the matrix-vector product $A\vec{x}$ is a linear combination of the columns of A:

$$A\vec{x} = \begin{pmatrix} | & | & | \\ \vec{a_1} & \vec{a_2} & \vec{a_3} \\ | & | & | \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1\vec{a_1} + x_2\vec{a_2} + x_3\vec{a_3} \tag{58}$$

Definition 3.2. Span
The Span of a set of vectors $\vec{v_1}, \ldots, \vec{v_p}$ in Re^n is:

$$\text{Span}\{\vec{v_1}, \ldots, \vec{v_p}\} = \{x_1\vec{v_1} + \cdots + x_2\vec{v_p} \mid x_1, \ldots, x_p \in \Re\}$$

For what vectors $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, does the equation have a solution?

$$\begin{pmatrix} 1 & 3 & 4 \\ 2 & 8 & 4 \\ 0 & 1 & -2 \end{pmatrix} \vec{x} = \vec{b} \tag{59}$$

Solution: Row Reduce $[A|b]$

$$\left(\begin{array}{ccc|c} 1 & 3 & 4 & b_1 \\ 2 & 8 & 4 & b_2 \\ 0 & 1 & -2 & b_3 \end{array}\right) \tag{60}$$

$$\left(\begin{array}{ccc|c} 1 & 3 & 4 & b_1 \\ 0 & 2 & -4 & -2b_1 + b_2 \\ 0 & 1 & -2 & b_3 \end{array}\right) \tag{61}$$

$$\left(\begin{array}{ccc|c} 1 & 3 & 4 & b_1 \\ 0 & 1 & -2 & b_3 \\ 0 & 2 & -4 & -2b_1 + b_2 \end{array}\right) \tag{62}$$

$$\left(\begin{array}{ccc|c} 1 & 3 & 4 & b_1 \\ 0 & 1 & -2 & b_3 \\ 0 & 0 & 0 & -2b_1 + b_2 - 2b_3 \end{array}\right) \tag{63}$$

$-2b_1 + b_2 - 2b_3 \neq 0$ is a pivot. Furthermore, the system will be consistent when $-2b_1 + b_2 - 2b_3 = 0$.

Q: When is $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ in the span of $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 3 \\ 8 \\ 1 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 4 \\ -2 \end{pmatrix}$?

Same solution. $\vec{b}$ is in the span of the vectors when it can be expressed as a linear combination. In other words, some real scalars x_1, x_2, x_3 can be used to create a linear combination of the three column vectors to result in $\vec{b}$.

Let A = $\begin{pmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{pmatrix}$ **and** $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$**. Is the equation** $A\vec{x} = \vec{b}$ **consistent for all possible** b_1, b_2, b_3**?**

Solution: Row Reduce A

$$\begin{pmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{pmatrix} \tag{64}$$

$$\begin{pmatrix} 1 & 3 & 4 \\ 0 & 7 & 5 \\ 0 & 0 & 0 \end{pmatrix} \tag{65}$$

$$\tag{66}$$

If A does not have a pivot in every row, then some choice of $\vec{b}$ will make a pivot in $[A|b]$ inconsistent.

3.2 Solution Sets of Linear Systems

> **Definition 3.3.** Homogeneity Linear systems of the form $A\vec{x} = \vec{0}$ are homogeneous. Linear systems of the form $A\vec{x} = \vec{b}$ are inhomogeneous.

Because homogeneous systems always have the trivial solution, $\vec{x} = \vec{0}$, the interesting question is how many solutions the system will have.

Solution Set for a Homogeneous System:

Row Reduce $[A|\vec{0}]$

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & 0 \\ 2 & -1 & 5 & 0 \\ 1 & 0 & -2 & 0 \end{array}\right) \tag{67}$$

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & 0 \\ 0 & -7 & -7 & 0 \\ 0 & -3 & -3 & 0 \end{array}\right) \tag{68}$$

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right) \tag{69}$$

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{70}$$

$$\left(\begin{array}{ccc|c} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{71}$$

Rewrite the Equations:

$$x_1 - 2x_3 = 0 \tag{72}$$

$$x_2 + x_3 = 0 \tag{73}$$

$$x_3 = x_3 \tag{74}$$

Move free variables to RHS. This results in parametric equation form.

$$x_1 = 2x_3 \tag{75}$$

$$x_2 = -x_3 \tag{76}$$

$$x_3 = x_3 \tag{77}$$

Therefore, the vector parametric form is:

$$\vec{x} = x_3 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \tag{78}$$

The solution is a line in $\Re^3$.

Inhomogeneous Case:

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 2 & -1 & 5 & 11 \\ 1 & 0 & -2 & 6 \end{array}\right) \tag{79}$$

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 0 & -7 & -7 & -7 \\ 0 & -3 & -3 & -3 \end{array}\right) \tag{80}$$

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{array}\right) \tag{81}$$

$$\left(\begin{array}{ccc|c} 1 & 0 & -2 & 6 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{82}$$

Rewrite the Equations:

$$x_1 - 2x_3 = 6 \tag{83}$$

$$x_2 + x_3 = 1 \tag{84}$$

$$x_3 = x_3 \tag{85}$$

Move free variables to RHS. This results in parametric equation form.

$$x_1 = 6 + 2x_3 \tag{86}$$

$$x_2 = 1 - x_3 \tag{87}$$

$$x_3 = x_3 \tag{88}$$

Therefore, the vector parametric form is:

$$\vec{x} = \begin{pmatrix} 6 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \tag{89}$$

The solution is a line in $\Re^3$. The particular solution is $\begin{pmatrix} 6 \\ 1 \\ 0 \end{pmatrix}$. The general solutions to $A\vec{x} = \vec{b}$ are general solutions of $A\vec{x} = \vec{0}$ plus a particular solution.

Find general solutions to both $A\vec{x} = \vec{b}$ and $A\vec{x} = \vec{0}$.

$$A = \begin{pmatrix} 1 & 3 & 0 & -4 \\ 2 & 6 & 0 & -8 \end{pmatrix} \tag{90}$$

$$\vec{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \tag{91}$$

$$\vec{b} = \begin{pmatrix} -9 \\ -18 \end{pmatrix} \tag{92}$$

Row reduce $[A|b]$.

$$\left(\begin{array}{cccc|c} 1 & 3 & 0 & -4 & -9 \\ 2 & 6 & 0 & -8 & -18 \end{array}\right) \tag{93}$$

$$\left(\begin{array}{cccc|c} 1 & 3 & 0 & -4 & -9 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \tag{94}$$

Rewrite the Equations:

$$x_1 + 3x_2 - 4x_4 = -9 \tag{95}$$

$$x_2 = x_2 \tag{96}$$

$$x_3 = x_3 \tag{97}$$

$$x_4 = x_4 \tag{98}$$

$$\tag{99}$$

Move free variables to RHS. This results in parametric equation form.

$$x_1 = -9 - 3x_2 + 4x_4 \tag{100}$$

$$x_2 = x_2 \tag{101}$$

$$x_3 = x_3 \tag{102}$$

$$x_4 = x_4 \tag{103}$$

$$\tag{104}$$

Therefore, the vector parametric form is:

$$\vec{x} = \begin{pmatrix} -9 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 4 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{105}$$

The solution is a 3D-object in $\Re^4$. The particular solution is $\begin{pmatrix} -9 \\ 0 \\ 0 \\ 0 \end{pmatrix}$. The general solutions to $A\vec{x} = \vec{b}$ are general solutions of $A\vec{x} = \vec{0}$ plus a particular solution. Therefore, the solution to $A\vec{x} = \vec{0}$ is:

$$\vec{x} = x_2 \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 4 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{106}$$

3.3 Linear Independence

Definition 3.4. An indexed set of vectors $\{\vec{v_1}, \ldots, \vec{v_p}\}$ in Re^n is said to be linearly independent if the vector equation:

$$x_1\vec{v_1} + x_2\vec{v_2} + \cdots + x_p\vec{v_p} = 0$$

has only the trivial solution. The set $\{\vec{v_1}, \ldots, \vec{v_p}\}$ is said to be linearly dependent if the vector equation

$$x_1\vec{v_1} + x_2\vec{v_2} + \cdots + x_p\vec{v_p} = 0$$

has an infinite number of solutions where some values in $x_1, \ldots, x_p$ are non-zero.

Ex. 1: Is the following linearly independent: $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

Solution:

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 3 & 1 \end{pmatrix} \tag{107}$$

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \tag{108}$$

$$\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{109}$$

$$\tag{110}$$

Rewrite the Equations:

$$x_1 - 2x_3 = 0 \tag{111}$$
$$x_2 + x_3 = 0 \tag{112}$$
$$x_3 = x_3 \tag{113}$$

Move free variables to RHS. This results in parametric vector form.

$$\vec{x} = x_3 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \tag{114}$$

Therefore, the system is linearly dependent because there are non-trivial solutions to the coefficient matrix. Specifically, one dependence relation is:

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 3 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{115}$$

Ex. 2: Is the following linearly independent: $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$

Solution:

$$\left(\begin{array}{ccc|c} 1 & 2 & 0 & 0 \\ 0 & 2 & 2 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right) \tag{116}$$
$$\begin{pmatrix} 1 & 2 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & 1 & 0 \end{pmatrix} \tag{117}$$
$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{118}$$
$$\tag{119}$$

Rewrite the Equations:

$$x_1 = 0 \tag{120}$$
$$x_2 = 0 \tag{121}$$
$$x_3 = 0 \tag{122}$$

Therefore, the system is linearly independent because there is only the trivial solution to the coefficient matrix.

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 0 \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{123}$$

Ex. 3: Is the following linearly independent: $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}$

Solution: Linearly dependent. A dependence relation is:

$$3 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{124}$$

4 Linear Transformations

4.1 Introduction

Functions can be expressed as so:

$$f:\ \Re \to \Re \tag{125}$$

A $m \times n$ matrix transformation can be expressed as:

$$T:\ \Re^n \to \Re^m,\ T(\vec{v}) = A\vec{v} \tag{126}$$

The vector $T(\vec{v})$ is the image of $\vec{v}$ under T. The set of all possible images $T(\vec{v})$ is the range.

Definition 4.1. A function T: $\Re^n \to \Re^m$ is linear if:

- $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$ for all $\vec{u}, \vec{v}$ in $\Re^n$.
- $T(c\vec{v}) = cT(\vec{v})$ for all $\vec{v}$ in $\Re^n$, and c in $\Re$.

Definition 4.2. The principle of superposition:
If T is linear, then:

$$T(c_1\vec{v_1} + \cdots + c_k\vec{v_k}) = c_1T(\vec{v_1}) + \cdots + c_kT(\vec{v_k})$$

4.2 Standard Basis Vectors

For example, in $\Re^3$,

$$\vec{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{127}$$

$$\vec{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{128}$$

$$\vec{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{129}$$

What is the linear transform $T(\vec{x}) = A\vec{x}$ to rotate counterclockwise by angle θ?

$$A = \begin{pmatrix} cos(\theta) & cos(\theta + \frac{\pi}{2}) \\ sin(\theta) & sin(\theta + \frac{\pi}{2}) \end{pmatrix} \tag{130}$$

$$= \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix} \tag{131}$$

Let $T(\vec{x} = A\vec{x}$ be the transformation which first reflects vectors in $\Re^2$ across the line $y = 0$ and then projects the resulting vector to the y-axis. Thus:

$$A = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \tag{132}$$

4.3 Onto and One-to-one Transformations

Definition 4.3. A linear transformation $T: \Re^n \rightarrow \Re^m$ is onto if for all $\vec{b} \in \Re^m$ there is a $\vec{x} \in \Re^n$ so that $T(\vec{x}) = \vec{b}$. In other words, $A\vec{x} = \vec{b}$ is always consistent. There are pivots in every row.

Definition 4.4. A linear transformation $T: \Re^n \rightarrow \Re^m$ is one-to-one if for all $\vec{b} \in \Re^m$ there is at most one (possibly no) so that $T(\vec{x}) = \vec{b}$. There are pivots in every column.

One-to-one but not onto?
Tall matrix. There is a pivot in every column, but not a pivot in every row.

Onto but not one-to-one?
Wide matrix. There is a pivot in every row, but not in every column.

5 Matrix Operations

5.1 Properties

Given that A is an $m \times n$ matrix.

The following are **facts**:

1. Associative: $(AB)C = A(BC)$
2. Left Distributive: $A(B + C) = AB + AC$
3. Right Distributive: $(A + B)C = AC + BC$
4. Identity: $I_m A = AI_n$

The following are **non-facts**:

1. Non-commutative: $AB \neq BA$
2. Non-cancellation: $AB = AC$, does not imply $B = C$
3. Zero Divisors: $AB = 0$ does not imply that either $A = 0$ or $B = 0$.

5.2 Transpose of a Matrix

A^T is the matrix whose columns are the rows of A.

Properties:

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(rA)^T = rA^T$
4. $(AB)^T = B^T A^T$

5.3 Matrix Powers

$$A^k = AA \dots A \tag{133}$$

When A is a diagonal matrix (non-zero terms only on the main diagonal), the power computation just exponentiates the main diagonal elements.

6 Inverse of a Matrix

6.1 Definition

Definition 6.1. $A \in R^{n\times n}$ is invertible if there is a $C \in R^{n\times n}$ so that:

$$AC = CA = I$$

6.2 Inverse of 2×2 Matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \tag{134}$$

6.3 Using Inverses to Solve Linear Systems

$$3x + 4y = 7 \tag{135}$$

$$5x + 6y = 7 \tag{136}$$

$$\begin{pmatrix} 3 & 4 \\ 5 & 6 \end{pmatrix}^{-1} = \begin{pmatrix} -3 & 2 \\ \frac{5}{2} & -\frac{3}{2} \end{pmatrix} \tag{137}$$

$$\begin{pmatrix} -3 & 2 \\ \frac{5}{2} & -\frac{3}{2} \end{pmatrix} \begin{pmatrix} 7 \\ 7 \end{pmatrix} = \begin{pmatrix} -7 \\ 7 \end{pmatrix} \tag{138}$$

6.4 Properties of Inverses

A and B are invertible $n \times n$ matrices.

1. $(A^{-1})^{-1} = A$
2. $(AB)^{-1} = B^{-1}A^{-1}$
3. $(A^T)^{-1} = (A^{-1})^T$

6.5 General Algorithm for Computing A^{-1}

$$\left(\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \tag{139}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \tag{140}$$

$$\tag{141}$$

Therefore, A^{-1} is $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Thus, the inverse of a matrix is the product of the elementary matrices that turned the A matrix to RREF.

6.6 Invertible Matrix Theorem

Criteria for having an inverse are below. All statements are equivalent. Thus, if one statement is true, all are true. If one statement is false, all are false. A is an $n \times n$ square matrix.

1. A is invertible.
2. A is row equivalent to I_n.
3. A has n pivotal columns.
4. $A\vec{x} = \vec{0}$ only has the trivial solution.
5. The columns of A are linearly independent.
6. The linear transformation $\vec{x} \to A\vec{x}$ is one-to-one.
7. The equation $A\vec{x} = \vec{b}$ has a solution for all $\vec{b} \in \Re^n$.
8. The columns of A span $\Re^n$.
9. The linear transformation $\vec{x} \to A\vec{x}$ is onto.
10. There is a $n \times n$ matrix C so that $CA = I_n$. (Left Inverse).
11. There is a $n \times n$ matrix D so that $AD = I_n$. (Right Inverse).
12. A^T is invertible.
13. The columns of A are a basis for $\Re^n$.
14. Col $A = \Re^n$
15. rank $A = \dim(\text{Col } A) = n$
16. Null $A = \{\vec{0}\}$
17. $(\text{Col } A)^\perp = \{\vec{0}\}$
18. $(\text{Nul } A)^\perp = \Re^n$
19. Row $A = \Re^n$
20. A has n nonzero singular values.

Theorem 6.2. If A and B are $n \times n$ matrices and $AB = I$, then $A = B^{-1}$ and $B = A^{-1}$.

Ex: Is this matrix invertible?

$$\begin{pmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{pmatrix} \tag{142}$$

$$\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{pmatrix} \tag{143}$$

Since the matrix has pivots in every column, the matrix is invertible.

Definition 6.3. A singular matrix is non-invertible. A non-singular matrix is invertible.

6.7 Partitioned Matrix

The matrix A can be represented as $\begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}$.

Ex: Find the inverse of $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$.

$$\begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \begin{pmatrix} X & Y \\ W & Z \end{pmatrix} = \begin{pmatrix} AX + BW & AY + BZ \\ 0X + CW & 0Y + CZ \end{pmatrix} = \begin{pmatrix} I_n & 0 \\ 0 & I_n \end{pmatrix} \tag{144}$$

Therefore, $Z = C^{-1}$, $W = 0$, $X = A^{-1}$, $Y = -A^{-1}BC^{-1}$.

$$\begin{pmatrix} X & Y \\ W & Z \end{pmatrix} = \begin{pmatrix} A^{-1} & -A^{-1}BC^{-1} \\ 0 & C^{-1} \end{pmatrix} \tag{145}$$

6.8 Strassen Algorithm

Matrix multiplication is $O(n^3)$. Strassen's Algorithm uses matrix partitions to get to $O(n^{2.803...})$.

7 Matrix Factorization

Examples:

$$A = LU \tag{146}$$

$$A = QR \tag{147}$$

$$A = PDP^{-1} \tag{148}$$

$$A = U\Sigma V^T \tag{149}$$

7.1 LU Decomposition

Theorem 7.1. If A is an $m \times n$ matrix that can be row reduced to echelon form without row exchanges, then $A = LU$. L is a lower triangular $m \times m$ matrix with 1's on the diagonal. U is an echelon form of A.

Definition 7.2. Types of Matrices
Upper Triangular Matrix: if $a_{i,j} = 0$ for $i > j$.
Lower Triangular Matrix: if $a_{i,j} = 0$ for $i < j$.
Diagonal: Both upper and lower, if $a_{i,j} = 0$ for $i \neq j$.

Example 1:

1. $4R_1 + R_2 \to R_2$
2. $-2R_1 + R_3 \to R_3$
3. $-5R_1 + R_3 \to R_3$

$$A = \begin{pmatrix} 4 & -3 & -1 & 5 \\ -16 & 12 & 2 & -17 \\ 8 & -6 & -12 & 22 \end{pmatrix} \tag{150}$$

$$\begin{pmatrix} 4 & -3 & -1 & 5 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & -10 & 12 \end{pmatrix} \tag{151}$$

$$U = \begin{pmatrix} 4 & -3 & -1 & 5 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & -3 \end{pmatrix} \tag{152}$$

Use the scaling factors to determine L,

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ 2 & 5 & 1 \end{pmatrix} \tag{153}$$

7.2 Why LU Decomposition works?

$$E_3 E_2 E_1 A = U \tag{154}$$

$$A = E_1^{-1} E_2^{-1} E_3^{-1} U \tag{155}$$

$$A = LU \tag{156}$$

LU decompositions do not always exist and the decompositions are not always unique.

7.3 Solving using LU Decomposition

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

Example:

$$A = LU = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \tag{157}$$

$$\vec{b} = \begin{pmatrix} 16 \\ 2 \\ -4 \\ 6 \end{pmatrix} \tag{158}$$

Step 1: Row reduce $(L|\vec{b})$.

$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 16 \\ 0 & 1 & 0 & 0 & -14 \\ 0 & 0 & 1 & 0 & 8 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right) \tag{159}$$

Step 2: Row reduce $(U|\vec{y})$.

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right) \tag{160}$$

Therefore, $\vec{x} = \begin{pmatrix} 6 \\ 10 \\ 4 \end{pmatrix}$.

8 Vector Subspaces

8.1 Definitions

Definition 8.1. A subset H of $\Re^n$ is a subspace if it is closed under scalar multiples and vector addition. That is: for any $c \in \Re$ and for $\vec{u}, \vec{v} \in H$.

1. $c\vec{u} \in H$
2. $\vec{u} + \vec{v} \in H$

Definition 8.2. Given an $m \times n$ matrix $A = [\vec{a_1} \dots \vec{a_n}]$

1. The column space of A, Col A, is the subspace of $\Re^m$ spanned by $\vec{a_1}, \dots, \vec{a_n}$.
2. The null space of A, Null A, is the subspace of $\Re^n$ spanned by the set of all vectors $\vec{x}$ that solve $A\vec{x} = \vec{0}$

9 Dimension and Rank

Definition 9.1. The dimension of a subspace H of $\Re^n$ is the number of basis vectors in any basis of H. We define $\dim\{0\} = 0$.

9.1 Choice of Basis

The standard basis of $\Re^3$ is $B_1 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$.

However, another possible choice of basis in $\Re^3$ is $B_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$.

9.2 Dimension of Null and Column Spaces

dim(Null A) is the number of free variables. dim(Col A) is the number of pivots.

10 Determinants

10.1 Definition

Definition 10.1. Definition of Determinant Suppose A is $n \times n$ and has elements a_{ij}.

1. If $n = 1$, $A = a_{11}$, and has determinant det $A = a_{11}$.
2. Inductive case: for $n > 1$,

$$\det A = a_{11}\det A_{11} - a_{12}\det A_{12} + \dots + (-1)^{1+n} a_{1n}\det A_{1n}$$

   where A_{ij} is the submatrix obtained by eliminating row i and column j of A.

Definition 10.2. Cofactors
The (i, j) cofactor of an $n \times n$ matrix A is:

$$C_{ij} = (-1)^{i+j}\det A_{ij}$$

Theorem 10.3. The determinant of a matrix A can be computed down any row or column of the matrix. For instance, down the jth column, the determinant is:

$$\det A = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}$$

10.2 Determinant of a 3×3 Matrix

$$\det \begin{pmatrix} 1 & -5 & 0 \\ 2 & 4 & -1 \\ 0 & 2 & 0 \end{pmatrix} = 1 \begin{vmatrix} 4 & -1 \\ 2 & 0 \end{vmatrix} - (-5) \begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0 \begin{vmatrix} 2 & 4 \\ 0 & 2 \end{vmatrix} = 2 \tag{161}$$

10.3 Determinants of Triangular Matrices

Theorem 10.4. If A is a triangular matrix then,

$$\det A = a_{11}a_{22}a_{33}\ldots a_{nn}$$

10.4 Properties of the Determinant

Theorem 10.5. For any square matrices A and B, we can show the following.

1. $\det A = \det A^T$
2. A is invertible iff $\det A \neq 0$.
3. $\det(AB) = (\det A)(\det B)$
4. $\det(2A) = 2^n \det A$

Theorem 10.6. Row Operations and the Determinant:
Let A be a square matrix.

1. If a multiple of a row of A is added to another row to produce B, then $\det B = \det A$.
2. If two rows are interchanged to produce B, then $\det B = -\det A$.
3. If one row of A is multiplied by a scaler k to produce B, then $\det B = k\det A$.

Example: Compute $\det A = \begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix}$.

1. $2R_1 + R_2 \to R_2$
2. $R_1 + R_3 \to R_3$

3. $R_2 \leftrightarrow R_3$

4. $-\frac{1}{5}R_3 \rightarrow R_3$

$$\begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix} \tag{162}$$

$$\begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix} \tag{163}$$

$$\begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & 1 \end{vmatrix} \tag{164}$$

$$E_4E_3E_2E_1A = B \tag{165}$$

$$-\frac{1}{5} * -1 * 1 * 1 * \det A = \det B = 3 \tag{166}$$

$$\det A = 15 \tag{167}$$

10.5 Volume from Determinants

The absolute value of the determinant of an $n \times n$ matrix represents the volume of a n dimensional parallepiped where the sides are the basis vectors of the matrix.

When a linear transformation occurs, the volume of the shape is scaled by the absolute value of the determinant.

Theorem 10.7. If $T_A : \Re^n \rightarrow \Re^n$, and S is some parallelogram in $\Re^n$, then:

$$\text{volume}(T_A(S)) = |\det(A)| \times \text{volume}(S)$$

11 Markov Chains

11.1 Introduction

1. A small town has two libraries, A and B.

2. After 1 month, among the books checked out of A,

   (a) 80% returned to A.

   (b) 20% returned to B.

3. After 1 month, among the books checked out of B,

   (a) 30% returned to A.

(b) 70% returned to B.

State after n months is:

$$X_n = 1000x_n = 1000 \times \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix}^n \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \tag{168}$$

Where:

- x_n is the nth element of the Markov Chain.
- X_n is the nth state of the system.

11.2 Steady State of a Markov Chain

Definition 11.1. Markov Chain Definitions:

1. A probability vector is a vector, $\vec{x}$, with non-negative elements that sum to 1.
2. A stochastic matrix is a square matrix, P, whose columns are probability vectors.
3. A Markov Chain is a sequence of probability vectors $\vec{x_k}$, and a stochastic matrix P, such that:
$$\vec{x_{k+1}} = P\vec{x_k},\ k = 0, 1, 2, \ldots$$
4. A steady-state vector for P is a vector $\vec{q}$ such that $P\vec{q} = \vec{q}$.

Example 1: Find a steady-state vector for the stochastic matrix.

$$\begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix} \tag{169}$$

Since:

$$P\vec{x} = \vec{x} \tag{170}$$
$$P\vec{x} - \vec{x} = \vec{0} \tag{171}$$
$$P\vec{x} - I\vec{x} = \vec{0} \tag{172}$$
$$(P - I)\vec{x} = \vec{0} \tag{173}$$

$$\text{rref}(P - I) = \begin{pmatrix} 1 & -\frac{3}{2} \\ 0 & 0 \end{pmatrix} \tag{174}$$

Determine x_2 by ensuring that $x_1 + x_2 = 1$.

$$x_2 = \frac{1}{\frac{3}{2} + 1} = \frac{2}{5} \tag{175}$$

Thus, $\vec{q} = \begin{pmatrix} \frac{3}{5} & \frac{2}{5} \end{pmatrix}$.

11.3 Convergence for a Regular Stochastic Matrix

Definition 11.2. A stochastic matrix P is regular if there is some k such that P^k only contains strictly positive entries.

Theorem 11.3. If P is a regular stochastic matrix, then P has a unique steady-state vector $\vec{q}$, and $x_k \vec{+} 1 = P\vec{x_k}$ converges to $\vec{q}$ as $k \to \infty$.

Example 2: Find a steady-state vector for the stochastic matrix.

$$\begin{pmatrix} 0.8 & 0.1 & 0.2 \\ 0.2 & 0.6 & 0.3 \\ 0.0 & 0.3 & 0.5 \end{pmatrix} \tag{176}$$

Since P is regular, a unique steady-state vector must exist.

$$P - I = \begin{pmatrix} -0.2 & 0.1 & 0.2 \\ 0.2 & -0.4 & 0.3 \\ 0.0 & 0.3 & -0.5 \end{pmatrix} \tag{177}$$

$$\text{rref}(P - I) = \begin{pmatrix} 1 & 0 & -\frac{11}{6} \\ 0 & 1 & -\frac{5}{3} \\ 0 & 0 & 0 \end{pmatrix} \tag{178}$$

$$\vec{q} = \begin{pmatrix} \frac{11}{27} \\ \frac{10}{27} \\ \frac{6}{27} \end{pmatrix} \tag{179}$$

12 Eigenvectors and Eigenvalues

12.1 Introduction

Definition 12.1. If $A \in \Re^{n\times n}$, and there is a $\vec{v} \neq \vec{0}$ in $\Re^n$, and

$$A\vec{v} = \lambda\vec{v}$$

then $\vec{v}$ is an eigenvector for A, and $\lambda \in \mathbb{C}$ is the corresponding eigenvalue.

Example 1: For $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, are the following eigenvectors?

1. $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$: Yes. $\lambda = 2$.

2. $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$: Yes. $\lambda = 0$.

3. $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$: No. The $\vec{0}$ cannot be an eigenvector.

Example 2: Confirm that $\lambda = 3$ is an eigenvalue of $A = \begin{pmatrix} 2 & -4 \\ -1 & -1 \end{pmatrix}$.

$$A\vec{v} = \lambda\vec{v} \tag{180}$$
$$A\vec{v} - \lambda I\vec{v} = \vec{0} \tag{181}$$
$$(A - \lambda I)\vec{v} = \vec{0} \tag{182}$$

In this case, $\lambda = 3$.

$$A - 3I = \begin{pmatrix} -1 & -4 \\ -1 & -4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 4 \\ 0 & 0 \end{pmatrix} \tag{183}$$

Thus, $\begin{pmatrix} -4 \\ 1 \end{pmatrix}$ is the eigenvector when $\lambda = 3$.

12.2 Eigenspace

Definition 12.2. Suppose $A \in \Re^{n\times n}$. The eigenvectors for a given λ span a subspace of $\Re^n$ called the λ-eigenspace of A. The λ-eigenspace of a matrix A is $\text{Nul}(A - \lambda I)$.

Example 1: let $A = \begin{pmatrix} 5 & -6 \\ 3 & -4 \end{pmatrix}$ and $\lambda = 2$.

$$A - 2I = \begin{pmatrix} 3 & -6 \\ 3 & -6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} \tag{184}$$

Thus, a basis for the eigenspace is $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Example 2: let $A = \begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix}$ and $\lambda = 2$.

$$A - 2I = \begin{pmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -\frac{1}{2} & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{185}$$

$$\vec{x} = x_1 \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \tag{186}$$

Therefore, the basis for the eigenspace is $\left\{ \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\}$.

12.3 IMT Theorems

Theorem 12.3. IMT (Continued.)

1. The diagonal elements of a triangular matrix are its eigenvalues.
2. A is invertible $\Leftrightarrow$ 0 is not an eigenvalue of A.
3. If $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ are eigenvectors that correspond to distinct eigenvalues, then $\vec{v_1}, \vec{v_2}, \ldots, \vec{v_k}$ are linearly independent.

12.4 Characteristic Polynomial

The quantity $\det(A - \lambda I)$ is the characteristic polynomial of A. The quantity $\det(A - \lambda I) = 0$ is the characteristic equation of A. The roots of the characteristic polynomial are the eigenvalues of A.

Example 1:

$$p(\lambda) = \det \begin{pmatrix} 5-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} = \lambda^2 - 6\lambda + 1 \tag{187}$$

$$\lambda = \frac{6 \pm \sqrt{32}}{2} = 3 \pm 2\sqrt{2} \tag{188}$$

General Case:

$$\det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} = \lambda^2 - (a+d)\lambda + ad - bc \tag{189}$$

Therefore, $p(\lambda) = \lambda^2 - \text{tr } A + \det A$.

12.5 Multiplicities

Definition 12.4. Algebraic Multiplicity
The algebraic multiplicity of an eigenvalue is its multiplicity as a root of the characteristic polynomial.

Definition 12.5. Geometric Multiplicity
The geometric multiplicity of an eigenvalue λ is the dimension of $\text{Null}(A - \lambda I)$.

Example of a 4×4 matrix with $\lambda = 3$ as the only eigenvalue, but the geometric multiplicity of $\lambda = 3$ is one.

$$A = \begin{pmatrix} 3 & 1 & 2 & 2 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix} \tag{190}$$

12.6 Similar Matrices

Definition 12.6. Two $n \times n$ matrices A and B are similar if there is a matrix P so that $A = PBP^{-1}$.

Theorem 12.7. If A and B are similar, then they have the same characteristic polynomial.

12.7 Applications of Eigenvalues and Eigenvectors

Consider the Markov Chain:

$$\vec{x_{k+1}} = P\vec{x_0} = \begin{pmatrix} 0.6 & 0.4 \\ 0.4 & 0.6 \end{pmatrix} \tag{191}$$

13 Diagonalizability

13.1 Powers of Diagonal Matrices

$$A = \begin{pmatrix} 3 & 0 \\ 0 & 0.5 \end{pmatrix} \tag{192}$$

$$A^2 = \begin{pmatrix} 3^2 & 0 \\ 0 & 0.5^2 \end{pmatrix} \tag{193}$$

$$A^k = \begin{pmatrix} 3^k & 0 \\ 0 & 0.5^k \end{pmatrix} \tag{194}$$

13.2 Diagonalization

Definition 13.1. Suppose $A \in \Re^{n \times n}$. We say that A is diagonalizable if it is similar to a diagonal matrix, D. The columns of P are the eigenvectors of A. D is a diagonal matrix with the corresponding eigenvalues along the main diagonal. That is, we can write:

$$A = PDP^{-1}$$

Suppose:

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \tag{195}$$

$$P = \begin{pmatrix} | & | & | \\ \vec{v_1} & \vec{v_2} & \vec{v_3} \\ | & | & | \end{pmatrix} \tag{196}$$

Example 1: Diagonalize if possible.

$$\begin{pmatrix} 2 & 6 \\ 0 & -1 \end{pmatrix} \tag{197}$$

$\lambda_1 = 2$, $\lambda_2 = -1$ because upper triangular. The basis for 2-eigenspace is $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$. The basis for -1-eigenspace is $\left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$.

$$\begin{pmatrix} 2 & 6 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix}^{-1} \tag{198}$$

Example 2: Diagonalize if possible.

$$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \tag{199}$$

$\lambda_1 = 3$, $\lambda_2 = 3$ because upper triangular. The basis for 3-eigenspace is $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$.

$$\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \neq \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}^{-1} \tag{200}$$

Not diagonalizable, since the rank $\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \neq 2$. P^{-1} does not exist due to IMT.

Example 3: Diagonalize if possible.

$$\begin{pmatrix} 7 & 4 & 16 \\ 2 & 5 & 8 \\ -2 & -2 & -5 \end{pmatrix} \tag{201}$$

$\lambda_1 = 3$, $\lambda_2 = 1$ because upper triangular.

$$A - I = \begin{pmatrix} 6 & 4 & 16 \\ 2 & 4 & 8 \\ -2 & -2 & -6 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{202}$$

The basis for 1-eigenspace is $\left\{ \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} \right\}$.

$$A - 3I = \begin{pmatrix} 4 & 4 & 16 \\ 2 & 2 & 8 \\ -2 & -2 & -8 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{203}$$

The basis for 3-eigenspace is $\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} -4 \\ 0 \\ 1 \end{pmatrix} \right\}$.

$$\begin{pmatrix} 7 & 4 & 16 \\ 2 & 5 & 8 \\ -2 & -2 & -5 \end{pmatrix} = \begin{pmatrix} -2 & -1 & -4 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 3 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} -2 & -1 & -4 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1} \tag{204}$$

13.3 Diagonalization Theorem

Theorem 13.2. An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

Example 3: Diagonalize if possible. $\lambda_1 = 1, \lambda_2 = -2$

$$\begin{pmatrix} 2 & 4 & 2 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{pmatrix} \tag{205}$$

$$A - I = \begin{pmatrix} 1 & 4 & 3 \\ -4 & -7 & -3 \\ 3 & 3 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{206}$$

The basis for 1-eigenspace is $\left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}$.

$$A + 2I = \begin{pmatrix} 4 & 4 & 3 \\ -4 & -4 & -3 \\ 3 & 3 & 3 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{207}$$

The basis for -2-eigenspace is $\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\}$.

$$\begin{pmatrix} 2 & 4 & 2 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{pmatrix} \neq \begin{pmatrix} 1 & -1 & * \\ -1 & 1 & * \\ 1 & 0 & * \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & * \end{pmatrix} \begin{pmatrix} 1 & -1 & * \\ -1 & 1 & * \\ 1 & 0 & * \end{pmatrix}^{-1} \tag{208}$$

Not diagonalizable, since the rank $\begin{pmatrix} 1 & -1 & * \\ -1 & 1 & * \\ 1 & 0 & * \end{pmatrix} \neq 3$ (algebraic multiplicity for one eigenvalue is greater than the geometric multiplicity). P^{-1} does not exist due to IMT.

13.4 Basis of Eigenvectors

Example 1: Express the vector $\vec{x_0} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$ as a linear combination of the vectors $\vec{v_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\vec{v_2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and find the coordinates of $\vec{x_0}$ in the basis $B = \{\vec{v_1}, \vec{v_2}\}$.

$$[\vec{x_0}]_B = \begin{pmatrix} \frac{9}{2} \\ -\frac{1}{2} \end{pmatrix} \tag{209}$$

Let $P = [\vec{v_1}, \vec{v_2}]$ and $D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, and find $[A^k \vec{x_0}]_B$, where $A = PDP^{-1}$, for $k = 1, 2, \ldots$.

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{210}$$

$$[A^k \vec{x_0}]_B = \begin{pmatrix} \frac{9}{2} \\ (-1)^{k+1} \frac{1}{2} \end{pmatrix} \tag{211}$$

Example 2: Express the vector $\vec{x_0} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$ as a linear combination of the vectors $\vec{v_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\vec{v_2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and find the coordinates of $\vec{x_0}$ in the basis $B = \{\vec{v_1}, \vec{v_2}\}$.

$$[\vec{x_0}]_B = \begin{pmatrix} \frac{9}{2} \\ -\frac{1}{2} \end{pmatrix} \tag{212}$$

Let $P = [\vec{v_1}, \vec{v_2}]$ and $D = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$, and find $[A^k \vec{x_0}]_B$, where $A = PDP^{-1}$, for $k = 1, 2, \ldots$.

$$A = \begin{pmatrix} 0.25 & 0.75 \\ 0.75 & 0.25 \end{pmatrix} \tag{213}$$

$$[A^k \vec{x_0}]_B = \begin{pmatrix} \frac{9}{2} \\ (-\frac{1}{2})^{k+1} \end{pmatrix} \tag{214}$$

Example 3: Express the vector $\vec{x_0} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}$ as a linear combination of the vectors $\vec{v_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\vec{v_2} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and find the coordinates of $\vec{x_0}$ in the basis $B = \{\vec{v_1}, \vec{v_2}\}$.

$$[\vec{x_0}]_B = \begin{pmatrix} \frac{9}{2} \\ -\frac{1}{2} \end{pmatrix} \tag{215}$$

Let $P = [\vec{v_1}, \vec{v_2}]$ and $D = \begin{pmatrix} 2 & 0 \\ 0 & \frac{3}{2} \end{pmatrix}$, and find $[A^k \vec{x_0}]_B$, where $A = PDP^{-1}$, for $k = 1, 2, \ldots$.

$$A = \begin{pmatrix} 1.75 & 0.25 \\ 0.25 & 1.75 \end{pmatrix} \tag{216}$$

$$[A^k \vec{x_0}]_B = \begin{pmatrix} \frac{9}{2} \cdot 2^k \\ -\frac{1}{2}(\frac{3}{2})^k \end{pmatrix} \tag{217}$$

14 Complex Eigenvalues

Theorem 14.1. Complex Eigenvalues

1. If $\lambda \in \mathbb{C}$ is a root of a real polynomial $p(x)$, then the conjugate λ is also a root of $p(x)$.
2. If λ is an eigenvalue of real matrix A with eigenvector $\vec{v}$, then $\bar{\lambda}$ is an eigenvalue of A with eigenvector $\vec{v}$.

15 Rotation-Dilation Matrix

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{218}$$

$$= \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix} \tag{219}$$

The eigenvalues of A are $a \pm bi$.

16 Inner Product and Orthogonality

16.1 Dot Products and Orthogonality

Definition 16.1. Dot Product

$$\vec{u} \cdot \vec{v} = \vec{u}^T \vec{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Definition 16.2. Orthogonality
Two vectors $\vec{u}$ and $\vec{w}$ are orthogonal if $\vec{u} \cdot \vec{v} = 0$. This is equivalent to:

$$||\vec{u} + \vec{w}||^2 = ||\vec{u}||^2 + ||\vec{w}||^2$$

16.2 Orthogonal Compliments and Sets

Definition 16.3. Orthogonal Compliments
Let W be a subspace of $\Re^n$. Vector $\vec{z} \in \Re^n$ is orthogonal to W if $\vec{z}$ is orthogonal to every vector in W.
The set of all vectors orthogonal to W is a subspace, the orthogonal complement of W or $W^\perp$.

$$W^\perp = \{\vec{x} \in \Re^n : \vec{z} \cdot \vec{w} = 0 \text{ for all } \vec{w} \in W\}$$

Definition 16.4. Row Space
Row A is the space spanned by the rows of matrix A.

Definition 16.5. A set of vectors $\{\vec{u_1}, \ldots, \vec{u_p}\}$ are an orthogonal set of vectors if for each $j \neq k, \vec{u_j} \perp \vec{v_k}$.

16.3 Projections

Definition 16.6. Projections

$$\text{proj}_{\vec{u}} \vec{v} = \frac{\vec{v} \cdot \vec{u}}{\vec{u} \cdot \vec{u}} \vec{u}$$

16.4 Inverse of Orthonormal Matrix

Theorem 16.7. Inverse of Orthonormal Matrix
The inverse of an orthonormal matrix A is $A^{-1} = A^T$. In other words, $AA^T = I$.

16.5 Orthogonal Decomposition Theorem

Theorem 16.8. Let W be a subspace of $\Re^n$. Then, each vector $\vec{y} \in \Re^n$ has the unique decomposition:

$$\vec{y} = \hat{y} + w^{\perp},\ \hat{y} \in W,\ w^{\perp} \in W^{\perp}$$

And, if $\vec{u_1}, \dots, \vec{u_p}$ is any orthogonal basis for W.

$$\hat{y} = \frac{\vec{y}\vec{u_1}}{\vec{u_1}\vec{u_1}}\vec{u_1} + \dots + \frac{\vec{y}\vec{u_p}}{\vec{u_p}\vec{u_p}}\vec{u_p}$$

We say that $\hat{y}$ is the orthogonal projection of $\vec{y}$ onto W.

16.6 Best Approximation Theorem

Theorem 16.9. Let W be a subspace of $\Re^n$, $\vec{y} \in \Re^n$, and $\hat{y}$ is the orthogonal projection of $\vec{y}$ onto W. Then for any $\vec{w} \neq \hat{y} \in W$, we have:

$$||\vec{y} - \hat{y}|| < ||\vec{y} - \vec{w}||$$

That is, $\hat{y}$ is the unique vector in W that is closest to $\vec{y}$.

17 Gram-Schmidt Process

Theorem 17.1. The Gram-Schmidt Process
Given a basis $\{\mathbf{x_1}, \dots, \mathbf{x_p}\}$ for a nonzero subspace W of $\Re^n$, define:

$$\mathbf{v_1} = \mathbf{x_1}$$

$$\mathbf{v_2} = \mathbf{x_2} - \frac{\mathbf{x_2} \cdot \mathbf{v_1}}{\mathbf{v_1} \cdot \mathbf{v_1}}\mathbf{v_1}$$

$$\mathbf{v_3} = \mathbf{x_3} - \frac{\mathbf{x_3} \cdot \mathbf{v_1}}{\mathbf{v_1} \cdot \mathbf{v_1}}\mathbf{v_1} - \frac{\mathbf{x_3} \cdot \mathbf{v_2}}{\mathbf{v_2} \cdot \mathbf{v_2}}\mathbf{v_2}$$

$$\vdots$$

$$\mathbf{v_p} = \mathbf{x_p} - \frac{\mathbf{x_p} \cdot \mathbf{v_1}}{\mathbf{v_1} \cdot \mathbf{v_1}}\mathbf{v_1} - \frac{\mathbf{x_p} \cdot \mathbf{v_2}}{\mathbf{v_2} \cdot \mathbf{v_2}}\mathbf{v_2} - \dots - \frac{\mathbf{x_p} \cdot \mathbf{v_{p-1}}}{\mathbf{v_{p-1}} \cdot \mathbf{v_{p-1}}}\mathbf{v_{p-1}}$$

Then $\{\mathbf{v_1}, \dots, \mathbf{v_p}\}$ is an orthogonal basis for W. In addition,

$$\text{Span}\{\mathbf{v_1}, \dots, \mathbf{v_p}\} = \text{Span}\{\mathbf{x_1}, \dots, \mathbf{x_p}\} \quad \text{for } 1 \leq k \leq p$$

Definition 17.2. A set of vectors from an orthonormal basis if the vectors are mutually orthogonal and have unit length.

17.1 QR Factorization

Theorem 17.3. QR Factorization
Any $m \times n$ matrix A with linearly independent columns has the QR factorization:

$$A = QR$$

where:

1. Q is $m \times n$, its columns are an orthonormal basis for Col A.
2. R is $n \times n$, upper triangular, with positive entries on its diagonal, and the length of the jth column of R is equal to the length of the jth column of A.

Example 1:

$$\vec{v_1} = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \ \vec{v_2} = \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} \tag{220}$$

Steps:

1. Determine Q matrix.
   (a) Do Gram-Schmidt to the columns of A. In this case, the columns of A are already orthogonal.
   (b) Normalize the basis.
2. Calculate $R = Q^T A$.

$$\{\vec{w_1}, \vec{w_2}\} = \left\{ \begin{pmatrix} \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{-2}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \\ \frac{1}{\sqrt{14}} \end{pmatrix} \right\} \tag{221}$$

$$Q = \begin{pmatrix} \frac{3}{\sqrt{13}} & \frac{-2}{\sqrt{14}} \\ \frac{2}{\sqrt{13}} & \frac{3}{\sqrt{14}} \\ 0 & \frac{1}{\sqrt{14}} \end{pmatrix} \tag{222}$$

$$R = \begin{pmatrix} \frac{3}{\sqrt{13}} & \frac{2}{\sqrt{13}} & 0 \\ \frac{-2}{\sqrt{14}} & \frac{3}{\sqrt{14}} & \frac{1}{\sqrt{14}} \end{pmatrix} \begin{pmatrix} 3 & -2 \\ 2 & 3 \\ 0 & 1 \end{pmatrix} \tag{223}$$

$$= \begin{pmatrix} \sqrt{13} & 0 \\ 0 & \sqrt{14} \end{pmatrix} \tag{224}$$

18 Least-Squares Problems

Definition 18.1. If A is $m \times n$ and $\vec{b}$ is in $\Re^m$, a least-squares solution of $A\vec{x} = \vec{b}$ is an $\hat{x}$ in $\Re^n$ such that:

$$||\vec{b} - A\hat{x}|| \leq ||\vec{b} - A\vec{x}||$$

for all $\vec{x}$ in $\Re^n$.

Example 1: Find a least-squares solution of the inconsistent system $A\vec{x} = \vec{b}$ for:

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} \tag{225}$$

1. Compute $A^T A$ and $A^T \vec{b}$
2. Solve $A^T A\hat{x} = A^T \vec{b}$

$$A^T A = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix} \tag{226}$$

$$A^T \vec{b} = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} = \begin{pmatrix} 19 \\ 11 \end{pmatrix} \tag{227}$$

$$[A^T A | A^T \vec{b}] = \left(\begin{array}{cc|c} 17 & 1 & 19 \\ 1 & 5 & 11 \end{array}\right) \tag{228}$$

$$\Rightarrow \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right) \tag{229}$$

$$\hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \tag{230}$$

$\hat{x}$ provides the weights to multiply by the columns of A in order to find the projection of $\vec{b}$.

Example 2: Fit a line $y = \alpha x + \beta$ that best fits the data (0,0.5), (1,1), (2,2.5), (3,3).

1. Plug in data into the model.
2. Compute $A^T A$ and $A^T \vec{b}$
3. Solve $A^T A\hat{x} = A^T \vec{b}$

$$0.5 = \alpha \cdot 0 + \beta \tag{231}$$
$$1 = \alpha \cdot 1 + \beta \tag{232}$$
$$2.5 = \alpha \cdot 2 + \beta \tag{233}$$
$$3 = \alpha \cdot 3 + \beta \tag{234}$$

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 0.5 \\ 1 \\ 2.5 \\ 3 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{235}$$

$$A^T A = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 14 & 6 \\ 6 & 4 \end{pmatrix} \tag{236}$$
$$A^T \vec{b} = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0.5 \\ 1 \\ 2.5 \\ 3 \end{pmatrix} = \begin{pmatrix} 15 \\ 7 \end{pmatrix} \tag{237}$$

$$A^T A \hat{x} = A^T \vec{b} \tag{238}$$
$$\hat{x} = (A^T A)^{-1} A^T \vec{b} \tag{239}$$
$$= \frac{1}{20} \begin{pmatrix} 4 & -6 \\ -6 & 14 \end{pmatrix} \begin{pmatrix} 15 \\ 7 \end{pmatrix} \tag{240}$$
$$= \begin{pmatrix} 0.9 \\ 0.4 \end{pmatrix} \tag{241}$$

The Sum of Squared Errors (SSE) $||\vec{b} - \hat{b}||^2 = \sum_{i=1}^{k} (\vec{b}_i - \hat{b}_i)^2$ is minimized. $\frac{1}{k}||\vec{b} - \hat{b}||^2 = \frac{1}{k}\sum_{i=1}^{k} (\vec{b}_i - \hat{b}_i)^2$ is equivalent to the Mean Squared Error (MSE) of the line.

18.1 Proof of Least Squares

$$\vec{b} - \hat{b} \in (\text{Col } A)^{\perp} = \text{Nul } A^T \tag{242}$$
$$A^T(\vec{b} - \hat{b}) = \vec{0} \tag{243}$$
$$A^T \vec{b} - A^T \hat{b} = \vec{0} \tag{244}$$
$$A^T \vec{b} - A^T A \vec{x} = \vec{0} \tag{245}$$
$$A^T A \vec{x} = A^T \vec{b} \tag{246}$$

18.2 Using Least Squares for Complex Models

Example 1: Fit a line $y = \alpha x^2 + \beta x + \gamma$ that best fits the data (0,1), (1,1), (2,3).

1. Plug in data into the model.
2. Compute $A^T A$ and $A^T \vec{b}$
3. Solve $A^T A \hat{x} = A^T \vec{b}$

$$\begin{align} 1 &= \alpha \cdot 0 + \beta \cdot 0 + \gamma \tag{247} \\ 1 &= \alpha \cdot 1 + \beta \cdot 1 + \gamma \tag{248} \\ 3 &= \alpha \cdot 4 + \beta \cdot 2 + \gamma \tag{249} \end{align}$$

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \tag{250}$$

$$\begin{align} A^T A &= \begin{pmatrix} 0 & 1 & 4 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 17 & 9 & 5 \\ 9 & 5 & 3 \\ 5 & 3 & 3 \end{pmatrix} \tag{251} \\ A^T \vec{b} &= \begin{pmatrix} 0 & 1 & 4 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 13 \\ 7 \\ 5 \end{pmatrix} \tag{252} \end{align}$$

$$\begin{align} A^T A \hat{x} &= A^T \vec{b} \tag{253} \\ \hat{x} &= (A^T A)^{-1} A^T \vec{b} \tag{254} \\ &= \begin{pmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ -\frac{3}{2} & 2 & -\frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 13 \\ 7 \\ 5 \end{pmatrix} \tag{255} \\ &= \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \tag{256} \end{align}$$

18.3 Unique Solutions for Least Squares

Theorem 18.2. Let A be any $m \times n$ matrix. These statements are equivalent.

1. The equation $A\vec{x} = \vec{b}$ has a unique least-squares solution for each $\vec{b} \in \Re^m$.
2. The columns of A are linearly independent.
3. The matrix $A^T A$ is invertible.

And, if these statements hold, the least square solution is:

$$\hat{x} = (A^T A)^{-1} A^T \vec{b}$$

18.4 Least Squares and QR Decomposition

Theorem 18.3. Least Squares and QR Decomposition
Let $m \times n$ matrix A have a QR decomposition. Then for each $\vec{b} \in \Re^m$ the equation $A\vec{x} = \vec{b}$ has a unique least squares solution

$$R\hat{x} = Q^T \vec{b}$$

(Remember, R is upper triangular, so the equation above is solved by back-substitution.)

19 Finite State Markov Chains

19.1 Steady-State Vector and Google Page Rank

The steady state vector can be found by applying $P\vec{q} = \vec{q}$, which means that the row is reduced $P - I$. By definition, the steady state vector $\vec{q}$ will have an eigenvalue of 1.
If P is regular, then $P^n \vec{e_i} = \vec{q}$. Therefore:

$$P = \begin{pmatrix} \vec{q} & \vec{q} & \dots & \vec{q} \end{pmatrix} \tag{257}$$

Theorem 19.1. Google Page Rank
If P is a regular $m \times m$ transition matrix with $m \geq 2$, then the following statements are all true.

1. There is a stochastic matrix Π such that:

$$\lim_{n\to\infty} P^n = \Pi$$

2. Each column of Π is the same probability vector $\vec{q}$.
3. For any initial probability vector $\vec{x_0}$,

$$\lim_{n\to\infty} P^n \vec{x_0} = \vec{q}$$

4. P has a unique eigenvector, $\vec{q}$, which has eigenvalue $\lambda = 1$.
5. The eigenvalues of P satisfy $|\lambda| \leq 1$.

Example 1:

1. Generate the Google Matrix.
2. Adjust matrix to fix columns that do not sum to 1 (aka have no links).
3. Use damping factor to adjust for user behavior.
4. Solve $G\vec{q} = \vec{q}$.

$$P = \begin{pmatrix} 0 & 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 & 0 \end{pmatrix} \tag{258}$$

$$P^* = \begin{pmatrix} 0 & 0.5 & 0.5 & 0 & 0.2 \\ 0.5 & 0 & 0.5 & 0.5 & 0.2 \\ 0 & 0.5 & 0 & 0 & 0.2 \\ 0.5 & 0 & 0 & 0 & 0.2 \\ 0 & 0 & 0 & 0.5 & 0.2 \end{pmatrix} \tag{259}$$

$$K = \begin{pmatrix} 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \end{pmatrix} \tag{260}$$

$$G = 0.85P^* + 0.15K \tag{261}$$

$$\vec{q} = \begin{pmatrix} 0.251 \\ 0.298 \\ 0.1766 \\ 0.156 \\ 0.116 \end{pmatrix} \tag{262}$$

Therefore, the page ranking is B, A, C, D, E.

Example 2:

1. Generate the Google Matrix.
2. Adjust matrix to fix columns that do not sum to 1 (aka no links).
3. Use damping factor to adjust for user behavior.
4. Solve $G\vec{q} = \vec{q}$.

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \end{pmatrix} \tag{263}$$

$$P^* = \begin{pmatrix} 0 & 0 & 1 & 0 & \frac{1}{5} \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} & \frac{1}{5} \\ \frac{1}{3} & 0 & 0 & \frac{1}{2} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 & \frac{1}{5} \\ 0 & \frac{1}{2} & 0 & 0 & \frac{1}{5} \end{pmatrix} \tag{264}$$

$$K = \begin{pmatrix} 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \\ 0.2 & 0.2 & 0.2 & 0.2 & 0.2 \end{pmatrix} \tag{265}$$

$$G = 0.85P^* + 0.15K \tag{266}$$

$$\vec{q} = \begin{pmatrix} 0.231 \\ 0.208 \\ 0.208 \\ 0.208 \\ 0.142 \end{pmatrix} \tag{267}$$

Therefore, the page ranking is A, (B, C, D), E.

20 Symmetric Matrices

20.1 Definition

Definition 20.1. Definition of Symmetric Matrices
Matrix A is symmetric if $A^T = A$.

20.2 Symmetry of AA^T

AA^T is also a symmetric matrix:

$$AA^T = \begin{pmatrix} — & a_1 & — \\ — & a_2 & — \\ & \vdots & \\ — & a_n & — \end{pmatrix} \begin{pmatrix} | & | & & | \\ a_1 & a_2 & \dots & a_n \\ | & | & & | \end{pmatrix} \tag{268}$$

$$= \begin{pmatrix} a_1 \cdot a_1 & \dots & a_1 \cdot a_n \\ \vdots & \ddots & \vdots \\ a_n \cdot a_1 & \dots & a_n \cdot a_n \end{pmatrix} \tag{269}$$

20.3 Eigenspaces of Symmetric Matrices

Theorem 20.2. A is a symmetric matrix, with eigenvectors $\vec{v_1}$ and $\vec{v_2}$ corresponding to two distinct eigenvalues. Then, $\vec{v_1}$ and $\vec{v_2}$ are orthogonal. More generally, eigenspaces associated with distinct eigenvalues are orthogonal subspaces.

20.4 Spectral Theorem

Theorem 20.3. An $n \times n$ symmetric matrix A has the following properties:

1. All eigenvalues of A are real.
2. The dimension of each eigenspace is full, that its dimension is equal to its algebraic multiplicity.
3. The eigenspace are mutually orthogonal.
4. A can be diagonalized: $A = PDP^{-1} = PDP^T$, where D is diagonal and P is an orthogonal matrix.

Definition 20.4. Spectral Decomposition
Suppose A can be diagonalized as $A = PDP^T$, then A has the decomposition

$$A = \lambda_1 U_1 U_1^T + \lambda_2 U_2 U_2^T + \cdots + \lambda_n U_n U_n^T = \sum_{k=1}^{n} \lambda_k U_k U_k^T$$

The following is true about the spectral decomposition:

1. $U_n U_n^T$ will always be rank 1.
2. $U_1 U_1^T \vec{x} = \mathrm{proj}_{U_1}(\vec{x})$
3. After ordering the terms largest to smallest by their corresponding eigenvalue, one can determine a rank n approximation of A by taking the first n terms.

Example:

$$\begin{aligned}
A &= \begin{pmatrix} 2 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 2 \end{pmatrix} && (270) \\
&= \lambda_1 U_1 U_1^T + \lambda_2 U_2 U_2^T + \lambda_3 U_3 U_3^T && (271) \\
&= 4 \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} + 1 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} && (272) \\
&= 4 \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} && (273)
\end{aligned}$$

21 Quadratic Forms

21.1 Definition

Definition 21.1. A quadratic form is a function $Q : \Re^n \to \Re$, given by

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

21.2 Converting between Quadratic Forms

Example 1:

$$Q_A(x_1, x_2) = \vec{x}^T A \vec{x} = \begin{pmatrix} x_1 x_2 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{274}$$

$$= 4x_1^2 + 3x_2^2 \tag{275}$$

$$Q_B(x_1, x_2) = \vec{x}^T A \vec{x} = \begin{pmatrix} x_1 x_2 \end{pmatrix} \begin{pmatrix} 4 & 1 \\ 1 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{276}$$

$$= 4x_1^2 + 2x_1 x_2 - 3x_2^2 \tag{277}$$

Example 2: Write Q in the form $\vec{x}^T A \vec{x}$.

$$Q_A(x_1, x_2) = \vec{x}^T A \vec{x} = 5x_1^2 - x_2^2 + 3x_3^2 + 6x_1 X_3 - 12x_2 x_3 \tag{278}$$

$$= \begin{pmatrix} x_1 x_2 x_3 \end{pmatrix} \begin{pmatrix} 5 & 0 & 3 \\ 0 & -1 & -6 \\ 3 & -6 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \tag{279}$$

$$\tag{280}$$

21.3 Change of Variables

Procedure to find $\vec{y} = P^T \vec{x}$:

$$A = PDP^T \tag{281}$$

$$Q_A(\vec{x}) = \vec{x}^T A \vec{x} \tag{282}$$

$$= \vec{x}^T PDP^T \vec{x} \tag{283}$$

$$= (P^T \vec{x})^T DP^T \vec{x} \tag{284}$$

$$= \vec{y}^T D \vec{y} \tag{285}$$

$$= Q_D(\vec{y}) \tag{286}$$

In order to remove cross-terms from the quadratic form, one can define $Q_A(\vec{x}) = Q_D(\vec{y}) = \vec{y}^T D \vec{y}$. P^T is a rotation matrix that transforms the quadratic form such that the cross-terms when written as a function of $\vec{x}$ are removed when written as a function of $\vec{y} = P^T \vec{x}$.

Example 1:

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} = PDP^T \tag{287}$$

$$P = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \\ -1 & 2 \end{pmatrix} \tag{288}$$

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 7 \end{pmatrix} \tag{289}$$

$$Q_A(x_1, x_2) = 3x_1^2 + 4x_1x_2 + 6x_2^2 \tag{290}$$

$$Q_D(y_1, y_2) = 2y_1^2 + 7y_2^2 \tag{291}$$

$$\vec{y} = P^T \vec{x} \tag{292}$$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{\sqrt{5}}x_1 - \frac{1}{\sqrt{5}}x_2 \\ \frac{1}{\sqrt{5}}x_1 + \frac{2}{\sqrt{5}}x_2 \end{pmatrix} \tag{293}$$

Theorem 21.2. Principal Axes Theorem
If A is a symmetric matrix then there exists an orthogonal change of variable $\vec{x} = P\vec{y}$ that transforms $\vec{x}^T A\vec{x}$ to $\vec{y}^T D\vec{y}$ with no cross-product terms.

21.4 Classifying Quadratic Forms

Definition 21.3. A quadratic form Q is:

1. positive definite if $Q(\vec{x}) > 0$ for all $\vec{x} \neq \vec{0}$.
2. negative definite if $Q(\vec{x}) < 0$ for all $\vec{x} \neq \vec{0}$.
3. positive semidefinite if $Q(\vec{x}) \geq 0$ for all $\vec{x}$.
4. negative semidefinite if $Q(\vec{x}) \leq 0$ for all $\vec{x}$.
5. indefinite if $Q(\vec{x}) \in \Re$ for all $\vec{x}$.

21.5 Quadratic Forms and Eigenvalues

Theorem 21.4. If A is a symmetric matrix with eigenvalues λ_i, then $Q = \vec{x}^T A\vec{x}$ is:

1. positive definite iff $\lambda_i > 0$.
2. negative definite iff $\lambda_i < 0$.
3. positive semidefinite iff $\lambda_i \geq 0$.
4. negative semidefinite iff $\lambda_i \leq 0$.
5. indefinite iff $\lambda_i > 0$ and some other $\lambda_i < 0$.

22 Constrained Optimization

22.1 Quadratic Forms

Example 1: Find the largest and smallest of $Q(\vec{x}) = 9x_1^2 + 4x_2^2 + 3x_3^2$ on the surface of an unit sphere.

$$\text{Max: } Q(1,0,0) = 9 \tag{294}$$
$$\text{Min: } Q(0,0,1) = 3 \tag{295}$$

If $Q(\vec{x}) = ax_1^2 + bx_2^2 + cx_3^2$ and $||v|| = 1$, then $Q(\vec{v}$ is a weighted average of a, b, c.

22.2 Using Eigenvalues for Constrained Optimization

Example 1: Find the largest and smallest of $Q(\vec{x}) = 3x_1^2 + 7x_2^2$ on the edge of an unit circle.

$$\text{Max: } Q(0,1) = 7 \tag{296}$$
$$\text{Min: } Q(1,0) = 3 \tag{297}$$

Example 3: Let $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 1 \\ 1 & 1 & 4 \end{pmatrix}$. Find the maximum value of the quadratic form $x^T \vec{A}\vec{x}$ subject to the constraint $\vec{x}^T\vec{x} = 1$ and find an unit vector at which this maximum value is attained.

$$Q(x_1, x_2, x_3) = 3x_1^2 + 3x_2^2 + 4x_3^2 + 4x_1x_2 + 2x_1x_3 + 2x_2x_3 \tag{298}$$
$$p(\lambda) = -(\lambda - 6)(\lambda - 3)(\lambda - 1) = 0 \tag{299}$$
$$\lambda_1 = 6, \lambda_2 = 3, \lambda_3 = 1 \tag{300}$$
$$Q_D = 6y_1^2 + 3y_2^2 + y_3^2 \tag{301}$$
$$\text{Max: } Q_D(1,0,0) = 6 \tag{302}$$

$$P\vec{y} = \vec{x} \Leftarrow \text{Unit Length Eigenvector in 6-eigenspace} \tag{303}$$

$$A - 6I = \begin{pmatrix} -3 & 2 & 1 \\ 2 & -3 & 1 \\ 1 & 1 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -2 \\ 0 & 5 & -5 \\ 0 & -5 & 5 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \tag{304}$$

$$\vec{x} = \begin{pmatrix} \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix} \tag{305}$$

$$\text{Max: } Q(\vec{x}) = 6 \tag{306}$$

22.3 Constrained Optimization and Eigenvalues

Theorem 22.1. If $Q = \vec{x}^T A\vec{x}$, A is a real $n \times n$ symmetric matrix, with eigenvalues:

$$\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n$$

and associated normalized eigenvectors:

$$\vec{u_1}, \vec{u_2}, \ldots, \vec{u_n}$$

Then, subject to the constraint $||\vec{x}|| = 1$,

- the maximum value of $Q(\vec{x}) = \lambda_1$, attained at $\vec{x} = \pm\vec{u_1}$.
- the minimum value of $Q(\vec{x}) = \lambda_n$, attained at $\vec{x} = \pm\vec{u_n}$.

23 Singular Value Decomposition

23.1 Introduction

$$A = U\Sigma V^T \tag{307}$$

Facts:

- U and V^T is an orthogonal matrix.
- Σ is a diagonal matrix and the same size as A.

Steps:

1. Compute $A^T A$.
2. Find eigenvalues of $A^T A$, call them σ_i^2.
3. Find orthonormal eigenvectors of $A^T A$, call them v_i.

4. Compute $u_i = \frac{1}{\sigma_i} A v_i$
5. $A = U\Sigma V^T$, where $U = \begin{pmatrix} u_1 & u_2 & \dots & u_m \end{pmatrix}$ and $V = \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}$ are both orthogonal matrices.

Theorem 23.1. Singular Value Decomposition
A $m \times n$ matrix with rank r and non-zero singular values $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$ has a decomposition $U\Sigma V^T$ where

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}_{m \times n} = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 & \\ 0 & \sigma_2 & \dots & \cdot & 0 \\ \cdot & \cdot & \ddots & \cdot & \\ 0 & 0 & \dots & \sigma_r & \\ & 0 & & & 0 \end{pmatrix}$$

U is $m \times m$ orthogonal matrix, and V is a $n \times n$ orthogonal matrix.

23.2 Computing the SVD

Example 1: Square, Full Rank Matrix

$$\begin{pmatrix} 3 & 0 \\ 8 & 3 \end{pmatrix} = \begin{pmatrix} \vec{u_1} & \vec{u_2} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} \vec{v_1} & \vec{v_2} \end{pmatrix}^T \tag{308}$$

Procedure:

1. Compute $A^T A$ and find λ's.
2. Find $\vec{v_1}, \vec{v_2}$ orthonormal eigenvectors of $A^T A$.
3. Find $\vec{u_1}, \vec{u_2}$ using $\vec{u_i} = \frac{1}{\sigma_i} A \vec{v_i}$.

$$A^T A = \begin{pmatrix} 3 & 8 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 8 & 3 \end{pmatrix} = \begin{pmatrix} 73 & 24 \\ 24 & 9 \end{pmatrix} \tag{309}$$

$$p(\lambda) = \lambda^2 - 82\lambda + 81 = 0 \tag{310}$$

$$= (\lambda - 81)(\lambda - 1) = 0 \tag{311}$$

$$\lambda = 81, 1 \tag{312}$$

$$\sigma = 9, 1 \tag{313}$$

$$A^T A - 81 I_2 \Rightarrow \begin{pmatrix} 1 & -3 \\ 0 & 0 \end{pmatrix} \quad v_1 = \begin{pmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix} \tag{314}$$

$$A^T A - I_2 \Rightarrow \begin{pmatrix} 1 & \frac{1}{3} \\ 0 & 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} -\frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{pmatrix} \tag{315}$$

$$\vec{u_1} = \frac{1}{9}\begin{pmatrix} 3 & 0 \\ 8 & 3 \end{pmatrix}\begin{pmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{pmatrix} \tag{316}$$

$$\vec{u_2} = \frac{1}{9}\begin{pmatrix} 3 & 0 \\ 8 & 3 \end{pmatrix}\begin{pmatrix} -\frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{pmatrix} = \begin{pmatrix} -\frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{pmatrix} \tag{317}$$

$$\begin{pmatrix} 3 & 0 \\ 8 & 3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \end{pmatrix}\begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ -\frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix} \tag{318}$$

Example 2: Non-square, Full Rank Matrix

$$\begin{pmatrix} 3 & -3 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \vec{u_1} & \vec{u_2} & \vec{u_3} \end{pmatrix}\begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \vec{v_1} & \vec{v_2} \end{pmatrix}^T \tag{319}$$

Procedure:

1. Compute $A^T A$ and find λ's.
2. Find $\vec{v_1}, \vec{v_2}$ orthonormal eigenvectors of $A^T A$.
3. Find $\vec{u_1}, \vec{u_2}$ using $\vec{u_i} = \frac{1}{\sigma_i} A\vec{v_i}$.
4. Find $\vec{u_3}$ such that it is orthogonal to $\vec{u_1}, \vec{u_2}$ and is of unit length.

$$A^T A = \begin{pmatrix} 3 & 0 & 1 \\ -3 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & -3 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 10 & -8 \\ -8 & 10 \end{pmatrix} \tag{320}$$

$$p(\lambda) = \lambda^2 - 20\lambda + 36 = 0 \tag{321}$$

$$= (\lambda - 18)(\lambda - 2) = 0 \tag{322}$$

$$\lambda = 18, 2 \tag{323}$$

$$\sigma = 3\sqrt{2}, \sqrt{2} \tag{324}$$

$$A^T A - 18I_2 \Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \tag{325}$$

$$A^T A - 2I_2 \Rightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \tag{326}$$

$$\vec{u_1} = \frac{1}{3\sqrt{2}} \begin{pmatrix} 3 & -3 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{327}$$

$$\vec{u_2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 3 & 0 \\ 8 & 3 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{328}$$

$$\vec{u_3} \cdot \vec{u_1} = 0, \;\; \vec{u_3} \cdot \vec{u_2} = 0, \;\; ||\vec{u_3}|| = 1 \tag{329}$$

$$\vec{u_3} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{330}$$

$$\begin{pmatrix} 3 & 0 \\ 8 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{331}$$

Example 3: Tall, Full Rank Matrix

$$\begin{pmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \vec{u_1} & \vec{u_2} & \vec{u_3} & \vec{u_4} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \vec{v_1} & \vec{v_2} \end{pmatrix}^T \tag{332}$$

Procedure:

1. Compute $A^T A$ and find λ's.
2. Find $\vec{v_1}, \vec{v_2}$ orthonormal eigenvectors of $A^T A$.
3. Find $\vec{u_1}, \vec{u_2}$ using $\vec{u_i} = \frac{1}{\sigma_i} A\vec{v_i}$ for nonzero σ.
4. Find $\vec{u_3}, \vec{u_4}$ such that it forms an orthogonal basis for $\Re^4$. Do this by finding the Nul $\begin{pmatrix} \vec{u_1}^T \\ \vec{u_2}^T \end{pmatrix}$, since Nul $\begin{pmatrix} \vec{u_1}^T \\ \vec{u_2}^T \end{pmatrix} = \left(\text{Row}\begin{pmatrix} \vec{u_1} & \vec{u_2} \end{pmatrix}\right)^{\perp}$.

$$A^T A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} \tag{333}$$

$$\lambda = 9, 4 \tag{334}$$

$$\sigma = 3, 2 \tag{335}$$

$$A^TA - 9I_2 \Rightarrow \begin{pmatrix} -5 & 0 \\ 0 & 0 \end{pmatrix} \quad v_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{336}$$

$$A^TA - 4I_2 \Rightarrow \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix} \quad v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{337}$$

$$\vec{u_1} = \frac{1}{3} \begin{pmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} \tag{338}$$

$$\vec{u_2} = \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & -3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{339}$$

$$\begin{pmatrix} \vec{u_1}^T \\ \vec{u_2}^T \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{340}$$

$$\vec{u_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ \vec{u_4} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{341}$$

$$\begin{pmatrix} 3 & 0 \\ 8 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{342}$$

Example 4: Tall, Non-Full Rank Matrix

$$\begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} \vec{u_1} & \vec{u_2} & \vec{u_3} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \vec{v_1} & \vec{v_2} \end{pmatrix}^T \tag{343}$$

Procedure:

1. Compute A^TA and find λ's.
2. Find $\vec{v_1}, \vec{v_2}$ orthonormal eigenvectors of A^TA.
3. Find $\vec{u_1}$ using $\vec{u_i} = \frac{1}{\sigma_i} A\vec{v_i}$ for nonzero σ.
4. Find $\vec{u_2}, \vec{u_3}$ such that it forms an orthogonal basis for $\Re^4$. One way to do this is by finding the Nul $\left(\vec{u_1}^T\right)$, since Nul $\left(\vec{u_1}^T\right) = \left(\text{Row}\left(\vec{u_1}\right)\right)^\perp$, and then performing the

Gram-Schmidt algorithm on the null space basis to get an orthonormal basis. Another way to do this is by performing the Gram-Schmidt algorithm on $\{\vec{u_1}, \vec{e_1}, \vec{e_2}\}$.

$$A^T A = \begin{pmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 9 & -9 \\ -9 & 9 \end{pmatrix} \tag{344}$$

$$p(\lambda) = \lambda^2 - 18\lambda = 0 \tag{345}$$

$$= \lambda(\lambda - 18) = 0 \tag{346}$$

$$\lambda = 18, 0 \tag{347}$$

$$\sigma = 3\sqrt{2}, 0 \tag{348}$$

$$A^T A - 18I_2 \Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad v_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \tag{349}$$

$$A^T A - 0I_2 \Rightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \tag{350}$$

$$\vec{u_1} = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{pmatrix} \tag{351}$$

$$\begin{pmatrix} \vec{u_1}^T \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & -2 & 2 \end{pmatrix} \tag{352}$$

$$\vec{x} = r \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \tag{353}$$

$$\tag{354}$$

Performing the Gram-Schmidt Algorithm on $\left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right\}$,

$$\vec{y_1} = \vec{x_1} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \tag{355}$$

$$\vec{y_2} = \vec{x_2} - \frac{\vec{x_2} \cdot \vec{y_1}}{\vec{y_1} \cdot \vec{y_1}} \vec{y_1} = \begin{pmatrix} -2 \\ 4 \\ 5 \end{pmatrix} \tag{356}$$

$$\vec{u_1} = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{pmatrix} \tag{357}$$

$$\vec{u_2} = \begin{pmatrix} -\frac{2}{\sqrt{45}} \\ \frac{4}{\sqrt{45}} \\ \frac{5}{\sqrt{45}} \end{pmatrix} \tag{358}$$

$$\begin{pmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} & \frac{2}{\sqrt{5}} & -\frac{2}{\sqrt{45}} \\ \frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{\sqrt{45}} \\ -\frac{2}{3} & 0 & \frac{5}{\sqrt{45}} \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{359}$$

Example 5: Wide, Full Rank Matrix

$$\begin{pmatrix} 3 & 0 & 1 \\ -3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \vec{u_1} & \vec{u_2} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \end{pmatrix} \begin{pmatrix} \vec{v_1} & \vec{v_2} & \vec{v_3} \end{pmatrix}^T \tag{360}$$

Procedure:

1. Since finding the SVD of a wide matrix is more computationally intensive, one can find the SVD of $B = A^T$ and then apply $A = (U\Sigma V^T)^T = V\Sigma^T U^T$.
2. Compute $B^T B$ and find λ's.
3. Find $\vec{v_1}, \vec{v_2}$ orthonormal eigenvectors of $B^T B$.
4. Find $\vec{u_1}, \vec{u_2}$ using $\vec{u_i} = \frac{1}{\sigma_i} A\vec{v_i}$ for nonzero σ.
5. Find $\vec{u_3}$ such that it forms an orthogonal basis for $\Re^3$.

$$B^T B = \begin{pmatrix} 3 & 0 & 1 \\ -3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -3 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 10 & -8 \\ -8 & 10 \end{pmatrix} \tag{361}$$

$$p(\lambda) = \lambda^2 - 20\lambda + 36 = 0 \tag{362}$$

$$= (\lambda - 18)(\lambda - 2) = 0 \tag{363}$$

$$\lambda = 18, 2, 0 \tag{364}$$

$$\sigma = 3\sqrt{2}, \sqrt{2} \tag{365}$$

$$B^TB - 18I_2 \Rightarrow \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \tag{366}$$

$$B^TB - 2I_2 \Rightarrow \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} \quad v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \tag{367}$$

$$\vec{u_1} = \frac{1}{3\sqrt{2}} \begin{pmatrix} 3 & -3 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{368}$$

$$\vec{u_2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 3 & -3 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{369}$$

$$\vec{u_3} = \vec{u_1} \times \vec{u_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \tag{370}$$

$$A^T = B = \begin{pmatrix} 3 & -3 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{371}$$

$$A = V\Sigma^TU^T = \begin{pmatrix} 3 & 0 & 1 \\ -3 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{372}$$

23.3 Low Rank Approximation of A using SVD

Theorem 23.2. Spectral Decomposition Theorem
Since the singular values in the Σ matrix are ordered from largest to smallest, the following can be used to derive the n-rank approximation of A_n.

$$A_n = \sum_{i=1}^{n} \sigma_i \vec{u_i} \vec{v_i}^T \tag{373}$$

Example 1: Compute the spectral decomposition.

$$A = \begin{pmatrix} 3 & -3 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3\sqrt{2} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{374}$$

$$= 3\sqrt{2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} + \sqrt{2} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{375}$$

$$= 3\sqrt{2} \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} + \sqrt{2} \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{376}$$

23.4 Singular Values

Definition 23.3. Singular Values
The singular values of A are defined to be:

$$\sigma_1 = \sqrt{\lambda_1} \geq \sigma_1 = \sqrt{\lambda_2} \cdots \geq \sigma_n = \sqrt{\lambda_n}$$

The matrix $A^T A$ is always symmetric, with non-negative eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. Let $\{\vec{v_1}, \ldots, \vec{v_n}\}$ be the associated orthonormal eigenvectors. Then

$$||A\vec{v_j}||^2 = (A\vec{v_j}) \cdot (A\vec{v_j}) \tag{377}$$

$$= \vec{v_j}^T A^T A \vec{v_j} \tag{378}$$

$$= \lambda_j \vec{v_j} \cdot \vec{v_j} \tag{379}$$

$$= \lambda_j ||\vec{v_j}||^2 \tag{380}$$

$$= \lambda_j \tag{381}$$

If A has rank r, then $\{A\vec{v_1}, \ldots, A\vec{v_r}\}$ is an orthogonal basis for Col$A$. For $1 \leq j < k \leq r$

$$(A\vec{v_j})^T A\vec{v_k} = \vec{v_j}^T A^T A \vec{v_k} \tag{382}$$

$$= \lambda_k \vec{v_j}^T \vec{v_k} \tag{383}$$

$$= \lambda_k \vec{v_j} \cdot \vec{v_k} = 0 \tag{384}$$

23.5 SVD and the Four Subspaces

Suppose the rank A is r. Given the SVD of A,

$$A = U \Sigma V^T \tag{385}$$

An orthonormal basis for the Col A are the first r columns of U. The remaining columns of U form an orthonormal basis for the Nul A^T. An orthonormal basis for the Row A are the first r columns of V. Thus, an orthonormal basis for the Row A are the first r rows of V^T. The remaining rows of V^T form an orthonormal basis for the Nul A.

In other words,

1. $A\vec{v_s} = \sigma_s \vec{u_s}$
2. $\vec{v_1}, \ldots, \vec{v_r}$ is an orthonormal basis for Row A.
3. $\vec{u_1}, \ldots, \vec{u_r}$ is an orthonormal basis for Col A.
4. $\vec{v_{r+1}}, \ldots, \vec{v_n}$ is an orthonormal basis for Nul A.
5. $\vec{u_{r+1}}, \ldots, \vec{u_n}$ is an orthonormal basis for Nul A^T.

23.6 Geometric Interpretation of SVD

Since U and V^T are orthogonal, both matrices are pure rotations. Since Σ is a diagonal matrix, it is a pure dilation along the elementary basis vector directions.

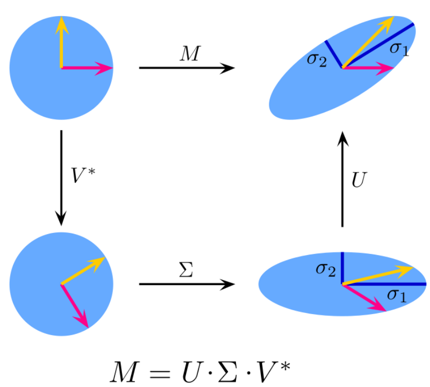

Figure 1: Geometric Interpretation of SVD

23.7 Algorithm to find SVD of A

Suppose A is $m \times n$ and has rank $r \leq n$.

1. Compute the squared singular values σ_i^2 of A^TA and construct Σ. This is done by finding the eigenvalues of A^TA, call then $\lambda_i \geq 0$ and set $\sigma_i = \sqrt{\lambda_i}$.
2. Compute the unit singular vectors $\vec{v_i}$ of A^TA and use them to form V. Since A^TA is symmetric, by the spectral theorem, the algebraic multiplicity is equal to the geometric multiplicity and we can find orthonormal basis $\{\vec{v_1}, \ldots, \vec{v_n}\}$ of $\Re^n$ constituting of eigenvectors of A^TA.
3. Compute an orthonormal basis for Col A using the following for any $\sigma_i \neq 0$:

$$\vec{u_i} = \frac{1}{\sigma_i} A\vec{v_i}, \quad i = 1, 2, \ldots, r$$

Extend the set $\{\vec{u_i}\}$ to form an orthogonal basis for $\Re^m$, use the basis for form U.

23.8 Condition Number of a Matrix

Definition 23.4. Condition Number
If A is invertible $n \times n$ matrix, the ratio:

$$\frac{\sigma_1}{\sigma_n}$$

is the condition number of A.
Note that:

- The condition number of a matrix describes the sensitivity of a solution to $A\vec{x} = \vec{b}$ is to errors in A.
- We could define the condition number for a rectangular matrix, but that would go beyond the scope of this course.