

Section 7.1 : Diagonalization of Symmetric Matrices

Chapter 7: Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. Symmetric matrices
2. Orthogonal diagonalization

Learning Objectives

1. Construct an orthogonal diagonalization of a symmetric matrix,
 $A = PDP^T$.

Symmetric Matrices

Definition

Matrix A is **symmetric** if $A^T = A$.

should be square

Example. Which of the following matrices are symmetric? Symbols $*$ and \star represent real numbers.

$$A = [*]$$

$$B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = B^T$$

$$C = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} = C^T$$

$$D = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$D^T = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

$$E = \begin{bmatrix} 4 & 2 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$E^T = \begin{bmatrix} 4 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$

$$F = \begin{bmatrix} 4 & 2 & 0 & 1 \\ 2 & 0 & 7 & 4 \\ 0 & 7 & 6 & 0 \\ 1 & 4 & 0 & 3 \end{bmatrix}$$

Examples

$$A \in \mathbb{R}^{n \times n}$$

- $A + A^T$: Symm.

$$(A + A^T)^T = A^T + A$$

- $A^T \cdot A$: Symm.

$$(A^T A)^T = A^T \cdot A$$

$A^T A$ is Symmetric for $A \in \mathbb{R}^{m \times n}$

A very common example: For **any** matrix A with columns a_1, \dots, a_n ,

$$A^T A = \begin{bmatrix} \text{---} a_1^T \text{---} \\ \text{---} a_2^T \text{---} \\ \vdots \\ \text{---} a_n^T \text{---} \end{bmatrix} \begin{bmatrix} | & | & \cdots & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & \cdots & | \end{bmatrix}$$

$$= \begin{bmatrix} a_1^T a_1 & a_1^T a_2 & \cdots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \cdots & a_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^T a_1 & a_n^T a_2 & \cdots & a_n^T a_n \end{bmatrix}$$

Entries are the dot products of columns of A .

Recall

① $u \cdot v = u^T \cdot v$ ↙ matrix multi. ↔ for real vectors

$$u \cdot v = \overline{u^T} \cdot v$$

Section 7.1 Slide 4

$$u = \begin{bmatrix} i \\ 1 \end{bmatrix} \quad u \cdot u = \overline{u^T} \cdot u = [-i \quad 1] \begin{bmatrix} i \\ 1 \end{bmatrix}$$

$$= (-i) \cdot i + 1 \cdot 1 = 2$$

② $(Ax) \cdot y = (Ax)^T \cdot y = x^T \cdot (A^T y)$

$$= x \cdot (A^T y)$$

Symmetric Matrices and their Eigenspaces

Theorem

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 to distinct eigenvalues are orthogonal subspaces.

Proof:

$$A \vec{v}_1 = \lambda_1 \vec{v}_1 \quad \lambda_1 \neq \lambda_2$$

$$A \vec{v}_2 = \lambda_2 \vec{v}_2$$

$$\begin{aligned} (\downarrow A \vec{v}_1) \cdot \vec{v}_2 &= \lambda_1 (\vec{v}_1 \cdot \vec{v}_2) \\ \parallel & \parallel \end{aligned}$$

$$\vec{v}_1 \cdot (A \vec{v}_2) = \lambda_2 (\vec{v}_1 \cdot \vec{v}_2)$$

$$\underbrace{(\lambda_1 - \lambda_2)}_0 \cdot \underbrace{(\vec{v}_1 \cdot \vec{v}_2)}_0 = 0$$

Example 1

Diagonalize A using an orthogonal matrix. Eigenvalues of A are given.

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \underline{\lambda = -1, 1}$$

Hint: Gram-Schmidt

$$\lambda = -1 : E_{-1} = \text{Nul}(A + I) = \text{Nul} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$= \text{Nul} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\lambda = 1 : E_1 = \text{Nul}(A - I) = \text{Nul} \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

$$= \text{Nul} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad v_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$A = \begin{bmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} | & | & | \\ P^{-1} & & \\ | & | & | \end{bmatrix} \begin{matrix} P^T \\ \\ \end{matrix}$$

P has orthonormal columns, square

$$\Rightarrow P^T \cdot P = I$$

$$P^T = P^{-1}$$

Spectral Theorem

Recall: If P is an orthogonal $n \times n$ matrix, then $P^{-1} = P^T$, which implies $A = PDP^T$ is diagonalizable and symmetric.

Theorem: Spectral Theorem

An $n \times n$ symmetric matrix A ^{← real} has the following properties.

1. All eigenvalues of A are real.
2. The dimension of each eigenspace is full, that it's dimension is **equal to its algebraic multiplicity**. Diagonalizable.
3. The eigenspaces are mutually orthogonal.
4. A can be diagonalized: $A = PDP^T$, where D is diagonal and P is orthogonal.

Proof (if time permits):

$$\textcircled{1} \quad A v = \lambda v, \quad \lambda \in \mathbb{C}, \quad v \in \mathbb{C}^n, \quad v \neq 0$$
$$\overline{v^T} \cdot v = v \cdot v \neq 0$$

Section 7.1 Slide 7

$$\begin{aligned} \overline{(\lambda v)} \cdot v &= \overline{A v} \cdot v = \overline{(A v)^T} \cdot v \\ &= \overline{v^T} (A^T \cdot v) = \overline{v^T} \cdot (\overline{A^T} \cdot v) \\ &= v \cdot A v \\ &= v \cdot (\lambda v) = \lambda (v \cdot v) \end{aligned}$$
$$\begin{aligned} (\lambda - \overline{\lambda}) \cdot (v \cdot v) &= 0 \\ \lambda &= \overline{\lambda} \\ \Rightarrow \lambda &= \text{real} \end{aligned}$$

② E : eigenspace for λ

If $x \in E^\perp$ then $Ax \in E^\perp$

Why? Let $y \in E$ then $Ay = \lambda y$

$$\begin{aligned} 0 &= \lambda(x \cdot y) = x \cdot (\lambda y) = x \cdot Ay \\ &= Ax \cdot y \Rightarrow Ax \in E^\perp \quad \checkmark \end{aligned}$$

11/13/24

Spectral Theorem

A : $n \times n$ real symm.

(i) Every eigenvalue is real.

(ii) A is orthogonally diagonalizable.

$$A = P \cdot D \cdot P^T = \begin{bmatrix} | & | & \dots & | \\ v_1 & v_2 & \dots & v_n \\ | & | & \dots & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} \begin{bmatrix} - & v_1^T \\ - & v_2^T \\ \vdots & \vdots \\ - & v_n^T \end{bmatrix}$$

\uparrow
orthogonal matrix

$$= \lambda_1 \cdot \underbrace{v_1 \cdot v_1^T}_{\in \mathbb{R}^{n \times n}} + \lambda_2 \cdot v_2 \cdot v_2^T + \dots + \lambda_n \cdot v_n \cdot v_n^T$$

(Spectral Decomposition)

$v_i \cdot v_i^T$ ① $n \times n$, symmetric ($(v_i \cdot v_i^T)^T = (v_i^T)^T \cdot v_i = v_i \cdot v_i^T$)

② $\text{Rank}(v_i \cdot v_i^T) = 1$

Ex

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} [1 \ 2 \ 3] = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$$

$$\begin{aligned} \textcircled{3} \quad (v_i \cdot v_i^T) \cdot y &= v_i \cdot \underbrace{(v_i^T \cdot y)} = (y \cdot v_i) \cdot v_i \\ &= \text{proj}_{v_i}(y) \end{aligned}$$

Spectral Decomposition of a Matrix

Spectral Decomposition

Suppose A can be orthogonally diagonalized as

$$A = PDP^T = [\vec{u}_1 \quad \cdots \quad \vec{u}_n] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{u}_1^T \\ \vdots \\ \vec{u}_n^T \end{bmatrix}$$

Then A has the decomposition

$$A = \lambda_1 \vec{u}_1 \vec{u}_1^T + \cdots + \lambda_n \vec{u}_n \vec{u}_n^T = \sum_{i=1}^n \lambda_i \vec{u}_i \vec{u}_i^T$$

Each term in the sum, $\lambda_i \vec{u}_i \vec{u}_i^T$, is an $n \times n$ matrix with rank 1.

Example 2

Construct a spectral decomposition for A whose orthogonal diagonalization is given.

$$\begin{aligned} A &= \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} = PDP^T \\ &= \begin{pmatrix} \underbrace{1/\sqrt{2}}_{v_1} & \underbrace{-1/\sqrt{2}}_{v_2} \\ \underbrace{1/\sqrt{2}}_{v_1} & \underbrace{1/\sqrt{2}}_{v_2} \end{pmatrix} \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \\ &= 4 \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} + 2 \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \\ &= 2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \end{aligned}$$

Section 7.2 : Quadratic Forms

Chapter 7: Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. Quadratic forms
2. Change of variables
3. Principle axes theorem
4. Classifying quadratic forms

Learning Objectives

1. Characterize and classify quadratic forms using eigenvalues and eigenvectors.
2. Express quadratic forms in the form $Q(\vec{x}) = \vec{x}^T A \vec{x}$.
3. Apply the principle axes theorem to express quadratic forms with no cross-product terms.

Motivating Question Does this inequality hold for all x, y ?

$$\underbrace{x^2 - 6xy + 9y^2}_{||} \geq 0$$

$$\frac{(x - 3y)^2}{\uparrow}$$

Diagonalization

Quadratic Forms

Definition

A **quadratic form** is a function $Q : \mathbb{R}^n \rightarrow \mathbb{R}$, given by

$$Q(\vec{x}) = \vec{x}^T A \vec{x} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix}$$

Matrix A is $n \times n$ and symmetric.

In the above, \vec{x} is a vector of variables.

Example 1

Compute the quadratic form $\vec{x}^T A \vec{x}$ for the matrices below.

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

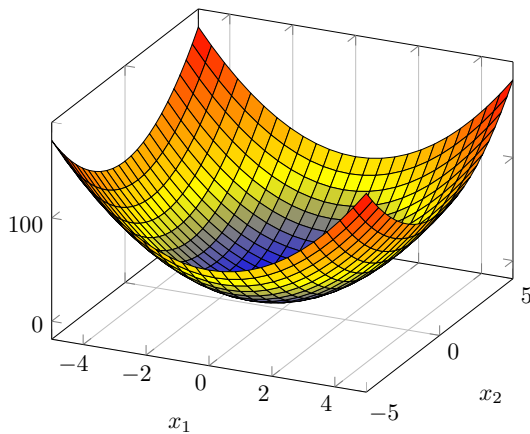
$$\begin{aligned} Q_A(x) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} 4x_1 & 3x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 4 \cdot x_1^2 + 3 \cdot x_2^2 \neq 0 \end{aligned}$$

$$\begin{aligned} Q_B(x) &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= \begin{bmatrix} 4x_1 + x_2 & x_1 - 3x_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &= (4x_1 + x_2) x_1 + (x_1 - 3x_2) x_2 \end{aligned}$$

$$\begin{aligned} &= 4x_1^2 + x_1 \cdot x_2 + x_1 \cdot x_2 - 3x_2^2 \\ &= 4x_1^2 + 2x_1 \cdot x_2 - 3x_2^2 \end{aligned}$$

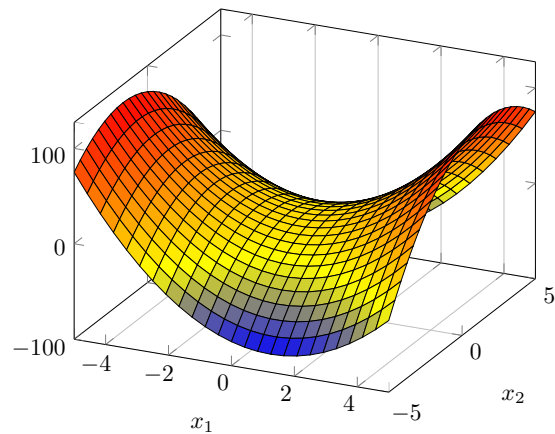
Example 1 - Surface Plots

The surfaces for Example 1 are shown below.



$$y = 4x_1^2 + 3x_2^2$$

Students are not expected to be able to sketch quadratic surfaces, but it is helpful to see what they look like.



$$y = 4x_1^2 + 2x_1x_2 - 3x_2^2$$

Example 2

Write Q in the form $\vec{x}^T A \vec{x}$ for $\vec{x} \in \mathbb{R}^3$.

$$Q(x) = 5x_1^2 - x_2^2 + 3x_3^2 + 6x_1x_3 - 12x_2x_3 + 0 \cdot x_1x_2$$
$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 5 & 0 & +3 \\ 0 & -1 & -6 \\ +3 & -6 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

In general,

$$\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots \\ & & & & \\ & & & & \\ & & & & \\ & & & & \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$
$$= \sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j$$

Change of Variable

If \vec{x} is a variable vector in \mathbb{R}^n , then a **change of variable** can be represented as

$$\vec{x} = P\vec{y}, \quad \text{or} \quad \vec{y} = P^{-1}\vec{x}$$

With this change of variable, the quadratic form $\vec{x}^T A \vec{x}$ becomes:

Ex

$$x_1^2 - 6x_1x_2 + 10x_2^2$$

$$= \underbrace{(x_1 - 3x_2)^2}_{y_1} + \underbrace{x_2^2}_{y_2}$$

$$y_1 = x_1 - 3x_2$$

$$y_2 = 0 \cdot x_1 + 1 \cdot x_2$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \longrightarrow \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Section 7.2 Slide 14

$$Q_A(x) = x^T \cdot A \cdot x = \underbrace{x^T \cdot P}_{y^T} \cdot D \cdot \underbrace{P^T x}_y$$

$$y^T = (P^T \cdot x)^T$$

$$= x^T \cdot P$$

$$= y^T \cdot D \cdot y$$

$$= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$$

Example 3

Make a change of variable $\vec{x} = P\vec{y}$ that transforms $Q = \vec{x}^T A \vec{x}$ so that it does not have cross terms. The orthogonal decomposition of A is given.

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

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

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

$$y = P^T \cdot x$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$Q(x) = 3x_1^2 + 4x_1x_2 + 6x_2^2$$

$$= y^T \cdot D \cdot y$$

$$= 2y_1^2 + 7y_2^2$$

$$= 2 \cdot \left(\frac{1}{\sqrt{5}} (2x_1 - x_2) \right)^2 + 7 \cdot \left(\frac{1}{\sqrt{5}} (x_1 + 2x_2) \right)^2$$

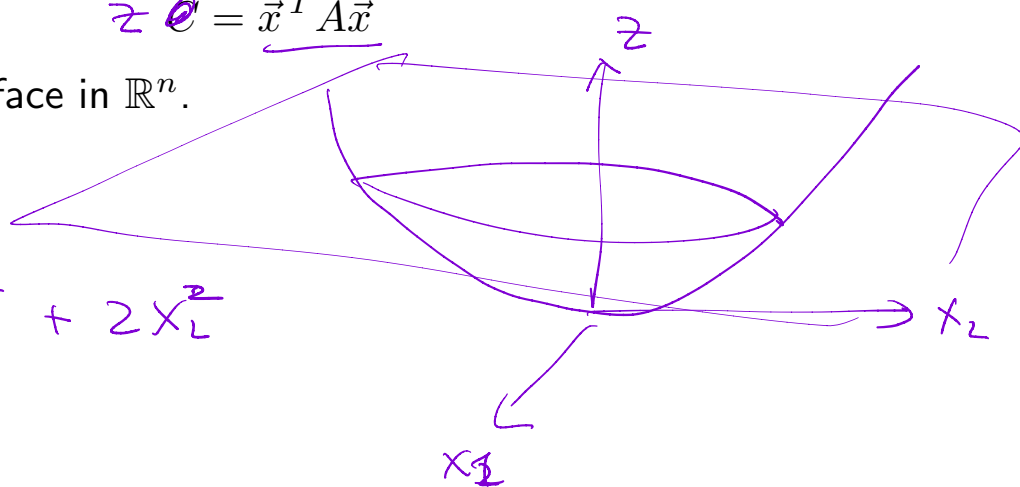
Geometry

Suppose $Q(\vec{x}) = \vec{x}^T A \vec{x}$, where $A \in \mathbb{R}^{n \times n}$ is symmetric. Then the set of \vec{x} that satisfies

$$z = \vec{x}^T A \vec{x}$$

defines a curve or surface in \mathbb{R}^n .

$$z = 4x_1^2 + 2x_2^2$$



$\{v_1, \dots, v_n\} = \text{ONB for } \mathbb{R}^n$

Note

$$\underbrace{(v_1 \cdot x)^2 + (v_2 \cdot x)^2 + \dots + (v_n \cdot x)^2}_{\| \text{PT. } x \|^2} = \|x\|^2$$

Pythagorean Thm.

Section 7.2 Slide 16

$$\begin{bmatrix} v_1 \cdot x \\ v_2 \cdot x \\ \vdots \\ v_n \cdot x \end{bmatrix} = \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} \cdot x = \begin{bmatrix} v_1 \cdot x \\ v_2 \cdot x \\ \vdots \\ v_n \cdot x \end{bmatrix}$$

Recall

$$\begin{aligned} Q(x) &= x^T \cdot A \cdot x = \sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j \\ &\quad \uparrow \text{real } n \times n \text{ symm.} \\ &= x^T \cdot (P \cdot D \cdot P^T) \cdot x = (x^T \cdot P) \cdot D \cdot (P^T x) \\ &\quad \text{y}^T = \underline{\underline{P^T x}} \\ &= y^T \cdot D \cdot y \\ &\quad \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2 \\ &= \lambda_1 (v_1 \cdot x)^2 + \lambda_2 (v_2 \cdot x)^2 + \dots + \lambda_n (v_n \cdot x)^2 \end{aligned}$$

Principle Axes Theorem

Theorem

If A is a $n \times n$ real symm. 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.

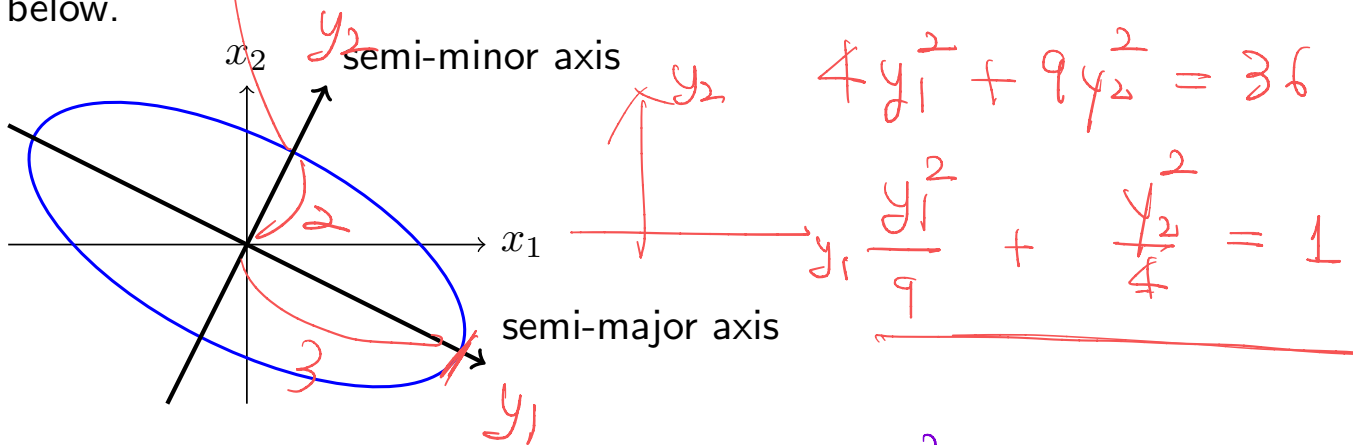
Proof (if time permits):

$$\text{or } y = P^T \cdot x$$

$$y_1 = 0 \Rightarrow 2x_1 - x_2 = 0 \\ x_2 = 2 \cdot x_1$$

Example 5

Compute the quadratic form $Q = \vec{x}^T A \vec{x}$ for $A = \begin{pmatrix} 5 & 2 \\ 2 & 8 \end{pmatrix}$, and find a change of variable that removes the cross-product term. A sketch of Q is below.



Eigenvalue : $\text{Det}(A - \lambda I) = \lambda^2 - 13\lambda + 36 = 0$

$$= (\lambda - 9)(\lambda - 4)$$

$\lambda = 4$: $\text{Nul}(A - 4I) = \text{Nul} \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \text{Nul} \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}$

$$v_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

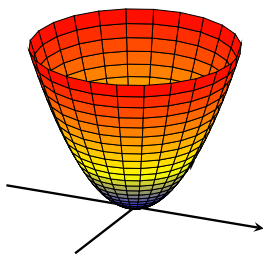
$\lambda = 9$: $\text{Nul}(A - 9I) = \text{Nul} \begin{pmatrix} -4 & 2 \\ 2 & -1 \end{pmatrix} = \text{Nul} \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}$

$$v_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

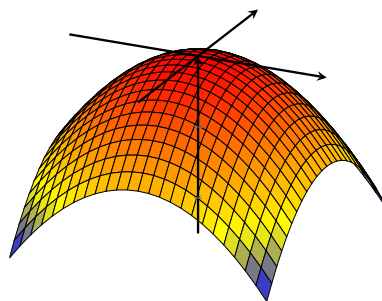
$$\begin{aligned}
 Q(x) &= 4 \cdot y_1^2 + 9 \cdot y_2^2 = 4 \cdot (v_1 \cdot x)^2 + 9 \cdot (v_2 \cdot x)^2 \\
 &= 4 \left(\frac{1}{\sqrt{5}} (2x_1 - x_2) \right)^2 + 9 \left(\frac{1}{\sqrt{5}} (x_1 + 2x_2) \right)^2 \\
 &= 36
 \end{aligned}$$

Classifying Quadratic Forms

$$Q = x_1^2 + x_2^2$$



$$Q = -x_1^2 - x_2^2$$



Definition

A quadratic form Q is

1. **positive definite** if $Q(x) > 0$ for all $\vec{x} \neq \vec{0}$. ($Q(x) = 0$ if $x = 0$)
2. **negative definite** if $Q(x) < 0$ for all $\vec{x} \neq \vec{0}$. ($Q(x) = 0$ if $x = 0$)
3. **positive semidefinite** if $Q(x) \geq 0$ for all \vec{x} .
4. **negative semidefinite** if $Q(x) \leq 0$ for all \vec{x} .
5. **indefinite** if there exist x_1, x_2 such that

$$Q(x_1) > 0, \quad Q(x_2) < 0$$

Quadratic Forms and Eigenvalues

Theorem

If A is a $n \times n$ real symm. matrix with eigenvalues λ_i , then $Q = \vec{x}^T A \vec{x}$ is

1. **positive definite** iff $\lambda_i \geq 0$ for all $i = 1, 2, \dots, n$
2. **negative definite** iff $\lambda_i < 0$ "
3. **indefinite** iff $\lambda_i \geq 0$ $\lambda_j < 0$ for some (i, j) .

Proof (if time permits):

Example 6

We can now return to our motivating question (from first slide): does this inequality hold for all x, y ?

$$Q(x) = \underbrace{x^2 - 6xy + 9y^2}_{\geq 0}$$

$$A = \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix}$$

$$\lambda: \quad \text{Det}(A - \lambda I) = \lambda^2 - 10\lambda + 0 = 0$$

$$\lambda = 0, 10 \geq 0$$

Q is Positive Semi-definite

$$Q(x) \geq 0.$$

Section 7.3 : Constrained Optimization

Chapter 7: Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. Constrained optimization as an eigenvalue problem
2. Distance and orthogonality constraints

Learning Objectives

1. Apply eigenvalues and eigenvectors to solve optimization problems that are subject to distance and orthogonality constraints.

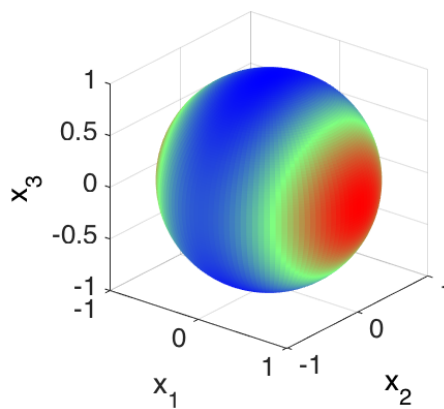
Example 1

The surface of a unit sphere in \mathbb{R}^3 is given by

$$1 = x_1^2 + x_2^2 + x_3^2 = \|\vec{x}\|^2$$

Q is a quantity we want to optimize

$$Q(\vec{x}) = \underline{9x_1^2 + 4x_2^2 + 3x_3^2}$$



Find the largest and smallest values of Q on the surface of the sphere.

$$3(x_1^2 + x_2^2 + x_3^2) \leq Q(x) \leq 9 \cdot (x_1^2 + x_2^2 + x_3^2)$$

$$3 \leq Q(x) \leq 9$$

↑ max?

$$Q(x) = 9 \quad \text{if} \quad x = \pm \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$Q(x) = 3 \quad \text{if} \quad x = \pm \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

A Constrained Optimization Problem

Suppose we wish to find the maximum or minimum values of

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

subject to

$$\|\vec{x}\| = 1$$

That is, we want to find

$$m = \min\{Q(\vec{x}) : \|\vec{x}\| = 1\}$$

$$M = \max\{Q(\vec{x}) : \|\vec{x}\| = 1\}$$

This is an example of a **constrained optimization** problem. Note that we may also want to know where these extreme values are obtained.

Constrained Optimization and Eigenvalues

Theorem

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

$$\lambda_1 \geq \lambda_2 \dots \geq \lambda_n$$

and associated normalized eigenvectors

$$\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$$

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 $\vec{x} = \pm \vec{u}_n$.

Proof:

11/18/24

$$Q(x) = x^T \cdot A \cdot x \quad \text{max/min of } Q \text{ under } \|x\| = 1.$$

$$= x^T \cdot P \cdot D \cdot P^T x$$

$$= (P^T x)^T \cdot D \cdot (P^T x) \stackrel{y}{=} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2$$

$$(\nu_1 \cdot x)^2 + (\nu_2 \cdot x)^2 + \dots + (\nu_n \cdot x)^2 = \|x\|^2 = 1$$

Example 2

Calculate the maximum and minimum values of $Q(\vec{x}) = \vec{x}^T A \vec{x}$, $\vec{x} \in \mathbb{R}^3$, subject to $\|\vec{x}\| = 1$, and identify points where these values are obtained.

$$Q(\vec{x}) = x_1^2 + 2x_2x_3$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\text{Det}(A - \lambda I) = \text{Det} \begin{bmatrix} 1-\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & 1 & -\lambda \end{bmatrix} = (1-\lambda) \cdot \text{det} \begin{bmatrix} -\lambda & 1 \\ 1 & -\lambda \end{bmatrix}$$

$$= (1-\lambda)(\lambda^2 - 1) = (1-\lambda)(\lambda+1)(\lambda-1)$$

$$\lambda = 1, -1, 1$$

$$\lambda = 1:$$

$$\text{Nul}(A - I) = \text{Nul} \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

$$\lambda = -1$$

$$\text{Nul}(A + I) = \text{Nul} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

$$v_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

$$Q(x) = 1 \cdot (v_1 \cdot x)^2 + 1 \cdot (v_2 \cdot x)^2 + (-1) \cdot (v_3 \cdot x)^2$$

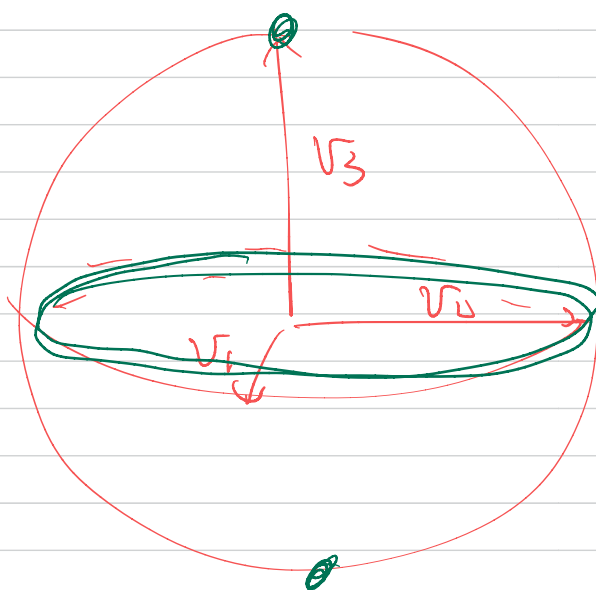
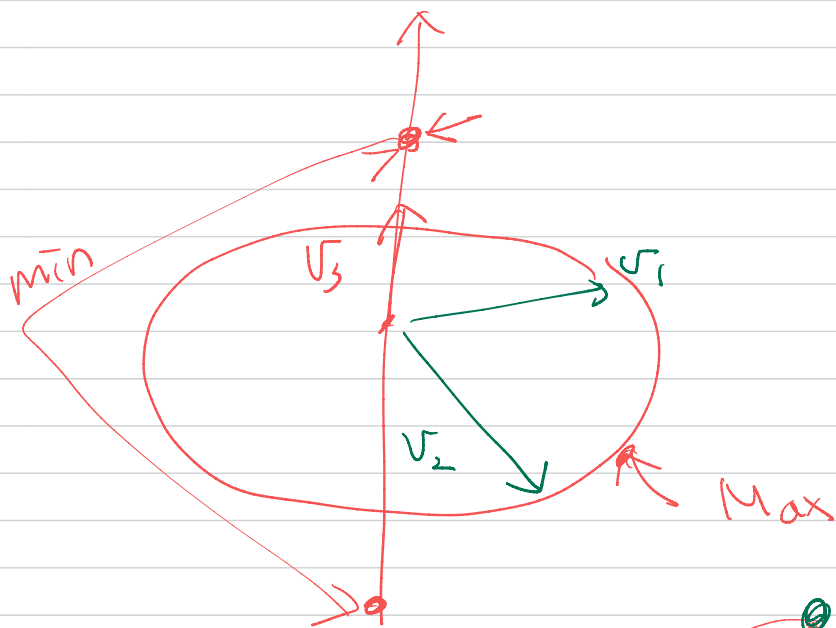
$$\|x\|^2 = (v_1 \cdot x)^2 + (v_2 \cdot x)^2 + (v_3 \cdot x)^2 = 1$$

Max : $Q(x) = 1$ when $x \cdot v_3 = 0$

Min : $Q(x) = -1$ when

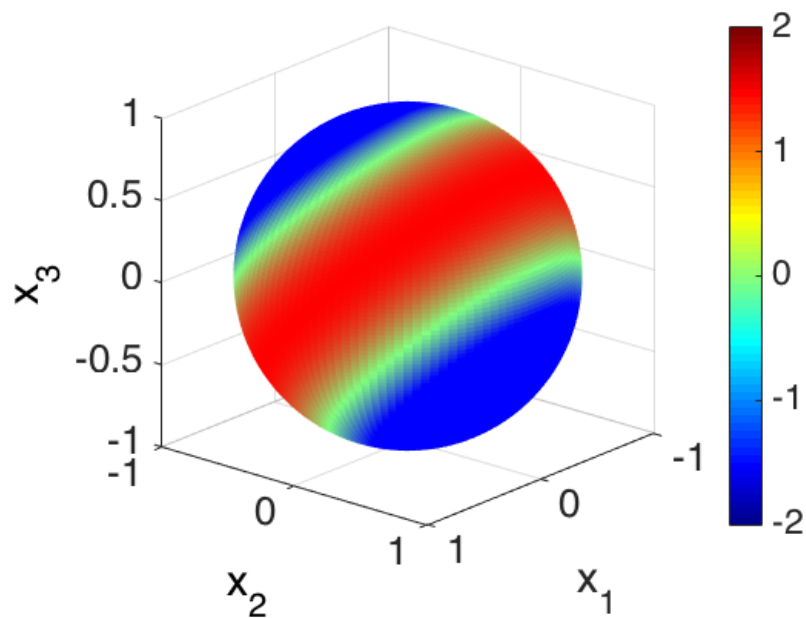
$$\begin{aligned} v_1 \cdot x &= 0 \\ v_2 \cdot x &= 0 \end{aligned}$$

$$x = v_3, -v_3$$



Example 2

The image below is the unit sphere whose surface is colored according to the quadratic from the previous example. Notice the agreement between our solution and the image.



An Orthogonality Constraint

Theorem

Suppose $Q = \vec{x}^T A \vec{x}$, A is a real $n \times n$ symmetric matrix, with eigenvalues

$$\lambda_1 \geq \lambda_2 \dots \geq \lambda_n$$

and associated eigenvectors

$$\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n$$

Subject to the constraints $\|\vec{x}\| = 1$ and $\vec{x} \cdot \vec{u}_1 = 0$,

- The maximum value of $Q(\vec{x}) = \lambda_2$, attained at $\vec{x} = \vec{u}_2$.
- The minimum value of $Q(\vec{x}) = \lambda_n$, attained at $\vec{x} = \vec{u}_n$.

Note that λ_2 is the second largest eigenvalue of A .

$$\max_{\|\vec{x}\|=1} \vec{x}^T A \cdot \vec{x} = \text{the largest eigenvalue}$$

$$\max_{\substack{\|\vec{x}\|=1 \\ \vec{x} \cdot \vec{v}_1 = 0}} \vec{x}^T A \cdot \vec{x} = \text{2}^{\text{nd}} \text{ largest eigenvalue.}$$

Example 3

Calculate the maximum value of $Q(\vec{x}) = \vec{x}^T A \vec{x}$, $\vec{x} \in \mathbb{R}^3$, subject to $\|\vec{x}\| = 1$ and to $\vec{x} \cdot \vec{u}_1 = 0$, and identify a point where this maximum is obtained.

$$Q(\vec{x}) = x_1^2 + 2x_2x_3, \quad \vec{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = v_1$$

$$\lambda : 1, 1, -1$$

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = v_2, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = v_3$$

$$Q(x) = 1 \cdot (v_1 \cdot x)^2 + 1 \cdot (v_2 \cdot x)^2 + (-1) \cdot (v_3 \cdot x)^2$$

$$\begin{cases} (v_1 \cdot x)^2 + (v_2 \cdot x)^2 + (v_3 \cdot x)^2 = \|x\|^2 = 1 \\ (v_1 \cdot x) = 0 \end{cases}$$

Section 7.3 Slide 30

$$\text{Max} : Q(x) = 1 \quad \text{when}$$

second largest eigenvalue \Rightarrow eigenvector.

$$\begin{aligned} v_1 \cdot x &= 0 \\ v_3 \cdot x &= 0 \end{aligned}$$

$$x = v_2, -v_2$$

$$\text{Min} : Q(x) = -1$$

when

$$x = \pm v_3$$

Example 4 (if time permits)

Calculate the maximum value of $Q(\vec{x}) = \vec{x}^T A \vec{x}$, $\vec{x} \in \mathbb{R}^3$, subject to $\|\vec{x}\| = 5$, and identify a point where this maximum is obtained.

$$Q(\vec{x}) = x_1^2 + 2x_2x_3$$

$$\begin{aligned} Q &= \vec{x}^T \cdot A \cdot \vec{x} \\ &= \lambda_1 (\vec{v}_1 \cdot \vec{x})^2 + \dots \\ (\vec{v}_1 \cdot \vec{x})^2 + \dots &= \|\vec{x}\|^2 = 25 \end{aligned}$$

Section 7.4 : The Singular Value Decomposition

Chapter 7: Orthogonality and Least Squares

Math 1554 Linear Algebra

Topics and Objectives

Topics

1. The Singular Value Decomposition (SVD) and some of its applications.

Learning Objectives

1. Compute the SVD for a rectangular matrix.
2. Apply the SVD to
 - ▶ estimate the rank and condition number of a matrix,
 - ▶ construct a basis for the four fundamental spaces of a matrix, and
 - ▶ construct a spectral decomposition of a matrix.

Recall (Spectral Decomposition)

A : $n \times n$ real symm.

\Rightarrow (i) eigenvalues are real

(ii) Diagonalizable $A = P \cdot D \cdot P^T$

(iii) $A = \lambda_1 \cdot v_1 \cdot v_1^T + \lambda_2 v_2 \cdot v_2^T + \dots + \lambda_n v_n \cdot v_n^T$

Why (ii) \Rightarrow (iii) :

From (ii) $\{v_1, \dots, v_n\}$: ONB, eigenvectors

$$P = [v_1 \dots v_n], \quad P^T \cdot P = I = \underline{P \cdot P^T}$$

$$A \cdot I = A (v_1 \cdot v_1^T + v_2 \cdot v_2^T + \dots + v_n \cdot v_n^T)$$

$$\begin{aligned} A &= (A v_1) \cdot v_1^T + (A v_2) \cdot v_2^T + \dots + (A v_n) \cdot v_n^T \\ &= \lambda_1 v_1 \cdot v_1^T + \lambda_2 v_2 \cdot v_2^T + \dots + \lambda_n v_n \cdot v_n^T. \end{aligned}$$

A : $m \times n$ real,

p.s.d.

$\Rightarrow A^T \cdot A$: $n \times n$ real symm. positive semi-definite.

$$Q(x) = (x^T \cdot A^T) A \cdot x = (Ax)^T \cdot Ax = \|Ax\|^2 \geq 0$$

eigenvalues for $A^T \cdot A$

$$\Rightarrow \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \lambda_{r+2} = \dots = \lambda_n = 0$$

eigenvectors for $A^T \cdot A$

$\{v_1, v_2, \dots, v_n\}$ ONB for \mathbb{R}^n

$$v_1 \cdot v_1^T + \dots + v_n \cdot v_n^T = I_n$$

$$\Rightarrow A = A v_1 \cdot v_1^T + \dots + A v_n \cdot v_n^T$$

$$\textcircled{1} A^T \cdot A \cdot v_{r+1} = \overset{=0}{\lambda_{r+1}} \cdot v_{r+1} = 0 = A^T A v_{r+1} = \dots = A^T \cdot A v_n$$

$$A^T \cdot A v_k = 0 \quad \text{for } k = r+1, \dots, n.$$

$$Q(v_k) = v_k^T \underbrace{A^T \cdot A \cdot v_k}_0 = 0 = \|A \cdot v_k\|^2$$

$$A v_k = 0, \quad \text{for } k = r+1, \dots, n.$$

$$\boxed{\text{Nul}(A) = \text{Nul}(A^T \cdot A)}$$

$$\Rightarrow A = A v_1 \cdot v_1^T + \dots + A v_r \cdot v_r^T$$

What is r ? $\Rightarrow r = \text{Rank}(A)$

$\{v_1, \dots, v_n\}$: basis

$\{A v_1, \dots, A v_r\}$
orthogonal basis for $\text{Col}(A)$

$$\downarrow A$$

$$\{A v_1, A v_2, \dots, A v_r, 0, \dots, 0\}$$

\Rightarrow orthogonal \Rightarrow lin indep.

$$\begin{aligned} (A v_i) \cdot (A v_j) &= (A v_i)^T \cdot (A v_j) \\ &= v_i^T \cdot \boxed{A^T \cdot A \cdot v_j} \\ &= v_i^T \cdot (\lambda_j v_j) = \lambda_j (v_i \cdot v_j) \end{aligned}$$

If $i \neq j$, $A v_i \cdot A v_j = 0$

$$\left\{ \underbrace{\frac{1}{\sqrt{\lambda_1}} A v_1}_{u_1}, \underbrace{\frac{1}{\sqrt{\lambda_2}} A v_2}_{u_2}, \dots, \underbrace{\frac{1}{\sqrt{\lambda_r}} A v_r}_{u_r} \right\}$$

~~orthogonal~~ orthonormal basis for $\text{Col}(A)$

$$\|A v_i\|^2 = v_i^T \cdot \underline{A^T A} \cdot v_i = \lambda_i \frac{v_i^T \cdot v_i}{\|v_i\|^2} = \lambda_i$$

Def $\sigma_i = \sqrt{\lambda_i}$: the singular value for A .

$$A = \sigma_1 \frac{1}{\sigma_1} A v_1 \cdot v_1^T + \overset{\sigma_2 \cdot \frac{1}{\sigma_2}}{\sqrt{A} v_2} \cdot v_2^T + \dots + \overset{\sigma_r \cdot \frac{1}{\sigma_r}}{\sqrt{A} v_r} \cdot v_r^T$$

$$= \sigma_1 u_1 \cdot v_1^T + \sigma_2 u_2 \cdot v_2^T + \dots + \sigma_r u_r \cdot v_r^T$$

: Singular value decomposition.

$\{v_1, v_2, \dots, v_n\}$: ONB for \mathbb{R}^n

$\{v_1, \dots, v_r\}$: ONB for $\underline{\text{Nul}(A)^\perp}$

$\{u_1, u_2, \dots, u_r\}$: ONB for $\text{Col}(A) \subseteq \mathbb{R}^m$

$\{v_{r+1}, v_{r+2}, \dots, v_n\}$: ONB for $\text{Nul}(A)$

$\{u_{r+1}, u_{r+2}, \dots, u_m\}$: ONB for $\text{Col}(A)^\perp$
 $\text{Nul}(A^T)$

From Gram-Schmidt

\uparrow
 ~~n~~ m ?

$$\Rightarrow A = \begin{matrix} m \times m \\ [u_1 \dots u_m] \end{matrix} \begin{matrix} m \times n \\ \begin{bmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \sigma_r & 0 \\ 0 & & & \ddots \end{bmatrix} \end{matrix} \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$= U \cdot \Sigma \cdot V^T$$

U : $m \times m$ orthogonal matrix

V : $n \times n$ " "

$$\text{Col}(A) = \text{Span} \{ A \cdot e_1, A e_2, \dots, A e_n \} \quad \{v_1, \dots, v_n\}$$

$$= \text{Span} \{ A v_1, A v_2, \dots, A v_n \}$$

$$= \text{Span} \left\{ \frac{1}{\sigma_1} A v_1, \frac{1}{\sigma_2} A v_2, \dots, \frac{1}{\sigma_r} A v_r \right\}$$

orthonormal

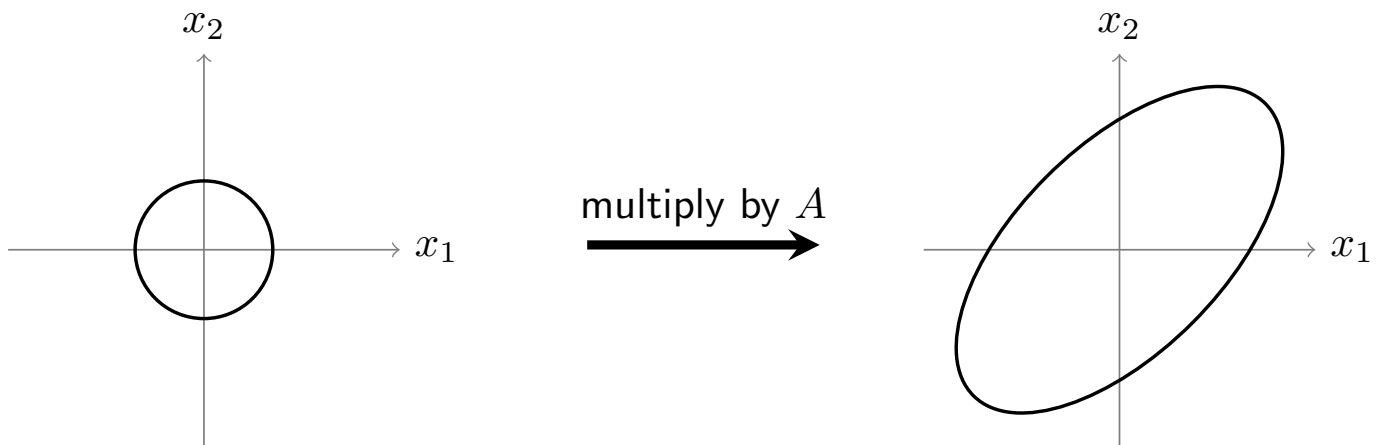
$$A e_1 = c_1 A v_1 + c_2 A v_2 + \dots + c_n A v_n$$

Example 1

The linear transform whose standard matrix is

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

maps the unit circle in \mathbb{R}^2 to an ellipse, as shown below. Identify the unit vector \vec{x} in which $\|A\vec{x}\|$ is maximized and compute this length.



Example 1 - Solution

Singular Values

positive semidefinite

The matrix $A^T A$ is always symmetric, with non-negative eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$. Let $\{\vec{v}_1, \dots, \vec{v}_n\}$ be the associated orthonormal eigenvectors. Then

$$\|A\vec{v}_j\|^2 = \lambda_j$$

If the A has rank r , then $\{A\vec{v}_1, \dots, A\vec{v}_r\}$ is an orthogonal basis for $\text{Col}A$:
For $1 \leq j < k \leq r$:

$$(A\vec{v}_j)^T A\vec{v}_k = 0 \quad j \neq k$$

Definition: $\sigma_1 = \sqrt{\lambda_1} \geq \sigma_2 = \sqrt{\lambda_2} \geq \dots \geq \sigma_n = \sqrt{\lambda_n}$ are the singular values of A .

$$\sigma_{r+1} = \dots = \sigma_n = 0$$

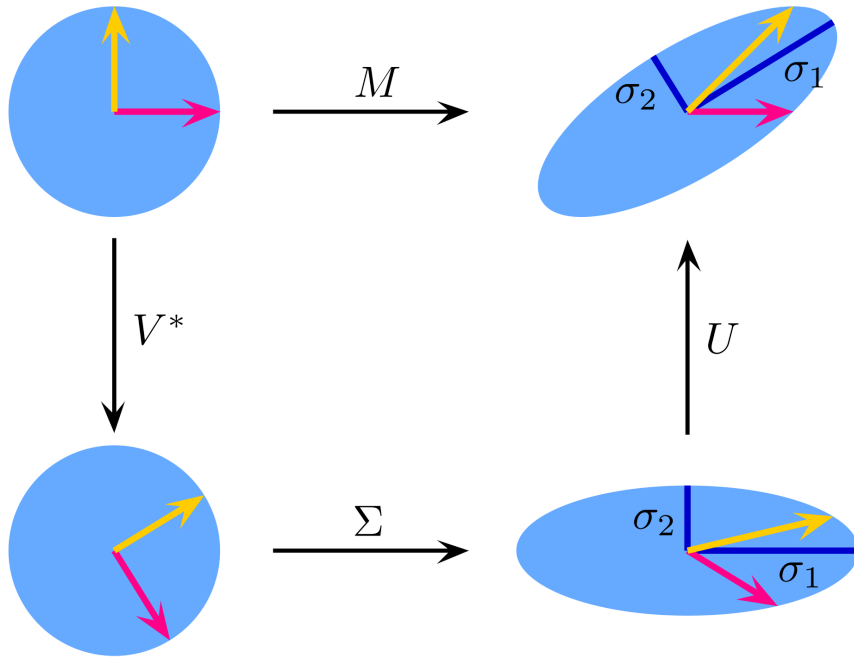
The SVD

Theorem: 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

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

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



$$M = U \cdot \Sigma \cdot V^T$$

Algorithm to find the SVD of A

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

1. Compute the squared singular values of $A^T A$, σ_i^2 , and construct Σ .
2. Compute the unit singular vectors of $A^T A$, \vec{v}_i , use them to form V .
3. Compute an orthonormal basis for $\text{Col}A$ using

$$\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i, \quad i = 1, 2, \dots, r$$

Extend the set $\{\vec{u}_i\}$ to form an orthonormal basis for \mathbb{R}^m , use the basis for form U .

11/22/24

Singular Value Decomposition.

A : $m \times n$ real

$A^T A$: $n \times n$ real symm. positive semidefinite.

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$$

$\{v_1, v_2, \dots, v_n\}$ eigenvectors for $A^T A$, ONB for \mathbb{R}^n

$$\text{Nul}(A) = \text{Nul}(A^T A) \Rightarrow Av_k = 0 \text{ for } k = r+1, \dots, n$$

$$A = \sigma_1 \boxed{\frac{1}{\sigma_1} A \cdot v_1} v_1^T + \sigma_2 \boxed{\frac{1}{\sigma_2} A v_2} v_2^T + \dots + \sigma_r \boxed{\frac{1}{\sigma_r} A v_r} v_r^T$$

$\sigma_i = \sqrt{\lambda_i}$: the singular values for A .

$$= \sigma_1 u_1 \cdot v_1^T + \dots + \sigma_r u_r v_r^T.$$

$\{u_1, \dots, u_r\}$: ONB for $\text{Col}(A)$

$\{u_{r+1}, \dots, u_m\}$: ONB for $\text{Col}(A)^\perp$

$\{v_{r+1}, \dots, v_n\}$: ONB for $\text{Nul}(A) = \text{Nul}(A^T A)$

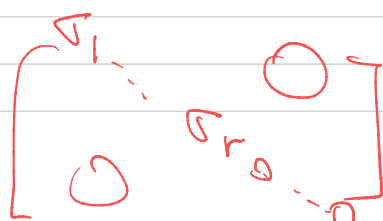
$\{v_1, \dots, v_r\}$: ONB for $\text{Nul}(A)^\perp$
= $\text{Col}(A^T)$

= $\text{Row}(A)$

$$A = \begin{matrix} m \times n \\ U \end{matrix} \cdot \begin{matrix} m \times n \\ \Sigma \end{matrix} \cdot \begin{matrix} n \times n \\ V^T \end{matrix}$$

$$[u_1, \dots, u_m]$$

$$[v_1, \dots, v_n]$$



Example 2: Write down the singular value decomposition for

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

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

$$\lambda_1 = 9 \quad \lambda_2 = 4$$

$$v_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad v_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$u_1 = \frac{1}{\sigma_1} \cdot A \cdot v_1 = \frac{1}{3} \cdot \begin{bmatrix} 0 \\ -3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} \quad \left. \begin{array}{l} \\ \\ \\ \end{array} \right\} \text{ONB for } \underline{\text{Col}(A)}$$

$$u_2 = \frac{1}{2} \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Section 7.4 Slide 40

$$\{u_3, u_4\} : \text{ONB for } \text{Col}(A)^\perp = \text{Nul}(A^T)$$

$$= \text{Nul} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ & & 1 & \uparrow \uparrow \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$= \sqrt{18} \cdot u_1 \cdot v_1^T$$

Example 3: Construct the singular value decomposition of

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 1/3 & & \\ -2/3 & & \\ 2/3 & & \end{bmatrix} \begin{bmatrix} \sqrt{18} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \leftarrow v^T$$

(It has rank 1.)

3×3 3×2 2×2

$$A^T A = \begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$$

$$\lambda^2 - 18\lambda + 0 = 0 \quad \lambda_1 = 18 \quad \lambda_2 = 0$$

$$= \lambda(\lambda - 18)$$

$$(r = 1 = \text{Rank}(A) = \text{Rank}(A^T A))$$

$$\text{Nul}(A - 18I) = \text{Nul} \begin{bmatrix} -9 & -9 \\ -9 & -9 \end{bmatrix} = \text{Nul} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

$$v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Section 7.4 Slide 41

$$\text{Nul}(A - 0I) = \text{Nul} \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix} = \text{Nul} \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}$$

$$v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{18}} \cdot \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} 2 \\ -4 \\ 4 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$$

$\{u_2, u_3\}$: ONB for $\text{Col}(A)^\perp = \text{Nul}(A^T)$

$$= \text{Nul} \begin{bmatrix} 1 & -2 & 2 \\ -1 & 2 & -2 \end{bmatrix}$$

$$= \text{Nul} \begin{bmatrix} 1 & -2 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix}$$

Basis, for

$$= x_2 \underbrace{\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}}_{y_1} + x_3 \underbrace{\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}}_{y_2}$$

$$x_1 = y_1 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

$$x_2 = y_2 - \frac{y_2 \cdot x_1}{x_1 \cdot x_1} x_1 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} - \frac{(-4)}{5} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

proj_{x₁}(y₂)

$$= \begin{bmatrix} -2 + \frac{8}{5} \\ \frac{4}{5} \\ 1 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} -2 \\ 4 \\ 5 \end{bmatrix}$$

$$u_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

$$u_3 = \frac{1}{\sqrt{45}} \begin{bmatrix} -2 \\ 4 \\ 5 \end{bmatrix}$$

Applications of the SVD

The SVD has been applied to many modern applications in CS, engineering, and mathematics (our textbook mentions the first four).

- Estimating the rank and condition number of a matrix
- Constructing bases for the four fundamental spaces
- Computing the pseudoinverse of a matrix
- Linear least squares problems
- Non-linear least-squares
https://en.wikipedia.org/wiki/Non-linear_least_squares
- Machine learning and data mining
<https://en.wikipedia.org/wiki/K-SVD>
- Facial recognition
<https://en.wikipedia.org/wiki/Eigenface>
- Principle component analysis
https://en.wikipedia.org/wiki/Principal_component_analysis
- Image compression

Students are expected to be familiar with the 1st two items in the list.

The Condition Number of a Matrix

$$\Rightarrow \text{Rank}(A) = n$$

If A is an invertible $n \times n$ matrix, the ratio

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

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$$

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.

$$V^* = \overline{V^T}$$

Example 4

For $A = U\Sigma V^T$, determine the rank of A , and orthonormal bases for $\text{Null}A$ and $(\text{Col}A)^\perp$.

$$U = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Col(A) is indicated by the first three columns (cyan), and Col(A)[⊥] is indicated by the fourth column (yellow).

$$\text{Rank}(A) = 3 = \text{Rank}(A^T \cdot A)$$

$$\Sigma = \begin{bmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$V^T = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sqrt{0.2} & 0 & 0 & 0 & \sqrt{0.8} \\ 0 & 0 & 0 & 1 & 0 \\ -\sqrt{0.8} & 0 & 0 & 0 & \sqrt{0.2} \end{bmatrix}$$

ONB for Null(A)[⊥] is indicated by the first three rows (cyan).
ONB for Null(A) is indicated by the last two rows (yellow).

Example 4 - Solution

The Four Fundamental Spaces

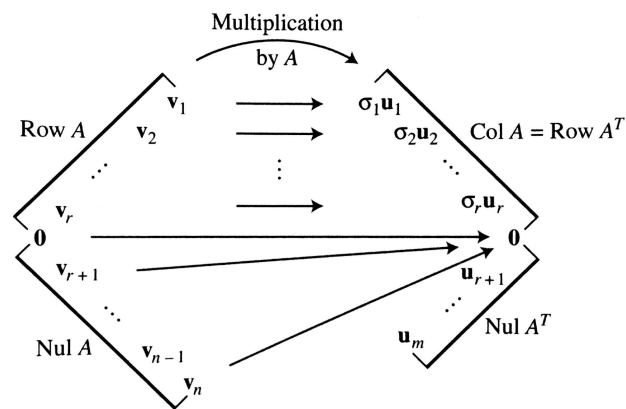


FIGURE 4 The four fundamental subspaces and the action of A .

1. $A\vec{v}_s = \sigma_s \vec{u}_s$.
2. $\vec{v}_1, \dots, \vec{v}_r$ is an orthonormal basis for $\text{Row } A$.
3. $\vec{u}_1, \dots, \vec{u}_r$ is an orthonormal basis for $\text{Col } A$.
4. $\vec{v}_{r+1}, \dots, \vec{v}_n$ is an orthonormal basis for $\text{Null } A$.
5. $\vec{u}_{r+1}, \dots, \vec{u}_m$ is an orthonormal basis for $\text{Null } A^T$.

The Spectral Decomposition of a Matrix

The SVD can also be used to construct the spectral decomposition for any matrix with rank r

$$A = \sum_{s=1}^r \sigma_s \vec{u}_s \vec{v}_s^T,$$

where \vec{u}_s, \vec{v}_s are the s^{th} columns of U and V respectively.

For the case when $A = A^T$, we obtain the same spectral decomposition that we encountered in Section 7.2.