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 $A^T = A$.

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

$$A = [*] \qquad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix}$$

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

$A^T A$ is Symmetric

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

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

$$= \underbrace{\begin{bmatrix} a_1^T a_1 & a_1^T a_2 & \cdots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \cdots & a_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^T a_1 & a_n^T a_2 & \cdots & a_n^T a_n \end{bmatrix}}_{\text{Entries are the dot products of columns of } A}$$

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:

Example 1

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

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

Hint: Gram-Schmidt

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 has the following properties.

1. All eigenvalues of A are real.
2. The dimenison of each eigenspace is full, that it's dimension is **equal to** it's algebraic multiplicity.
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):

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?

$$x^2 - 6xy + 9y^2 \geq 0$$

Quadratic Forms

Definition

A **quadratic form** is a function $Q : \mathbb{R}^n \to \mathbb{R}$, 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.

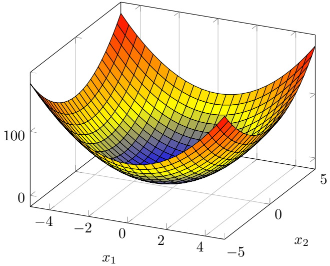
Example 1

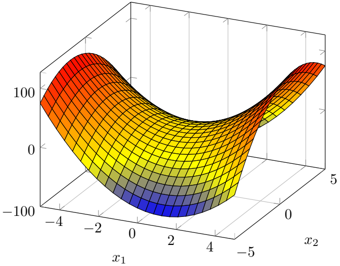
Compute the quadratic form $\vec{x}^T A\vec{x}$ for the matrices below.

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

Example 1 - Surface Plots

The surfaces for Example 1 are shown below.





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

Example 2

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

$$Q(x) = 5x_1^2 - x_2^2 + 3x_3^2 + 6x_1x_3 - 12x_2x_3$$

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:

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.

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

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

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

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

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

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

Principle Axes Theorem

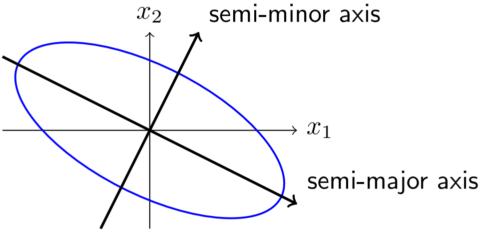
Theorem

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

Proof (if time permits):

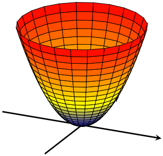
Example 5

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.

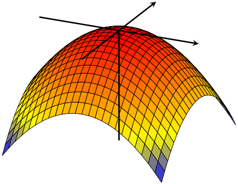


Classifying Quadratic Forms

$Q = x_1^2 + x_2^2$



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



Definition

A quadratic form Q is

1. **positive definite** if ____________ for all $\vec{x} \neq \vec{0}$.
2. **negative definite** if ____________ for all $\vec{x} \neq \vec{0}$.
3. **positive semidefinite** if ____________ for all $\vec{x}$.
4. **negative semidefinite** if ____________ for all $\vec{x}$.
5. **indefinite** if ________________________

Quadratic Forms and Eigenvalues

Theorem

If A is a ______________________ matrix with eigenvalues λ_i, then $Q = \vec{x}^T A \vec{x}$ is

1. **positive definite** iff λ_i ___________
2. **negative definite** iff λ_i ___________
3. **indefinite** iff λ_i ___________

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?

$$x^2 - 6xy + 9y^2 \geq 0$$

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.

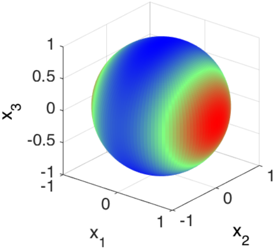
Example 1

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

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



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

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

$$\lambda_1 \geq \lambda_2 \ldots \geq \lambda_n$$

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

Proof:

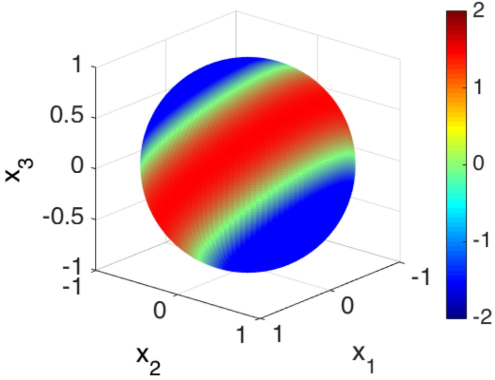
Example 2

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.

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

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

$$\lambda_1 \geq \lambda_2 \ldots \geq \lambda_n$$

and associated eigenvectors

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

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 $\vec{x} = \vec{u}_*$.
- 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.

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.

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

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.

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

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

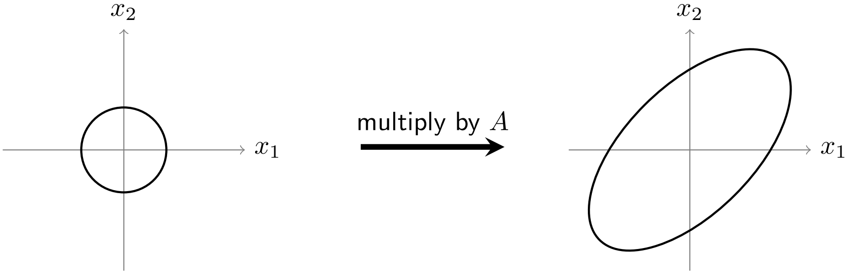
1. Compute the SVD for a rectangular matrix.
2. Apply the SVD to
   - estimate the rank and condition number of a matrix,
   - construct a basis for the four fundamental spaces of a matrix, and
   - construct a spectral decomposition of a matrix.

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

The matrix A^TA 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

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

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

$$(A\vec{v}_j)^T A\vec{v}_k =$$

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

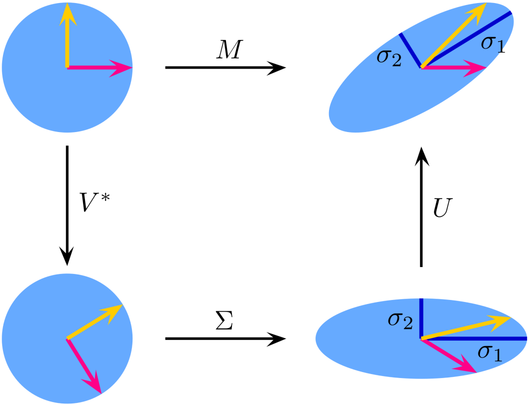
The SVD

Theorem: Singular Value Decomposition

A $m \times n$ matrix with rank r and non-zero singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$ has a decomposition $U\Sigma V^T$ where

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

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



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

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 Col$A$ using

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

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

Example 2: Write down the singular value decomposition for

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

Example 3: Construct the singular value decomposition of

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

(It has rank 1.)

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 1^{st} two items in the list.

The Condition Number of a Matrix

If A is an 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.

Example 4

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

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

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

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

Example 4 - Solution

The Four Fundamental Spaces

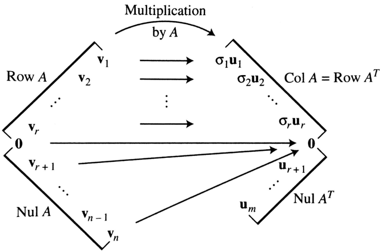


FIGURE 4 The four fundamental subspaces and the action of A.

1. $A\vec{v}_s = \sigma_s\vec{u}_s$.
2. $\vec{v}_1, \ldots, \vec{v}_r$ 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 Null$A$.
5. $\vec{u}_{r+1}, \ldots, \vec{u}_n$ is an orthonormal basis for NullA^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.