MATH254: Linear Algebra
Lecture 18

Moira MacNeil

February 25, 2025

Today

1. Orthogonal Complements and Projections

Reminders:

- Assignment 4 is posted and due Friday March 7

Lemma 8.1.1

Orthogonal Lemma

Let $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_m\}$ be an orthogonal set in $\mathbb{R}^n$. Given $\mathbf{x}$ in $\mathbb{R}^n$, write

$$\mathbf{f}_{m+1} = \mathbf{x} - \frac{\mathbf{x}\cdot\mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}\cdot\mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\mathbf{x}\cdot\mathbf{f}_m}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

Then:

1. $\mathbf{f}_{m+1}\cdot\mathbf{f}_k = 0$ for $k = 1, 2, \dots, m$. orthog. to all vect in orthog. set
2. If $\mathbf{x}$ is not in span$\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$, then $\mathbf{f}_{m+1} \neq \mathbf{0}$ and $\{\mathbf{f}_1, \dots, \mathbf{f}_m, \mathbf{f}_{m+1}\}$ is an orthogonal set.

Theorem 8.1.2

Theorem

Let U be a subspace of $\mathbb{R}^n$.

1. Every orthogonal subset $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ in U is a subset of an orthogonal basis of U.
2. U has an orthogonal basis.

Gram-Schmidt Orthogonalization Algorithm

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is any basis of a subspace U of $\mathbb{R}^n$, construct $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m$ in U successively as follows:

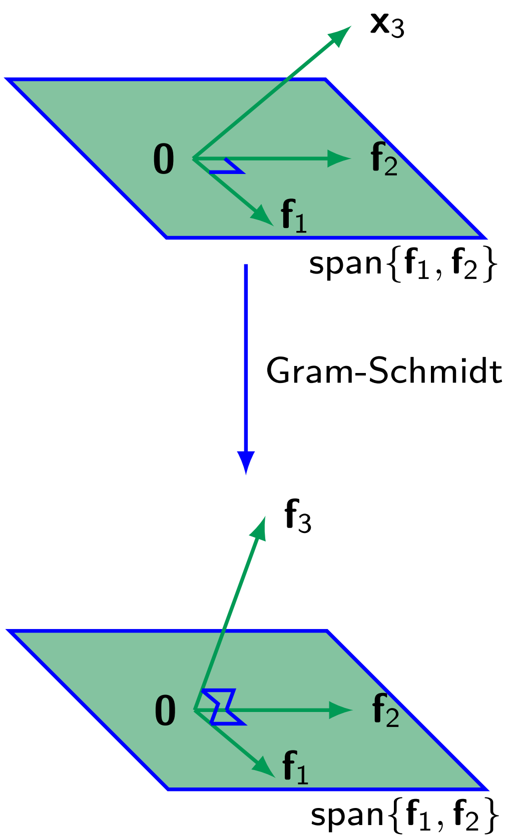
converts any basis to an orthog. basis

$$\begin{aligned}
\mathbf{f}_1 &= \mathbf{x}_1 \\
\mathbf{f}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 \\
\mathbf{f}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\
&\vdots \\
\mathbf{f}_k &= \mathbf{x}_k - \frac{\mathbf{x}_k \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \frac{\mathbf{x}_k \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \frac{\mathbf{x}_k \cdot \mathbf{f}_{k-1}}{\|\mathbf{f}_{k-1}\|^2}\mathbf{f}_{k-1}
\end{aligned}$$

for each $k = 2, 3, \ldots, m$. Then

1. $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is an orthogonal basis of U.
2. $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ for each $k = 1, 2, \ldots, m$.

A Remark on the Gram-Schmidt Orthogonalization Algorithm



- The vector $\frac{\mathbf{x}\cdot\mathbf{f}_i}{||\mathbf{f}_i||^2}\mathbf{f}_i$ is unchanged if a nonzero scalar multiple of $\mathbf{f}_i$ is used in place of $\mathbf{f}_i$
- If a newly constructed $\mathbf{f}_i$ is multiplied by a nonzero scalar at some stage of the Gram-Schmidt algorithm, the subsequent $\mathbf{f}$s will be unchanged
- Very useful in actual calculations

Example 8.1.4

Find an orthogonal basis of the row space of $A = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 3 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} x_1 \\ x_2 \\ x_3 \end{matrix}$

Observe $\{x_1, x_2, x_3\}$ is indep. (verify).

Use G.S.

$f_1 = x_1$

$f_2 = x_2 - \frac{x_2 \cdot f_1}{\|f_1\|^2} f_1 = (3,2,0,1) - \frac{4}{4}(1,1,-1,-1)$

$= (2, 2, -1, 2)$

$f_3 = x_3 - \frac{x_3 \cdot f_1}{\|f_1\|^2} f_1 - \frac{x_3 \cdot f_2}{\|f_2\|^2} f_2$

$$= (1,0,1,0) - \frac{0}{4} f_1 - \frac{1}{13}(2,2,-1,2)$$

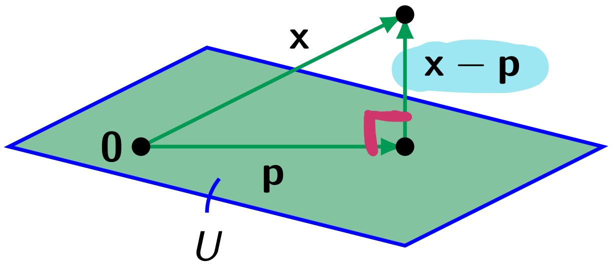
$$= (1,0,1,0) - \left(\frac{2}{13}, \frac{2}{13}, \frac{-1}{13}, \frac{2}{13}\right)$$

$$= \left(\frac{11}{13}, \frac{-2}{13}, \frac{14}{13}, \frac{11}{13}\right)$$

$$\Rightarrow (11, -2, 14, 11)$$

Point Closest to a Plane

- Given a point $\mathbf{x}$ and a plane U through the origin in $\mathbb{R}^3$, we want to find the point $\mathbf{p}$ in the plane that is closest to $\mathbf{x}$
- Geometrically, $\mathbf{p}$ must be chosen such that $\mathbf{x} - \mathbf{p}$ is perpendicular to the plane
- U is a subspace of $\mathbb{R}^3$ (it contains the origin)
- $\mathbf{x} - \mathbf{p}$ is perpendicular to U means that $\mathbf{x} - \mathbf{p}$ is orthogonal to every vector in U



Orthogonal Complement

Orthogonal Complement of a Subspace of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$, define the **orthogonal complement** $U^\perp$ of U (pronounced "U-perp") by

$$U^\perp = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{y} = 0 \text{ for all } \mathbf{y} \text{ in } U\}$$

The set of vectors that are orthogonal to every vector in U.

Lemma 8.1.6

Lemma

Let U be a subspace of $\mathbb{R}^n$.

1. $U^\perp$ is a subspace of $\mathbb{R}^n$.
2. $\{\mathbf{0}\}^\perp = \mathbb{R}^n$ and $(\mathbb{R}^n)^\perp = \{\mathbf{0}\}$.
3. If $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$, then $U^\perp = \{\underline{\mathbf{x} \text{ in } \mathbb{R}^n} \mid \underline{\mathbf{x} \cdot \mathbf{x}_i} = 0 \text{ for } i = 1, 2, \ldots, k\}$.

Example 8.1.7

Find $U^{\perp}$ if $U = \text{span}\{(1,-1,2,0),(1,0,-2,3)\}$ in $\mathbb{R}^4$.

$x_1 = (x, y, z, w) \in U^{\perp}$ iff it is orthog. to both $x_2 = (1,-1,2,0)$ and $x_3 = (1,0,-2,3)$

$$x_1 \cdot x_2 = x - y + 2z = 0$$
$$x_1 \cdot x_3 = x - 2z + 3w = 0$$

$$\begin{bmatrix} 1 & -1 & 2 & 0 \\ 1 & 0 & -2 & 3 \end{bmatrix} \to \begin{bmatrix} 1 & -1 & 2 & 0 \\ 0 & 1 & -4 & 3 \end{bmatrix}$$

let $z = s$
$w = t$

$y = 4s - 3t$
$x = -2s + 4s - 3t$

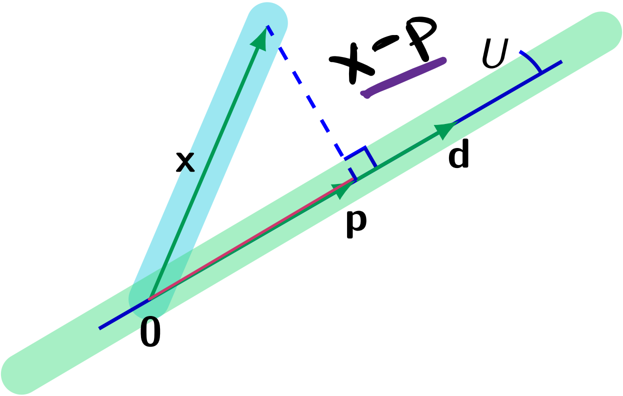
General Sol

$$s\begin{bmatrix}2\\4\\1\\0\end{bmatrix}+t\begin{bmatrix}-3\\-3\\0\\1\end{bmatrix}$$

$$U^{\perp}=\text{span}\{(2,4,1,0),(-3,-3,0,1)\}$$

Recall: Projection in $\mathbb{R}^3$

- Consider $\mathbf{x}$ and $\mathbf{d} \neq \mathbf{0}$ in $\mathbb{R}^3$
- The projection $\mathbf{p} = \text{proj}_{\mathbf{d}}\mathbf{x}$ of $\mathbf{x}$ on $\mathbf{d}$ is $\text{proj}_{\mathbf{d}}\mathbf{x} = \left(\frac{\mathbf{x}\cdot\mathbf{d}}{\|\mathbf{d}\|^2}\right)\mathbf{d}$ where $\mathbf{x} - \mathbf{p}$ is orthogonal to $\mathbf{d}$
- The line $U = \mathbb{R}\mathbf{d} = \{t\mathbf{d} \mid t \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$, $\{\mathbf{d}\}$ is an orthogonal basis of U, and $\mathbf{p} \in U$ and $\mathbf{x} - \mathbf{p} \in U^{\perp}$
- We can generalize this for any vector $\mathbf{x}$ in $\mathbb{R}^n$ and any subspace U of $\mathbb{R}^n$



Projection onto a Subspace

Projection onto a Subspace of $\mathbb{R}^n$

Let U be a subspace of $\mathbb{R}^n$ with orthogonal basis $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$. If $\mathbf{x}$ is in $\mathbb{R}^n$, the vector

$$\text{proj}_U \mathbf{x} = \frac{\mathbf{x} \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 + \frac{\mathbf{x} \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 + \cdots + \frac{\mathbf{x} \cdot \mathbf{f}_m}{\|\mathbf{f}_m\|^2}\mathbf{f}_m$$

is called the **orthogonal projection** of $\mathbf{x}$ on U. For the zero subspace $U = \{\mathbf{0}\}$, we define

$$\text{proj}_{\{\mathbf{0}\}} \mathbf{x} = \mathbf{0}$$

Theorem 8.1.9

Projection Theorem

If U is a subspace of $\mathbb{R}^n$ and $\mathbf{x}$ is in $\mathbb{R}^n$, write $\mathbf{p} = \text{proj}_U \mathbf{x}$. Then:

1. $\mathbf{p}$ is in U and $\mathbf{x} - \mathbf{p}$ is in $U^\perp$.
2. $\mathbf{p}$ is the vector in U closest to $\mathbf{x}$ in the sense that

$$\|\mathbf{x} - \mathbf{p}\| < \|\mathbf{x} - \mathbf{y}\| \quad \text{for all } \mathbf{y} \in U, \mathbf{y} \neq \mathbf{p}$$

Example 8.1.10

Let $U = \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}$ in $\mathbb{R}^4$ where $\mathbf{x}_1 = (1, 1, 0, 1)$ and $\mathbf{x}_2 = (0, 1, 1, 2)$. If $\mathbf{x} = (3, -1, 0, 2)$, find the vector in U closest to $\mathbf{x}$ and express $\mathbf{x}$ as the sum of a vector in U and a vector orthogonal to U.

$\{x_1, x_2\}$ indep. (verify), not orthog.

Use GS to find orthog. basis $\{f_1, f_2\}$

$f_1 = x_1 = (1,1,0,1)$

$$f_2 = x_2 - \frac{x_2 \cdot f_1}{\|f_1\|^2} f_1 = (0,1,1,2) - \frac{3}{3}(1,1,0,1)$$

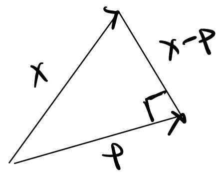
$$= (-1, 0, 1, 1)$$

Compute $\text{proj}_U x = p = \dfrac{x \cdot f_1}{\|f_1\|^2} f_1 + \dfrac{x \cdot f_2}{\|f_2\|^2} f_2$

p is pt in U closest to x

$$= \frac{4}{3} f_1 + \frac{-1}{3} f_2$$

$$= \tfrac{1}{3}(5, 4, -1, 3)$$



$$x - p = \tfrac{1}{3}(4, -7, 1, 3)$$

$$x = p + (x-p) = \tfrac{1}{3}(5, 4, -1, 3) + \tfrac{1}{3}(4, -7, 1, 3)$$

Example 8.1.11

Find the point in the plane with equation $2x + y - z = 0$ that is closest to the point $(2, -1, -3)$. $= X$

Plane is subspace U of $\mathbb{R}^3$ whose pts satisfy $z = 2x + y$

$$U = \{(s, t, 2s + t) \mid s, t \in \mathbb{R}\}$$
$$= \text{span}\{(0,1,1), (1,0,2)\}$$

Use GS to get orthog. basis

$$f_1 = (0,1,1)$$
$$f_2 = (1,0,2) - \frac{2}{2}(0,1,1) = (1,-1,1)$$

$$\text{proj}_U x = \frac{x \cdot f_1}{\|f_1\|^2} f_1 + \frac{x \cdot f_2}{\|f_2\|^2} f_2$$

$$= \frac{-4}{2} f_1 + 0 f_2$$

$$= (0, -2, -2) = \text{pt in } U \text{ closest to } x$$

Theorem 8.1.12

Theorem

Let U be a fixed subspace of $\mathbb{R}^n$. If we define $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(\mathbf{x}) = \text{proj}_U \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

1. T is a linear operator.
2. $\text{im}\, T = U$ and $\ker T = U^\perp$.
3. $\dim U + \dim U^\perp = n$.

Recap

Today we saw:

- Gram-Schmidt Orthogonalization Algorithm
- Orthogonal Complements
- Projections

Next time: Orthogonal Diagonalization

MATH254: Linear Algebra
Lecture 19

Moira MacNeil

February 26, 2025

Last Time

1. Gram-Schmidt Orthogonalization Algorithm
2. Orthogonal Complements
3. Projections

Today

1. Finish Orthogonal Complements and Projections
2. Orthogonal Diagonalization

Reminders:

- Assignment 4 is due Friday March 7

Recall: Orthogonal Complement

Orthogonal Complement of a Subspace of $\mathbb{R}^n$

If U is a subspace of $\mathbb{R}^n$, define the **orthogonal complement** $U^\perp$ of U (pronounced "U-perp") by

$$U^\perp = \{\mathbf{x} \text{ in } \mathbb{R}^n \mid \mathbf{x} \cdot \mathbf{y} = 0 \text{ for all } \mathbf{y} \text{ in } U\}$$

The set of vectors that are orthogonal to every vector in U.

Theorem 8.1.12

Theorem

Let U be a fixed subspace of $\mathbb{R}^n$. If we define $T : \mathbb{R}^n \to \mathbb{R}^n$ by

$$T(\mathbf{x}) = \text{proj}_U \mathbf{x} \quad \text{for all } \mathbf{x} \text{ in } \mathbb{R}^n$$

1. T is a linear operator.
2. $\text{im}\, T = U$ and $\ker T = U^\perp$.
3. $\dim U + \dim U^\perp = n$.

Proof of 1.

If $U = \{0\}$ then $U^\perp = \mathbb{R}^n$

$T(x) = \text{proj}_{\{0\}} x = 0$ for all x

$\therefore$ $T = 0$ so 1, 2, 3 hold

Assume $U \neq \{0\}$

1. If $\{f_1, f_2, \ldots, f_m\}$ is an ortho-normal basis of U, $\|f_i\|^2 = 1$

$$T(x) = \frac{(x \cdot f_1)}{\cancel{\|f_1\|^2}} f_1 + (x \cdot f_2) f_2 + \cdots + (x \cdot f_m) f_m$$

for all $x \in \mathbb{R}^n$

by definition of projection

let $x, y \in \mathbb{R}^n$ $r \in \mathbb{R}$

$(x+y) \cdot f_i = x \cdot f_i + y \cdot f_i$ for all $i = 1, \ldots, m$

so addition preserved

$(rx) \cdot f_i = r(x \cdot f_i)$ for all $i = 1, \ldots, m$

so scalar multiplication preserved.

Recall: Diagonalization

- An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors
- The matrix P with these eigenvectors as columns is a diagonalizing matrix for A, that is

$$D = P^{-1}AP \text{ is diagonal.}$$

- The nice bases of $\mathbb{R}^n$ are the orthogonal ones
- A natural question: which $n \times n$ matrices have an orthogonal basis of eigenvectors?
- The symmetric matrices $\rightarrow$ main result today

Recall: Orthonormal Vectors

- An orthogonal set of vectors is orthonormal if $\|\mathbf{v}\| = 1$ for each vector $\mathbf{v}$ in the set,
- An orthogonal set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ can be "normalized" to an orthonormal set $\{\frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1, \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2, \ldots, \frac{1}{\|\mathbf{v}_k\|}\mathbf{v}_k\}$
- If matrix A has n orthogonal eigenvectors, they can (by normalizing) be taken to be orthonormal
- The diagonalizing matrix P has orthonormal columns $\rightarrow$ these matrices are very easy to invert!

Theorem 8.2.1

Theorem

The following conditions are equivalent for an $n \times n$ matrix P.

1. P is invertible and $P^{-1} = P^T$.
2. The rows of P are orthonormal.
3. The columns of P are orthonormal.

Orthogonal Matrices

Orthogonal Matrices

An $n \times n$ matrix P is called an **orthogonal matrix** if it satisfies one (and hence all) of the conditions:

1. P is invertible and $P^{-1} = P^T$.
2. The rows of P are orthonormal.
3. The columns of P are orthonormal.

Given (2) and (3) orthonormal matrix might be a better name. But orthogonal matrix is standard.

Example 8.2.3

The rotation matrix $\begin{array}{cc} x_1 & x_2 \\ \multicolumn{2}{c}{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}} \end{array}$ is orthogonal for any angle θ.

3. $x_1 \cdot x_2 = -\cos\theta\sin\theta + \cos\theta\sin\theta = 0$

$\therefore$ orthogonal

$$\|x_1\| = \sqrt{\cos^2\theta + \sin^2\theta} = \|x_2\|$$

$$= \sqrt{1} = 1$$

Orthogonal matrices have many nice properties

- They are easy to invert $\rightarrow$ simply take the transpose
- If $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a linear operator, T is distance preserving if and only if its matrix is orthogonal
- The matrices of rotations and reflections about the origin in $\mathbb{R}^2$ and $\mathbb{R}^3$ are all orthogonal
- However, it is not enough that the rows of a matrix A are merely orthogonal for A to be an orthogonal matrix

orthonormal rows

Example 8.2.4

$x_2 \cdot x_3 = 1$

The matrix $\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{bmatrix} 2 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -1 & 1 \end{bmatrix} \end{array}$ has orthogonal rows but the columns are not orthogonal.

However, if the rows are normalized, the resulting matrix $\begin{bmatrix} \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ is orthogonal (so the columns are now orthonormal).

Example 8.2.5

If P and Q are orthogonal matrices, then PQ is also orthogonal, as is $P^{-1} = P^T$.

P and Q are invertible $(P^{-1} = P^T, Q^{-1} = Q^T)$
PQ is also invertible

$$(PQ)^{-1} = Q^{-1}P^{-1} = Q^T P^T = (PQ)^T$$

Therefore PQ is orthog.
(Pt 1. of Thm)

Orthogonally Diagonalizable Matrices

Orthogonally Diagonalizable Matrices

An $n \times n$ matrix A is said to be **orthogonally diagonalizable** when an orthogonal matrix P can be found such that $P^{-1}AP = P^TAP$ is diagonal.

Theorem 8.2.7

Principal Axes Theorem

The following conditions are equivalent for an $n \times n$ matrix A.

1. A has an orthonormal set of n eigenvectors.
2. A is orthogonally diagonalizable.
3. A is symmetric.

Proof $(2) \Rightarrow (3)$

Assume A ortnog. diag. then

$D = P^T A P$ is diag for some

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

Then $A = PDP^T$

But $D = D^T$

So $A^T = P^{TT} D^T P^T = PDP^T = A$

$\therefore$ A is symmetric

Principal Axes for a Matrix

- A set of orthonormal eigenvectors of a symmetric matrix A is called a set of **principal axes** for A
- Principal Axes Theorem is also called the **Real Spectral Theorem** because the eigenvalues of a (real) symmetric matrix are real,
- The set of distinct eigenvalues is called the **spectrum** of the matrix
- The spectral theorem is a more general result for matrices with complex entries

Example 8.2.8

Find an orthogonal matrix P such that $P^{-1}AP$ is diagonal, where $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ -1 & 2 & 5 \end{bmatrix}$.

$$c_A(x) = \det(xI - A)$$

$$= \begin{vmatrix} x-1 & 0 & 1 \\ 0 & x-1 & -2 \\ 1 & -2 & x-5 \end{vmatrix} = \begin{vmatrix} x-1 & 0 & 1 \\ 2(x-1) & x-1 & 0 \\ 1 & -2 & x-5 \end{vmatrix}$$

$$= (x-1)\left[(x-1)(x-5)\right] + -4(x-1) - (x-1)$$

$$= (x-1)\left[(x-1)(x-5) - 5\right]$$

$$= (x-1)(x^2 - 6x) = x(x-1)(x-6)$$

Eigenvalues $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 6$

$\lambda_1 = 0$ sols $(\lambda_1 I - A)x = 0$

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

$x_3 = t \quad x_2 = -2t \quad x_1 = t$

Eigenvector $\quad x_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$

$\lambda_2 = 1$

$x_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

$\lambda_3 = 6$

$x_3 = \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix}$

Orthogonal eigen vectors! (verify)

Normalize

$\|x_1\| = \sqrt{1^2 + (-2)^2 + 1^2} = \sqrt{6}$

$\|x_2\| = \sqrt{5}$

$\|x_3\| = \sqrt{30}$

Then $P = \left[\frac{1}{\sqrt{6}} x_1 \quad \frac{1}{\sqrt{5}} x_2 \quad \frac{1}{\sqrt{30}} x_3 \right]$

$$= \frac{1}{\sqrt{30}} \begin{bmatrix} \sqrt{5} & 2\sqrt{6} & -1 \\ -2\sqrt{5} & \sqrt{6} & 2 \\ \sqrt{5} & 0 & 5 \end{bmatrix}$$

is orthogonal

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

$$D = P^T A P = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

Recap

Today we saw:

- Orthogonally Diagonalizable Matrices
- Principal Axes Theorem

Next time: More on orthogonal diagonalization

MATH254: Linear Algebra
Lecture 20

Moira MacNeil

February 28, 2025

Last Time

1. Orthogonal Diagonalization

Today

1. Orthogonal Diagonalization continued

Reminders:

- Assignment 4 is due next Friday March 7

Recall: Gram-Schmidt Orthogonalization Algorithm

If $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is any basis of a subspace U of $\mathbb{R}^n$, construct $\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m$ in U successively as follows:

$$\begin{aligned}
\mathbf{f}_1 &= \mathbf{x}_1 \\
\mathbf{f}_2 &= \mathbf{x}_2 - \tfrac{\mathbf{x}_2 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 \\
\mathbf{f}_3 &= \mathbf{x}_3 - \tfrac{\mathbf{x}_3 \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \tfrac{\mathbf{x}_3 \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 \\
&\vdots \\
\mathbf{f}_k &= \mathbf{x}_k - \tfrac{\mathbf{x}_k \cdot \mathbf{f}_1}{\|\mathbf{f}_1\|^2}\mathbf{f}_1 - \tfrac{\mathbf{x}_k \cdot \mathbf{f}_2}{\|\mathbf{f}_2\|^2}\mathbf{f}_2 - \cdots - \tfrac{\mathbf{x}_k \cdot \mathbf{f}_{k-1}}{\|\mathbf{f}_{k-1}\|^2}\mathbf{f}_{k-1}
\end{aligned}$$

for each $k = 2, 3, \ldots, m$. Then

1. $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ is an orthogonal basis of U.
2. $\text{span}\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_k\} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ for each $k = 1, 2, \ldots, m$.

Recall: Orthogonal Matrices

Orthogonal Matrices

An $n \times n$ matrix P is called an **orthogonal matrix** if it satisfies one (and hence all) of the conditions:

1. P is invertible and $P^{-1} = P^T$.
2. The rows of P are orthonormal.
3. The columns of P are orthonormal.

Recall: Theorem 8.2.7

Principal Axes Theorem

The following conditions are equivalent for an $n \times n$ matrix A.

1. A has an orthonormal set of n eigenvectors.
2. A is orthogonally diagonalizable.
3. A is symmetric.

Theorem 8.2.9

Theorem

If A is an $n \times n$ symmetric matrix, then

$$(A\mathbf{x}) \cdot \mathbf{y} = \mathbf{x} \cdot (A\mathbf{y})$$

for all columns $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. converse is also true :)

Proof. Recall $x \cdot y = x^T y$ for all columns x, y

$A^T = A$

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

Theorem 8.2.10

Theorem

If A is a symmetric matrix, then eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof. Let $Ax = \lambda x$ $\quad Ay = \mu y$ where $\mu \neq \lambda$

$$\lambda(x \cdot y) = (\lambda x) \cdot y = (Ax) \cdot y = x \cdot (Ay) = x \cdot (\mu y)$$
$$= \mu (x \cdot y)$$

$$\lambda(x \cdot y) = \mu (x \cdot y) \rightarrow \underbrace{(\lambda - \mu)}_{\neq 0}(x \cdot y) = 0$$
$$\Rightarrow x \cdot y = 0$$

Diagonalizing a Symmetric Matrix

Procedure to diagonalize a symmetric $n \times n$ matrix:

1. Find the distinct eigenvalues
2. Find orthonormal bases for each eigenspace (using Gram-Schmidt algorithm if necessary)

Then the set of all these basis vectors is orthonormal and contains n vectors.

Example 8.2.11

Orthogonally diagonalize the symmetric matrix $A = \begin{bmatrix} 8 & -2 & 2 \\ -2 & 5 & 4 \\ 2 & 4 & 5 \end{bmatrix}$.

$$c_A(x) = \det(xI - A) = \begin{vmatrix} x-8 & 2 & -2 \\ 2 & x-5 & -4 \\ -2 & -4 & x-5 \end{vmatrix}$$

$$= \begin{vmatrix} x-8 & 2 & -2 \\ 0 & x-9 & x-9 \\ -2 & -4 & x-5 \end{vmatrix} = (x-8)\left[(x-9)(x-5) + 4(x-9)\right] - 2\left[2(x-9) + 2(x-9)\right]$$

$$= (x-9)\left[(x-8)(x-5)-4\right] = (x-9)\left[x^2-9x\right]$$
$$= x(x-9)^2$$

$\lambda_1 = 0$ $\quad\quad$ $\lambda_2 = 9$ (mult. $=2$)

solve $(\lambda_1 I - A)\,x_1 = 0$

$$\begin{bmatrix} -8 & 2 & -2 \\ 0 & -9 & -9 \\ -2 & -4 & -5 \end{bmatrix} \to \begin{bmatrix} 1 & -1/4 & 1/4 \\ 0 & -9 & -9 \\ 0 & -4.5 & -4.5 \end{bmatrix} \to$$

$$\begin{bmatrix} 1 & -1/4 & 1/4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$x_3 = t$
$x_2 = -t$
$x_1 = -1/2\,t$

$$x_1 = \begin{bmatrix} -1/2 \\ -1 \\ 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$$

$E_0(A) = \operatorname{span}\{x_1\}$

solve $(\lambda_2 I - A)\,x = 0$

$$\begin{bmatrix} 1 & 2 & -2 \\ 0 & 0 & 0 \\ -2 & -4 & 4 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$x_2 = s$
$x_3 = t$
$x_1 = -2s + 2t$

$x_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$ $\quad x_3 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$ $\quad x_2 \cdot x_3 = -4$

$\to$ NOT ORTHOG!

$E_q(A) = \text{span}\{x_2, x_3\}$

Use GS to find orthog. basis

$$f_1 = x_2$$
$$f_2 = x_3 - \frac{x_3 \cdot f_1}{\|f_1\|^2} f_1 = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

Normalize: $\left\{ \frac{1}{3} x_1, \frac{1}{\sqrt{5}} f_1, \frac{1}{3\sqrt{5}} f_2 \right\}$
(orthonormal eigenvectors)

$$P = \begin{bmatrix} \frac{1}{3} x_1 & \frac{1}{\sqrt{5}} f_1 & \frac{1}{3\sqrt{5}} f_2 \end{bmatrix}$$

$$= \frac{1}{3\sqrt{5}} \begin{bmatrix} \sqrt{5} & -6 & 2 \\ 2\sqrt{5} & 3 & 4 \\ -2\sqrt{5} & 0 & 5 \end{bmatrix}$$ is orthog. st $P^{-1}AP$ is diag.

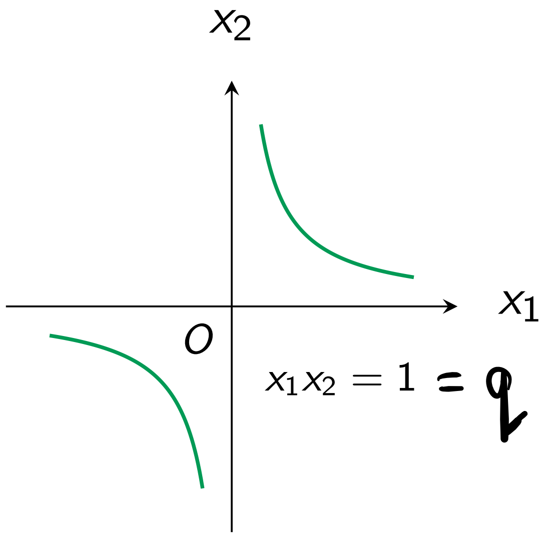
A nicer diag. matrix P

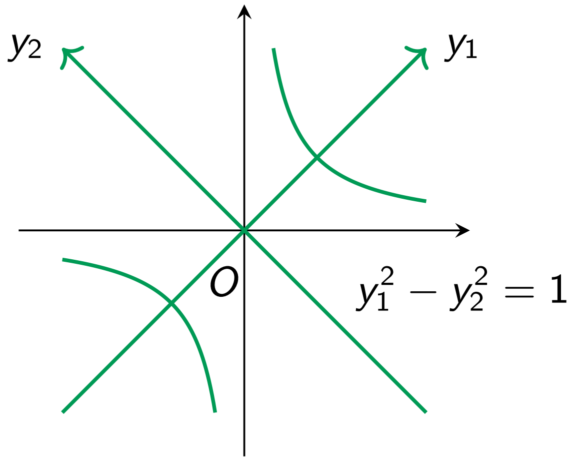
$y_2 = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ $\quad$ $y_3 \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$ are in $Eq(A)$ and orthog.

$Q = [\, 1/3\, x_1 \quad 1/3\, y_2 \quad 1/3\, y_3 \,]$ is also orthog. and $Q^{-1} A Q$ is diag.

Principle Axes and Quadratic Forms

- If A is symmetric and a set of orthogonal eigenvectors of A is given, the eigenvectors are called principal axes of A $\rightarrow$ from geometry
- An expression $q = ax_1^2 + bx_1x_2 + cx_2^2$ is called a **quadratic form** in the variables x_1 and x_2
- The graph of the equation $q = 1$ is called a **conic** in x_1 and x_2, e.g., $q = x_1x_2 = 1$
- Set $x_1 = y_1 + y_2$ and $x_2 = y_1 - y_2$ $\rightarrow$ $q = y_1^2 - y_2^2$ (no cross term!)
- The y_1 and y_2 axes are the principal axes for the conic





Example 8.2.12

Find principal axes for the quadratic form $q = x_1^2 - 4x_1x_2 + x_2^2$.

First express q as a matrix

$$q = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

NOT SYMMETRIC

$$q = x_1^2 - 2x_1x_2 - 2x_2x_1 + x_2^2$$

$$q = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x^T A x$$

SYMMETRIC $\to$ A

Eigenvalues of A are $\lambda_1 = 3$ $\lambda_2 = -1$

Eigenvectors (orthogonal) $x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ $x_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\|x_1\| = \|x_2\| = \sqrt{2}$ so

$P = 1/\sqrt{2} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ $D = \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} = P^T A P$

$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ where $y = P^T x$

$\Leftrightarrow$ $x = Py$ (since $P^{-1} = P^T$)

solve for ys

Thus $y_1 = \frac{1}{\sqrt{2}} (x_1 - x_2)$

$y_2 = \frac{1}{\sqrt{2}} (x_1 + x_2)$

Rewrite q in terms of y

$$q = x^T A x = (Py)^T A (Py)$$
$$= y^T P^T A P y$$
$$= y^T D y$$
$$= 3y_1^2 - y_2^2$$

Theorem 8.2.13

Triangulation Theorem

If A is an $n \times n$ matrix with n real eigenvalues, an orthogonal matrix P exists such that $P^T AP$ is upper triangular.

Corollary 8.2.14

Corollary

If A is an $n \times n$ matrix with real eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ (possibly not all distinct), then $\det A = \lambda_1 \lambda_2 \ldots \lambda_n$ and $\operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.

Recap

Today we saw:

- Orthogonal diagonalization
- Principal axes and quadratic forms

Next time: Positive Definite Matrices

MATH254: Linear Algebra
Lecture 21

Moira MacNeil

March 4, 2025

Last Time

1. Orthogonal Diagonalization

Today

1. Positive Definite Matrices
2. Introduction to Complex Matrices

Reminders:

- Assignment 4 is due on Friday (March 7)
- Midterm 2 is next Friday, March 14 → MIDTERM INFO POSTED

→ PRACTICE QUESTIONS COMING SOON
↳ FOR SEC 8.3 and 8.7

Positive Definite Matrices

Positive Definite Matrices

A square matrix is called **positive definite** if it is symmetric and all its eigenvalues λ are positive, that is $\lambda > 0$.

Theorem 8.3.2

Theorem

If A is positive definite, then it is invertible and $\det A > 0$.

Proof.

If A is $n \times n$ and has eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$ then $\det A = \lambda_1 \lambda_2 \cdots \lambda_n > 0$ by the principal axes thm and since A is PD.

Theorem 8.3.3

Theorem

A symmetric matrix A is positive definite if and only if $\underbrace{\mathbf{x}^T A\mathbf{x}}_{\text{quadratic forms}} > 0$ for every column $\mathbf{x} \neq \mathbf{0}$ in $\mathbb{R}^n$.

The positive definite matrices are the symmetric matrices A for which the quadratic form $q = \mathbf{x}^T A\mathbf{x}$ takes only positive values.

Example 8.3.4

(*) Every PD matrix can be factored as U^TU where U is upper Δ with pos on main diag.

If U is any invertible $n \times n$ matrix, show that $A = U^T U$ is positive definite. (The converse is also true!)

If $x \in \mathbb{R}^n$ where $x \neq 0$

$$x^TAx = x^T(U^TU)x = (x^TU^T)(Ux)$$
$$= (Ux)^TUx$$
$$= \|Ux\|^2 > 0$$

because $Ux \neq 0$ (U is invertible)

Principal Submatrices

- If A is any $n \times n$ matrix, let ${}^{(r)}A$ denote the $r \times r$ submatrix in the upper left corner of A
- ${}^{(r)}A$ is the matrix obtained from A by deleting the last $n - r$ rows and columns
- The matrices ${}^{(1)}A, {}^{(2)}A, {}^{(3)}A, \ldots, {}^{(n)}A = A$ are called the **principal submatrices** of A

Example 8.3.5

If $A = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$ then:

$^{(1)}A = [10]$ delete 3-1 row/cols

$^{(2)}A = \begin{bmatrix} 10 & 5 \\ 5 & 3 \end{bmatrix}$ delete 3-2 rows/cols

$^{(3)}A = A$ delete 0 rows cols

Lemma 8.3.6

Lemma

If A is positive definite, so is each principal submatrix ${}^{(r)}A$ for $r = 1, 2, \ldots, n$.

Proof. $A = \begin{bmatrix} {}^{(r)}A & P \\ Q & R \end{bmatrix}$ If $y \neq 0 \in \mathbb{R}^r$ let $x = \begin{bmatrix} y \\ 0 \end{bmatrix} \in \mathbb{R}^n$ $(x \neq 0)$

A is PD $\therefore$

$$0 < x^T A x = \begin{bmatrix} y^T & 0 \end{bmatrix} \begin{bmatrix} {}^{(r)}A & P \\ Q & R \end{bmatrix} \begin{bmatrix} y \\ 0 \end{bmatrix} = y^T\, {}^{(r)}A\, y$$

$\therefore$ ${}^{(r)}A$ is PD for any $r = 1, \ldots, n$ $\square$

Theorem 8.3.7

Theorem

The following conditions are equivalent for a symmetric $n \times n$ matrix A:

1. A is positive definite.
2. $\det({}^{(r)}A) > 0$ for each $r = 1, 2, \ldots, n$.
3. $A = U^T U$ where U is an upper triangular matrix with positive entries on the main diagonal.

The factorization in (3) is unique (called the **Cholesky factorization** of A).

Cholesky Factorization Algorithm

Algorithm for the Cholesky Factorization

If A is a positive definite matrix, the Cholesky factorization $A = U^T U$ can be obtained as follows:

Step 1. Carry A to an upper triangular matrix U_1 with positive diagonal entries using row operations each of which adds a multiple of a row to a lower row.

Step 2. Obtain U from U_1 by dividing each row of U_1 by the square root of the diagonal entry in that row.

Example

Find the Cholesky factorization of $A = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$.

Step 0: is A PD?

$\det({}^{(1)}A) = 10 > 0$

$\det({}^{(2)}A) = 5 > 0$

$\det({}^{(3)}A) = \det(A) = 3 > 0$

Step 1:

$$A = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix} \to \begin{bmatrix} 10 & 5 & 2 \\ 0 & 1/2 & 1 \\ 0 & 1 & 13/5 \end{bmatrix} \to \begin{bmatrix} 10 & 5 & 2 \\ 0 & 1/2 & 1 \\ 0 & 0 & 3/5 \end{bmatrix} = U_1$$

Step 2:

$$U = \begin{bmatrix} \sqrt{10} & 5/\sqrt{10} & 2/\sqrt{10} \\ 0 & 1/\sqrt{2} & \sqrt{2} \\ 0 & 0 & \sqrt{3}/\sqrt{5} \end{bmatrix}$$

$U^T U = A$

Complex Eigenvalues

- Until now, examples have been chosen so that the roots of characteristic polynomials are real
- This need not be true, even when the characteristic polynomial has real coefficients
- For example, if $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ then $c_A(x) = x^2 + 1$ has roots i and $-i$

$i^2 = -1$

Complex Vector Space

- $\mathbb{C}^n \to$ the set of all n-tuples of complex numbers
- We define vector operations on $\mathbb{C}^n$ as follows:

$$(v_1, v_2, \ldots, v_n) + (w_1, w_2, \ldots, w_n) = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$$
$$u(v_1, v_2, \ldots, v_n) = (uv_1, uv_2, \ldots, uv_n) \quad \text{for } u \text{ in } \mathbb{C}$$

- $\mathbb{C}^n$ satisfies the axioms for a vector space (with complex scalars)
- Thus we can have spanning sets, linearly independent subsets and bases for $\mathbb{C}^n$
- Definitions are identical to the real case, except that the scalars are allowed to be complex numbers.
- The standard basis of $\mathbb{R}^n$ remains a basis of $\mathbb{C}^n$, called the **standard basis** of $\mathbb{C}^n$

e_i

Complex Matrices

- $A = [a_{ij}]$ is called a **complex matrix** if every entry a_{ij} is a complex number
- If $z = a + bi$ is a complex number, the **conjugate** of z is the complex number:

$$\bar{z} = a - bi \quad \text{where } z = a + bi$$

- Define the **conjugate** of $A = [a_{ij}]$ to be the matrix

$$\overline{A} = \left[\ \overline{a}_{ij}\ \right]$$

obtained from A by conjugating every entry

- For all (complex) matrices of appropriate size:

$$\overline{A + B} = \overline{A} + \overline{B} \quad \text{and} \quad \overline{AB} = \overline{A}\,\overline{B}$$

Standard Inner Product

Standard Inner Product in ~~$\mathbb{R}^n$~~ $\mathbb{C}^n$

Given $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ in $\mathbb{C}^n$, define their **standard inner product** $\langle \mathbf{z}, \mathbf{w} \rangle$ by

$$\langle \mathbf{z}, \mathbf{w} \rangle = z_1\overline{w}_1 + z_2\overline{w}_2 + \cdots + z_n\overline{w}_n = \mathbf{z} \cdot \overline{\mathbf{w}}$$

where $\overline{w}$ is the conjugate of the complex number w.

Gives scalar in $\mathbb{C}^n$ and if $z, w \in \mathbb{R}^n$

then $\langle z, w \rangle = z \cdot w$

Example 8.7.2

$\overline{z} = (2, 1+i, -2i, 3+i)$ $\overline{w} = (1+i, -1, i, 3-2i)$

If $\mathbf{z} = (2, 1-i, 2i, 3-i)$ and $\mathbf{w} = (1-i, -1, -i, 3+2i)$, then

$$\langle \mathbf{z}, \mathbf{w} \rangle = 2(1+i) + (1-i)(-1) + (2i)(i) + (3-i)(3-2i) = 6-6i$$

$$\langle \mathbf{z}, \mathbf{z} \rangle = 2 \cdot 2 + (1-i)(1+i) + (2i)(-2i) + (3-i)(3+i) = 20$$

Theorem 8.7.3

Theorem

Let $\mathbf{z}$, $\mathbf{z}_1$, $\mathbf{w}$, and $\mathbf{w}_1$ denote vectors in $\mathbb{C}^n$, and let λ denote a complex number.

1. $\langle \mathbf{z}+\mathbf{z}_1, \mathbf{w}\rangle = \langle \mathbf{z}, \mathbf{w}\rangle + \langle \mathbf{z}_1, \mathbf{w}\rangle$ and $\langle \mathbf{z}, \mathbf{w}+\mathbf{w}_1\rangle = \langle \mathbf{z}, \mathbf{w}\rangle + \langle \mathbf{z}, \mathbf{w}_1\rangle$.
2. $\langle \lambda\mathbf{z}, \mathbf{w}\rangle = \lambda\langle \mathbf{z}, \mathbf{w}\rangle$ and $\langle \mathbf{z}, \lambda\mathbf{w}\rangle = \overline{\lambda}\langle \mathbf{z}, \mathbf{w}\rangle$.
3. $\langle \mathbf{z}, \mathbf{w}\rangle = \overline{\langle \mathbf{w}, \mathbf{z}\rangle}$.
4. $\langle \mathbf{z}, \mathbf{z}\rangle \geq 0$, and $\langle \mathbf{z}, \mathbf{z}\rangle = 0$ if and only if $\mathbf{z} = \mathbf{0}$.

PROOF.

(3) $z = (z_1, \cdots, z_n)$ $w = (w_1, \cdots, w_n) \in \mathbb{C}^n$

$$\overline{\langle w, z\rangle} = \overline{(w_1\overline{z_1} + \cdots + w_n\overline{z_n})} = \overline{w_1\overline{z_1}} + \cdots + \overline{w_n\overline{z_n}}$$
$$= z_1\overline{w_1} + \cdots + z_n\overline{w_n}$$
$$= \langle z, w\rangle$$

(4) let $w = z = (w_1, \cdots, w_n)$

$$\langle w, w\rangle = w_1\overline{w_1} + \cdots + w_n\overline{w_n}$$

$$\left[\begin{array}{l} |w_j| = \sqrt{a_j^2 + b_j^2} \geq 0 \text{ where } \quad w_j = a_j + b_j i \\ |w_j|^2 = w_j\overline{w_j} \geq 0 \end{array}\right]$$

$$\langle w, w\rangle = |w_1|^2 + \cdots + |w_n|^2$$

nonnegative real

equal to zero iff $w = 0$

This will give us vector norm in $\mathbb{C}^n$ $\square$

Recap

Today we saw:

- Positive Definite Matrices
- Cholesky Factorization
- Standard Inner Product

Next time: Complex Matrices

MATH254: Linear Algebra
Lecture 22

Moira MacNeil

March 5, 2025

Last Time

1. Positive Definite Matrices
2. Cholesky Factorization
3. Standard Inner Product

Today

1. More on Complex Matrices

Reminders:

- Assignment 4 is due on Friday (March 7)
- Midterm 2 is next Friday, March 14

Recall: Standard Inner Product

$z = a + bi \in \mathbb{C}$

$\bar{z} = a - bi$

Standard Inner Product in $\mathbb{C}^n$

Given $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ in $\mathbb{C}^n$, define their **standard inner product** $\langle \mathbf{z}, \mathbf{w} \rangle$ by

$$\langle \mathbf{z}, \mathbf{w} \rangle = z_1 \overline{w}_1 + z_2 \overline{w}_2 + \cdots + z_n \overline{w}_n = \underline{\mathbf{z}} \cdot \underline{\overline{\mathbf{w}}}$$

where $\overline{w}$ is the conjugate of the complex number w.

Norm and Length of Complex Vectors

Norm and Length in $\mathbb{C}^n$

As for the dot product on $\mathbb{R}^n$, property (4) enables us to define the **norm** or **length** $\|\mathbf{z}\|$ of a vector $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ in $\mathbb{C}^n$:

$$\|\mathbf{z}\| = \sqrt{\langle \mathbf{z}, \mathbf{z} \rangle} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}$$

For a complex number $z = a + bi$, we define $|z| = \sqrt{a^2 + b^2}$

Theorem 8.7.5

Theorem

If $\mathbf{z}$ is any vector in $\mathbb{C}^n$, then

1. $\|\mathbf{z}\| \geq 0$ and $\|\mathbf{z}\| = 0$ if and only if $\mathbf{z} = \mathbf{0}$.
2. $\|\lambda\mathbf{z}\| = |\lambda|\|\mathbf{z}\|$ for all complex numbers λ.

A vector $\mathbf{u}$ in $\mathbb{C}^n$ is called a **unit vector** if $\|\mathbf{u}\| = 1$.

Normalize $\frac{1}{\|v\|} v$

Example 8.7.6

$|z| = \sqrt{a^2 + b^2} \quad z = a + bi$

In $\mathbb{C}^4$, find a unit vector $\mathbf{u}$ that is a positive real multiple of $\mathbf{z} = (1 - i, i, 2, 3 + 4i)$.

$$\|z\| = \sqrt{\langle z, z\rangle} = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}$$

$$= \sqrt{(1^2 + 1^2) + (1^2) + 2^2 + (3^2 + 4^2)}$$

$$= \sqrt{32} = 4\sqrt{2}$$

$$u = \frac{1}{4\sqrt{2}} z$$

Conjugate Transpose

Conjugate Transpose in $\mathbb{C}^n$

The **conjugate transpose** A^H of a complex matrix A is defined by

$$A^H = (\overline{A})^T = \overline{(A^T)}$$

We have $A^H = A^T$ when A is real. (Other notations for A^H are A^* and $A^\dagger$.)

Example 8.7.8

$$\begin{bmatrix} 3 & 1-i & 2+i \\ 2i & 5+2i & -i \end{bmatrix}^H = \begin{bmatrix} 3 & -2i \\ 1+i & 5-2i \\ 2-i & i \end{bmatrix}$$

Theorem 8.7.9

Theorem

Let A and B denote complex matrices, and let λ be a complex number.

1. $(A^H)^H = A$.
2. $(A + B)^H = A^H + B^H$.
3. $(\lambda A)^H = \overline{\lambda} A^H$.
4. $(AB)^H = B^H A^H$.

Hermitian Matrices

Hermitian Matrices

A square complex matrix A is called **hermitian** if $A^H = A$, equivalently if $\overline{A} = A^T$.

Hermitian matrices have real entries on the main diagonal and the "reflection" of each nondiagonal entry must be the conjugate of that entry.

Example 8.7.11

Show $\begin{bmatrix} 3 & i & 2+i \\ -i & -2 & -7 \\ 2-i & -7 & 1 \end{bmatrix}$ is hermitian, whereas $\begin{bmatrix} 1 & i \\ i & -2 \end{bmatrix}$ and $\begin{bmatrix} 1 & i \\ -i & i \end{bmatrix}$ are not.

$$= \begin{bmatrix} 3 & i & 2+i \\ -i & -2 & -7 \\ 2-i & -7 & 1 \end{bmatrix}$$

$$\neq \begin{bmatrix} 1 & -i \\ -i & -2 \end{bmatrix}$$

$$\neq \begin{bmatrix} 1 & i \\ -i & -i \end{bmatrix}$$

Theorem 8.7.12

A sym. $\in \mathbb{R}$ $(Ax)\cdot y = x\cdot(Ay)$ for any $x, y \in \mathbb{R}^n$

Theorem

An $n \times n$ complex matrix A is hermitian if and only if

$$\langle A\mathbf{z}, \mathbf{w}\rangle = \langle \mathbf{z}, A\mathbf{w}\rangle$$

for all n-tuples $\mathbf{z}$ and $\mathbf{w}$ in $\mathbb{C}^n$.

Proof. $(\Rightarrow)$ A is hermitian $\underline{A^T = \overline{A}}$, if $z, w \in \mathbb{C}^n$

then $\langle z, w\rangle = z^T \overline{w}$

$\langle Az, w\rangle = (Az)^T \overline{w} = z^T \underline{A^T}\, \overline{w} = z^T \overline{A}\overline{w}$

$= z^{\dagger}\,\overline{(Aw)} = \langle z, Aw \rangle$

($\Leftarrow$) e_j is col j of identity matrix

$A = [a_{ij}]$ then condition gives

$\overline{a_{ij}} = \langle e_i, Ae_j \rangle \boxed{=} \langle Ae_i, e_j \rangle = a_{ij}$

Thus $\bar{A} = A^T$ $\therefore$ A hermitian.

Theorem 8.7.13

Theorem

Let A denote a hermitian matrix.

1. The eigenvalues of A are real.
2. Eigenvectors of A corresponding to distinct eigenvalues are orthogonal.

Proof. let λ, μ be eigenvalues of A with eigenvectors z, w (non zero) then $Az = \lambda z$ and $Aw = \mu w$

$$\lambda \langle z, w \rangle = \langle \lambda z, w \rangle = \langle Az, w \rangle = \langle z, Aw \rangle$$

$$= \langle z, \mu w \rangle = \bar{\mu} \langle z, w \rangle \quad (*)$$

(1) If $\lambda = \mu$ and $z = w$ $(*)$ becomes

$$\lambda \langle z, z \rangle = \bar{\lambda} \langle z, z \rangle$$

$$\langle z, z \rangle = \|z\|^2 \neq 0 \Rightarrow \underbrace{\lambda = \bar{\lambda}}_{\text{only true if } \lambda \text{ is real}}$$

similarly μ is real

$\square$

(2) Since μ, λ are real $(*)$ becomes

$$\lambda \langle z, w \rangle = \mu \langle z, w \rangle$$

if $\lambda \neq \mu \Rightarrow \langle z, w \rangle = 0$

$\therefore$ z, w are orthogonal. $\square$

Orthogonality in $\mathbb{C}^n$

Orthogonal and Orthonormal Vectors in $\mathbb{C}^n$

As in the real case, a set of nonzero vectors $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_m\}$ in $\mathbb{C}^n$ is called **orthogonal** if $\langle \mathbf{z}_i, \mathbf{z}_j \rangle = 0$ whenever $i \neq j$, and it is **orthonormal** if, in addition, $\|\mathbf{z}_i\| = 1$ for each i.

Theorem 8.7.15

Theorem

The following are equivalent for an $n \times n$ complex matrix A.

1. A is invertible and $A^{-1} = A^H$.
2. The rows of A are an orthonormal set in $\mathbb{C}^n$.
3. The columns of A are an orthonormal set in $\mathbb{C}^n$.

Unitary Matrices

Unitary Matrices

A square complex matrix U is called **unitary** if $U^{-1} = U^H$.

A real matrix is unitary if and only if it is orthogonal.

Example 8.7.17

The matrix $A = \begin{bmatrix} 1+i & 1 \\ 1-i & i \end{bmatrix}$ has orthogonal columns, but the rows are not orthogonal.

Normalizing the columns gives the unitary matrix $\frac{1}{2}\begin{bmatrix} 1+i & \sqrt{2} \\ 1-i & \sqrt{2}i \end{bmatrix}$.

Example 8.7.18

Consider the hermitian matrix $A = \begin{bmatrix} 3 & 2+i \\ 2-i & 7 \end{bmatrix}$. Find the eigenvalues of A, find two orthonormal eigenvectors, and so find a unitary matrix U such that $\underline{U^H A U}$ is diagonal.

$$c_A(x) = \det(xI - A) = \begin{vmatrix} x-3 & -2-i \\ -2+i & x-7 \end{vmatrix}$$

$$= (x-3)(x-7) - (-2+i)(-2-i)$$

$$= x^2 - 10x + 21 - [4 - i^2]$$

$$= x^2 - 10x + 16 = (x-2)(x-8)$$

So $\lambda_1 = 2 \qquad \lambda_2 = 8$

Eigenvectors: solve $(\lambda_i I - A)x = 0$

$$\begin{bmatrix} -1 & -2-i \\ -2+i & -5 \end{bmatrix} \to \begin{bmatrix} 1 & 2+i \\ 0 & 0 \end{bmatrix} \qquad x_1 = \begin{bmatrix} 2+i \\ -1 \end{bmatrix}$$

Add $-(-2+i)$ Row 1 to Row 2

$$-5-(-2-i)(-2+i) = -5-(4-i^2)$$
$$= -5+5 = 0$$

$$\begin{bmatrix} 5 & -2-i \\ -2+i & 1 \end{bmatrix} \to \begin{bmatrix} 1 & \frac{-2-i}{5} \\ 0 & 0 \end{bmatrix} \qquad \begin{array}{l} x_2 = t \\ x_1 = -\left(\frac{-2-i}{5}\right) \end{array}$$

$$x_2 = \begin{bmatrix} \frac{-2-i}{5} \\ 1 \end{bmatrix} = \begin{bmatrix} 2+i \\ 5 \end{bmatrix}$$

x_1, x_2 orthog. $\to$ normalize

$$\|x_1\| = \sqrt{2^2+1^2+1^2} = \sqrt{6}$$

$$\|x_2\| = \sqrt{(-2)^2+(-1)^2+5^2} = \sqrt{30}$$

$$U = \begin{bmatrix} \frac{1}{\sqrt{6}} x_1 & \frac{1}{\sqrt{30}} x_2 \end{bmatrix}$$

$$U^H A U = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}$$

Recap

Today we saw:

- Norm and Length in $\mathbb{C}^n$
- Conjugate Transpose
- Hermitian and Unitary Matrices

Next time: Unitary Diagonalization, Quadratic Forms

MATH254: Linear Algebra
Lecture 23

Moira MacNeil

March 7, 2025

Last Time

1. Norm and Length in $\mathbb{C}^n$
2. Conjugate Transpose A^H
3. Hermitian and Unitary Matrices

Today

1. Unitary Diagonalization
2. An Application to Quadratic Forms → NOT ON MIDTERM

Reminders:

- Midterm 2 is next Friday, March 14

Recall: Theorem 8.7.15

$A^H = (\bar{A})^T$

Theorem

The following are equivalent for an $n \times n$ complex matrix A.

1. A is invertible and $A^{-1} = A^H$.
2. The rows of A are an orthonormal set in $\mathbb{C}^n$.
3. The columns of A are an orthonormal set in $\mathbb{C}^n$.

Recall: Unitary Matrices

Unitary Matrices

A square complex matrix U is called **unitary** if $U^{-1} = U^H$.

A real matrix is unitary if and only if it is orthogonal.

Theorem 8.7.19

Schur's Theorem

If A is any $n \times n$ complex matrix, there exists a unitary matrix U such that

$$U^H A U = T$$

is upper triangular. Moreover, the entries on the main diagonal of T are the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of A (including multiplicities).

Corollary 8.7.20

Corollary

Let A be an $n \times n$ complex matrix, and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ denote the eigenvalues of A, including multiplicities. Then

$$\det A = \lambda_1 \lambda_2 \cdots \lambda_n \quad \text{and} \quad \operatorname{tr} A = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

Schur's theorem

- Schur's theorem asserts that every complex matrix can be "unitarily triangularized"
- This does not mean that that every complex matrix can be "unitarily diagonalized"
- For example, if $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, there is no invertible complex matrix U at all such that $U^{-1}AU$ is diagonal.

Theorem 8.7.21

Hermitian: $A^H = A$

Spectral Theorem

If A is hermitian, there is a unitary matrix U such that $U^H AU$ is diagonal.

A is called **unitarily diagonalizable** in this case.

Proof. By Schur's theorem, let $U^H A U = T$ where T is upper triangular and U is unitary.

$$\underset{LT}{T^H} = (U^H A U)^H = U^H A^H U^{HH} = U^H A U = \underset{UT}{T}$$

T is both upper and lower tri $\rightarrow$ diag. !

Example 8.7.22

Show that the non-hermitian matrix $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is unitarily diagonalizable.

A is non-herm. $A^H = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

$c_A(x) = \begin{vmatrix} x & -1 \\ 1 & x \end{vmatrix} = x^2+1 = (x-i)(x+i)$

$\lambda_1 = i \quad \lambda_2 = -i$

Eigenvectors

$\lambda_1 = i, \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \to \begin{bmatrix} i & -1 \\ 0 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}$ $x_2 = t$, $x_1 = -it$

$$x_1 = \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

$$\lambda_2 = -i, \quad \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \to \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix} \qquad \begin{aligned} x_2 &= t \\ x_1 &= it \end{aligned}$$

$$x_2 = \begin{bmatrix} i \\ 1 \end{bmatrix}$$

$$\begin{aligned} \langle x_1, x_2 \rangle &= x_1 \cdot \overline{x_2} \\ &= (-i)(-i) + 1(1) \\ &= i^2 + 1 = 0 \end{aligned}$$

this was the issue! we need to use inner prod since these vectors are in $\mathbb{C}^n$

A good reminder for all of us (especially me) :)

$$\|x_1\| = \|x_2\| = \sqrt{2}$$

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \text{ is unitary}$$

$$U^H A U = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \text{ is diagonal}$$

Normal Matrix

Normal Matrix

An $n \times n$ complex matrix N is called **normal** if $NN^H = N^H N$.

Every hermitian or unitary matrix is normal, as is, for example, the matrix $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ from the previous example.

Theorem 8.7.23

$A^H A = AA^H$

Theorem

An $n \times n$ complex matrix A is unitarily diagonalizable if and only if A is normal.

Proof. ($\Longrightarrow$) Assume $U^H A U = D$ where U unitary

D diagonal. D is normal $DD^H = D^H D$.

$DD^H = (U^H A U)(U^H A U)^H = U^H A (U U^H) A^H U$

$= U^H (A A^H) U$

$\updownarrow =$

$D^H D = U^H (A^H A) U$

$$U^H(AA^H)U = U^H(A^HA)U$$

$$AA^H = A^HA \Rightarrow A \text{ is normal.}$$

$\square$

Theorem 8.7.24

Cayley-Hamilton Theorem

If A is an $n \times n$ complex matrix, then $c_A(A) = 0$; that is, A is a root of its characteristic polynomial.

Example: Cayley-Hamilton Theorem

Let $A = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}$, find the characteristic polynomial $c_A(x)$, and compute $c_A(A)$.

$$c_A(x) = \det \begin{bmatrix} x-1 & -3 \\ 1 & x-2 \end{bmatrix} = (x-1)(x-2)+3$$
$$= x^2 - 3x + 5$$

$$c_A(A) = A^2 - 3A + 5I_2$$
$$= \begin{bmatrix} -2 & 9 \\ -3 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 9 \\ -3 & 6 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

MIDTERM TOPICS STOP HERE

Quadratic Forms

Quadratic Form

A **quadratic form** q in the n variables $x_1, x_2, \ldots, x_n$ is a linear combination of terms $x_1^2, x_2^2, \ldots, x_n^2$, and cross terms $x_1x_2, x_1x_3, x_2x_3, \ldots$.

$x_1 x_2$

$x_2 x_1 \quad \ldots$

Quadratic Forms as Matrix Products

- If $n = 3$, q has the form x_1, x_2, x_3

$$q = \underbrace{a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2} + \underbrace{a_{12}x_1x_2 + a_{21}x_2x_1 + a_{13}x_1x_3 + a_{31}x_3x_1 + a_{23}x_2x_3 + a_{32}x_3x_2}_{\text{cross terms}}$$

- In general, this can be written compactly as

$$q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} \qquad \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} A \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

- If $i \neq j$, the separate terms $a_{ij}x_ix_j$ and $a_{ji}x_jx_i$ can be replaced by

$$\frac{1}{2}(a_{ij} + a_{ji})x_ix_j \quad \text{and} \quad \frac{1}{2}(a_{ij} + a_{ji})x_jx_i$$

  without altering the quadratic form.

- May assume that x_ix_j and x_jx_i have the same coefficient in the sum for q
  $\to$ **assume that A is symmetric**.

Example 8.9.2

Write $q = x_1^2 + 3x_3^2 + 2x_1x_2 - x_1x_3$ in the form $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, where A is a symmetric 3×3 matrix.

$a_{12} = 2 \quad a_{21} = 0$

CROSS TERMS $\quad 2x_1x_2 = x_1x_2 + x_2x_1$

$a_{13} = -1 \quad a_{31} = 0$

$-x_1x_3 + 0x_3x_1 = -\tfrac{1}{2}x_1x_3 - \tfrac{1}{2}x_3x_1$

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

$$q(x) = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 1 & -1/2 \\ 1 & 0 & 0 \\ -1/2 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x^T A x$$

Change of Variables

Can always assume

- Given a quadratic form $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, where A is symmetric, the problem is to find new variables $y_1, y_2, \ldots, y_n$, related to $x_1, x_2, \ldots, x_n$, such that q has no cross terms $x_i x_j$
- We want to find $q = \mathbf{y}^T D\mathbf{y}$ where D is diagonal
- D is the matrix obtained when the symmetric matrix A is orthogonally diagonalized!
- By Principal Axes Theorem, we can find an orthogonal matrix P where $P^T AP = D$
- Define variables $\mathbf{y}$ by $\mathbf{x} = P\mathbf{y} \iff \mathbf{y} = P^T\mathbf{x}$ (these are the principal axes)
- Substituting into $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ gives

$$q = (P\mathbf{y})^T A(P\mathbf{y}) = \mathbf{y}^T(P^T AP)\mathbf{y} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

which has no cross terms

Theorem 8.9.3

Diagonalization Theorem

Let $q = \mathbf{x}^T A\mathbf{x}$ be a quadratic form in the variables $x_1, x_2, \dots, x_n$, where $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ and A is a symmetric $n \times n$ matrix. Let P be an orthogonal matrix such that $P^T AP$ is diagonal, and define new variables $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$ by

$$\mathbf{x} = P\mathbf{y} \quad \text{equivalently} \quad \mathbf{y} = P^T\mathbf{x}$$

If q is expressed in terms of these new variables $y_1, y_2, \dots, y_n$, the result is

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

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A repeated according to their multiplicities.

Recap

Today we saw:

- Unitary Diagonalization
- An Introduction to Quadratic Forms

Next time: More on Quadratic Forms

MATH254: Linear Algebra
Lecture 24

Moira MacNeil

March 11, 2025

Last Time

1. Unitary Diagonalization
2. Introduction to Quadratic Forms

Today

1. Continue with Application to Quadratic Forms

Reminders:

- Midterm 2 is on Friday, March 14

FORMULA SHEET – DIY :)

8.5 x 11 inches 1 sided

Recall: Quadratic Forms

Quadratic Form

A **quadratic form** q in the n variables $x_1, x_2, \ldots, x_n$ is a linear combination of terms $x_1^2, x_2^2, \ldots, x_n^2$, and cross terms $x_1x_2, x_1x_3, x_2x_3, \ldots$.

- Can be written compactly as

$$q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$$

- If $i \neq j$, the separate terms $a_{ij}x_ix_j$ and $a_{ji}x_jx_i$ can be replaced by

$$\frac{1}{2}(a_{ij} + a_{ji})x_ix_j \quad \text{and} \quad \frac{1}{2}(a_{ij} + a_{ji})x_jx_i$$

without altering the quadratic form $\implies$ A is symmetric

Theorem 8.9.3

Diagonalization Theorem

Let $q = \mathbf{x}^T A\mathbf{x}$ be a quadratic form in the variables $x_1, x_2, \dots, x_n$, where $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ and A is a symmetric $n \times n$ matrix. Let P be an orthogonal matrix such that $P^T AP$ is diagonal, and define new variables $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$ by

$$\mathbf{x} = P\mathbf{y} \quad \text{equivalently} \quad \mathbf{y} = P^T\mathbf{x}$$

If q is expressed in terms of these new variables $y_1, y_2, \dots, y_n$, the result is

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

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of A repeated according to their multiplicities.

Principal Axes

- Let $q = \mathbf{x}^T A\mathbf{x}$ be a quadratic form where A is a symmetric matrix and let $\lambda_1, \ldots, \lambda_n$ be the (real) eigenvalues of A repeated according to their multiplicities

- A corresponding set $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ of orthonormal eigenvectors for A is called a set of **principal axes** for the quadratic form q

- The orthogonal matrix P in Diagonalization Theorem is $P = \begin{bmatrix} \mathbf{f}_1 & \cdots & \mathbf{f}_n \end{bmatrix}$, so the variables X and Y are related by

$$\mathbf{x} = P\mathbf{y} = y_1\mathbf{f}_1 + y_2\mathbf{f}_2 + \cdots + y_n\mathbf{f}_n$$

Ch 5

- The coefficients are $y_i = \mathbf{x} \cdot \mathbf{f}_i$ by the expansion theorem. Hence q itself is easily computed:

$$q = q(\mathbf{x}) = \lambda_1(\mathbf{x} \cdot \mathbf{f}_1)^2 + \cdots + \lambda_n(\mathbf{x} \cdot \mathbf{f}_n)^2$$

Example 8.9.4

Find new variables y_1, y_2, y_3, and y_4 such that

$$q = 3(x_1^2 + x_2^2 + x_3^2 + x_4^2) + 2x_1x_2 - 10x_1x_3 + 10x_1x_4 + 10x_2x_3 - 10x_2x_4 + 2x_3x_4$$

has diagonal form, and find the corresponding principal axes.

$X = [x_1 \ x_2 \ x_3 \ x_4]^T$

$$A = \begin{bmatrix} 3 & 1 & -5 & 5 \\ 1 & 3 & 5 & -5 \\ -5 & 5 & 3 & 1 \\ 5 & -5 & 1 & 3 \end{bmatrix}$$

$c_A(x) = \det(xI - A)$

$= (x-12)(x+8)(x-4)^2$

Eigenvalues

$\lambda_1 = 12$, $\lambda_2 = -8$

$\lambda_3 = \lambda_4 = 4$

Orthonormal eigenvectors
(might involve using GS algo
to orthog. x_3, x_4)

$$f_1 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \quad f_2 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \quad f_3 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$f_4 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \quad P = \begin{bmatrix} f_1 & f_2 & f_3 & f_4 \end{bmatrix}$$

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

P is orthogonal.
$P^{-1}AP = P^TAP$ is diagonal

New vars y are related to
x vars
$$x = Py$$
$$y = P^Tx$$

$x \cdot f_1$

$$y_1 = \tfrac{1}{2}(x_1 - x_2 - x_3 + x_4)$$
$$y_2 = \tfrac{1}{2}(x_1 - x_2 + x_3 - x_4)$$
$$y_3 = \tfrac{1}{2}(x_1 + x_2 + x_3 + x_4)$$
$$y_4 = \tfrac{1}{2}(x_1 + x_2 - x_3 - x_4)$$

Substitute y's back into the $q(x)$

$$q = 12y_1^2 - 8y_2^2 + 4y_3^2 + 4y_4^2$$

$$= \lambda_1(x \cdot f_1)^2 + \lambda_2(x \cdot f_2)^2 + \lambda_3(x \cdot f_3)^2 + \lambda_4(x \cdot f_4)^2$$

Theorem 8.9.5

Theorem

If $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is a quadratic form given by a symmetric matrix A, then A is uniquely determined by q.

Proof. Let $q(x) = x^T B x$ and $q(x) = x^T A x$ for all x where $B^T = B$. If $C = A - B$, then $C^T = C$ $\quad$ $x^T C x = 0$ for all x,

since $x^T (A - B) x = (x^T A - x^T B) x = x^T A x - x^T B x$

$= q - q = 0$

Want to show $A=B \Leftrightarrow C=0$

Given $y \in \mathbb{R}^n$

$$0 = (x+y)^T C (x+y) = \overbrace{x^T C x}^{=0} + x^T C y + y^T C x + y^T C y = 0$$

$$= x^T C y + y^T C x$$

$$y^T C x = (x^T C y)^T = x^T C y \quad (1 \times 1)$$

$$0 = 2(x^T C y) \rightarrow x^T C y = 0$$

for all $x, y \in \mathbb{R}^n$

If e_j is col j of I_n

then $e_i^T C e_j = 0$

this is the (i,j)th entry of C

this holds for any (i,j) pair

so all of $C = 0$.

$\square$

Different Ways to Express the Same Quadratic Form

- A quadratic form q in variables x_i can be written in several ways . For example, if $q = 2x_1^2 - 4x_1x_2 + x_2^2$ then

$$q = 2(x_1 - x_2)^2 - x_2^2 \quad \text{and} \quad q = -2x_1^2 + (2x_1 - x_2)^2$$

- Let $q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, $\rightarrow$ new variables $\mathbf{y} = (y_1, y_2, \ldots, y_n)^T$ are linear combinations of x_i, then $\mathbf{y} = A\mathbf{x}$ for $n \times n$ matrix A
- Since we want to solve for the x_i in terms of y_i $\rightarrow$ matrix A invertible
- Hence suppose U is an invertible matrix and that the new variables $\mathbf{y}$ are given by

$$\mathbf{y} = U^{-1}\mathbf{x}, \quad \text{equivalently } \mathbf{x} = U\mathbf{y}$$

- Then $q = q(\mathbf{x}) = (U\mathbf{y})^T A(U\mathbf{y}) = \mathbf{y}^T \underline{(U^T AU)}\mathbf{y} \rightarrow q$ has matrix $U^T AU$ with respect to the new variables $\mathbf{y}$

Congruence

Congruent Matrices

Two $n \times n$ matrices A and B are called **congruent**, written $A \overset{c}{\sim} B$, if $B = U^T A U$ for some invertible matrix U.

1. $A \overset{c}{\sim} A$ *for all* A.
2. *If* $A \overset{c}{\sim} B$, *then* $B \overset{c}{\sim} A$.
3. *If* $A \overset{c}{\sim} B$ *and* $B \overset{c}{\sim} C$, *then* $A \overset{c}{\sim} C$.
4. *If* $A \overset{c}{\sim} B$, *then* A *is symmetric if and only if* B *is symmetric.*
5. *If* $A \overset{c}{\sim} B$, *then* $\mathrm{rank} A = \mathrm{rank} B$.

↳ converse is not always true

Example 8.9.6

Find the rank of the symmetric matrices $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Are they congruent?

$\operatorname{rank} A = 2 = \operatorname{rank} B$

A and B are not congruent

Assume $A \overset{c}{\sim} B$ then there is an invert. U such that $B = U^T A U = U^T I U = U^T U$

$\det B = -1 = (\det U)^2$

contradiction since $(\det U)^2 \geq 0$

Theorem 8.9.7

Sylvester's Law of Inertia

If $A \overset{c}{\sim} B$, then A and B have the same number of positive eigenvalues, counting multiplicities.

Complete Diagonalization

- The **index** of a symmetric matrix A is the number of positive eigenvalues of A
- If $q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is a quadratic form, the **index** and **rank** of q are defined to be, respectively, the index and rank of the matrix A. Hence the index and rank depend only on q and not on the way it is expressed.
- Now let $q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ be any quadratic form in n variables, of index k and rank r, where A is symmetric. We claim that new variables $\mathbf{z}$ can be found so that q is **completely diagonalized**—that is,

$$q(\mathbf{z}) = z_1^2 + \cdots + z_k^2 - z_{k+1}^2 - \cdots - z_r^2$$

$D_3(1,1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

- If $k \le r \le n$, let $D_n(k, r)$ denote the $n \times n$ diagonal matrix whose main diagonal consists of k ones, followed by $r - k$ minus ones, followed by $n - r$ zeros.

Complete Diagonalization continued

- We seek new variables $\mathbf{z}$ such that $q(\mathbf{z}) = \mathbf{z}^T D_n(k,r)\mathbf{z}$
- To determine $\mathbf{z}$: diagonalize A with an orthogonal matrix P_0 such that

$$P_0^T A P_0 = D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_r, 0, \ldots, 0)$$

- By reordering the columns of P_0, assume that $\lambda_1, \ldots, \lambda_k$ are positive and $\lambda_{k+1}, \ldots, \lambda_r$ are negative.
- Let $D_0 = \operatorname{diag}\left(\frac{1}{\sqrt{\lambda_1}}, \ldots, \frac{1}{\sqrt{\lambda_k}}, \frac{1}{\sqrt{-\lambda_{k+1}}}, \ldots, \frac{1}{\sqrt{-\lambda_r}}, 1, \ldots, 1\right)$

orthogonal columns

- Then $D_0^T D D_0 = D_n(k,r)$, so if $\mathbf{x} = (P_0 D_0)\mathbf{z}$, we obtain

$$q(\mathbf{z}) = \mathbf{z}^T D_n(k,r)\mathbf{z} = z_1^2 + \cdots + z_k^2 - z_{k+1}^2 - \cdots - z_r^2$$

Example 8.9.8

Completely diagonalize the quadratic form q in Example 8.9.4 and find the index and rank.

Eigenvals $12, -8, 4, 4$

index $A = 3$ (pos eigenvals) rank $A = 4$

Orthog. eigenvectors f_1, f_2, f_3, f_4 (see prev Eg)

$$P_0 = \begin{array}{cccc} 12 & 4 & 4 & -8 \\ \end{array}$$

$$P_0 = [f_1 \; f_3 \; f_4 \; f_2] \text{ orthogonal}$$

$$P_0^T A P_0 = \operatorname{diag}(12, 4, 4, -8)$$

$$D_0 = \operatorname{diag}\left(\frac{1}{\sqrt{12}}, \frac{1}{2}, \frac{1}{2}, -\frac{1}{\sqrt{8}}\right)$$

$$x = (D_0 P_0) z \quad \rightarrow \quad z = D_0^{-1} P_0^T x$$

$$z_1 = \sqrt{3}(x_1 - x_2 - x_3 + x_4)$$

$$z_2 = x_1 + x_2 + x_3 \; x_4$$

$$z_3 = x_1 + x_2 - x_3 - x_4$$

$$z_4 = \sqrt{2}(x_1 - x_2 + x_3 - x_4)$$

$$q = z_1^2 + z_2^2 + z_3^2 - z_4^2$$

Theorem 8.9.9

Theorem

Let A and B be symmetric $n \times n$ matrices, and let $0 \leq k \leq r \leq n$.

1. A has index k and rank r if and only if $A \overset{c}{\sim} D_n(k, r)$.
2. $A \overset{c}{\sim} B$ if and only if they have the same rank and index.

Proof.

Recap

Today we saw:

- Diagonalization to find change of variables for quadratic forms
- Congruence
- Complete diagonalization

Next time: Chapter 9 - Change of Basis