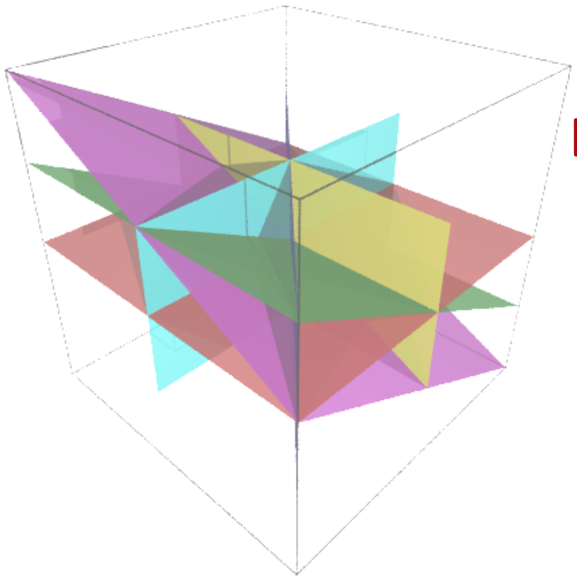

Math 201

Section F03

December 3, 2021

Spectral theorem

Spectral theorem. Let A be an $n \times n$ symmetric matrix over $\mathbb{R}$.

Spectral theorem

Spectral theorem. Let A be an $n \times n$ symmetric matrix over $\mathbb{R}$. Then A is diagonalizable over $\mathbb{R}$,

Spectral theorem

Spectral theorem. Let A be an $n \times n$ symmetric matrix over $\mathbb{R}$. Then A is diagonalizable over $\mathbb{R}$, and there exists an orthonormal basis for $\mathbb{R}^n$ (with respect to the standard inner product) consisting of eigenvectors for A.

Spectral theorem

Spectral theorem. Let A be an $n \times n$ symmetric matrix over $\mathbb{R}$. Then A is diagonalizable over $\mathbb{R}$, and there exists an orthonormal basis for $\mathbb{R}^n$ (with respect to the standard inner product) consisting of eigenvectors for A.

Let P be the matrix whose columns are these basis vectors. Then

$$P^{-1}AP = \text{diag}(\lambda_1, \ldots, \lambda_n).$$

Orthogonal matrix

Definition. A matrix $P \in M_{n\times n}(\mathbb{R})$ is *orthogonal* if its columns form an orthonormal set in $\mathbb{R}^n$.

Orthogonal matrix

Definition. A matrix $P \in M_{n \times n}(\mathbb{R})$ is *orthogonal* if its columns form an orthonormal set in $\mathbb{R}^n$.

Lemma. $P \in M_{n \times n}(\mathbb{R})$ is orthogonal if and only if $P^{-1} = P^t$.

Spectral theorem

Spectral theorem. If A is a real $n \times n$ symmetric matrix, then there exists a real diagonal matrix D and an orthogonal matrix P such that

$$P^t AP = D,$$

or equivalently,

$$A = PDP^t.$$

Proof of the spectral theorem

Step 1. The characteristic polynomial of A splits over $\mathbb{R}$ (and, thus, the eigenvalues of A are all real).

Proof of the spectral theorem

Step 1. The characteristic polynomial of A splits over $\mathbb{R}$ (and, thus, the eigenvalues of A are all real).

Proof. By the fundamental theorem of algebra, the characteristic polynomial splits over $\mathbb{C}$:

$$p_A(t) = \prod_{k=1}^{n} (\lambda_k - t)$$

with $\lambda_k \in \mathbb{C}$.

Proof of the spectral theorem

Step 1. The characteristic polynomial of A splits over $\mathbb{R}$ (and, thus, the eigenvalues of A are all real).

Proof. By the fundamental theorem of algebra, the characteristic polynomial splits over $\mathbb{C}$:

$$p_A(t) = \prod_{k=1}^{n} (\lambda_k - t)$$

with $\lambda_k \in \mathbb{C}$. We must show each $\lambda_k \in \mathbb{R}$.

Proof of the spectral theorem

Step 1. The characteristic polynomial of A splits over $\mathbb{R}$ (and, thus, the eigenvalues of A are all real).

Proof. By the fundamental theorem of algebra, the characteristic polynomial splits over $\mathbb{C}$:

$$p_A(t) = \prod_{k=1}^{n} (\lambda_k - t)$$

with $\lambda_k \in \mathbb{C}$. We must show each $\lambda_k \in \mathbb{R}$. Fix $\lambda = \lambda_k$ for some k. Take nonzero $v \in \mathbb{C}^n$ such that

$$Av = \lambda v.$$

Proof of the spectral theorem

Step 1. The characteristic polynomial of A splits over $\mathbb{R}$ (and, thus, the eigenvalues of A are all real).

Proof. By the fundamental theorem of algebra, the characteristic polynomial splits over $\mathbb{C}$:

$$p_A(t) = \prod_{k=1}^{n} (\lambda_k - t)$$

with $\lambda_k \in \mathbb{C}$. We must show each $\lambda_k \in \mathbb{R}$. Fix $\lambda = \lambda_k$ for some k. Take nonzero $v \in \mathbb{C}^n$ such that

$$Av = \lambda v.$$

miracle occurs here

Proof of the spectral theorem

Step 2. Induct on n.

Proof of the spectral theorem

Step 2. Induct on n. The case $n = 1$ is trivial.

Proof of the spectral theorem

Step 2. Induct on n. The case $n = 1$ is trivial. Assume $n > 1$, and take an eigenvalue-eigenvector pair $\lambda_1 \in \mathbb{R}$ and $v_1 \in \mathbb{R}^n$.

Proof of the spectral theorem

Step 2. Induct on n. The case $n = 1$ is trivial. Assume $n > 1$, and take an eigenvalue-eigenvector pair $\lambda_1 \in \mathbb{R}$ and $v_1 \in \mathbb{R}^n$. Complete to an ordered orthonormal basis for $\mathbb{R}^n$:

$$\langle v_1, \ldots, v_n \rangle.$$

Proof of the spectral theorem

Step 2. Induct on n. The case $n = 1$ is trivial. Assume $n > 1$, and take an eigenvalue-eigenvector pair $\lambda_1 \in \mathbb{R}$ and $v_1 \in \mathbb{R}^n$. Complete to an ordered orthonormal basis for $\mathbb{R}^n$:

$$\langle v_1, \ldots, v_n \rangle.$$

Let Q be the orthogonal matrix with columns $v_1, \ldots, v_n$,

Proof of the spectral theorem

Step 2. Induct on n. The case $n = 1$ is trivial. Assume $n > 1$, and take an eigenvalue-eigenvector pair $\lambda_1 \in \mathbb{R}$ and $v_1 \in \mathbb{R}^n$. Complete to an ordered orthonormal basis for $\mathbb{R}^n$:

$$\langle v_1, \ldots, v_n \rangle.$$

Let Q be the orthogonal matrix with columns $v_1, \ldots, v_n$, and define

$$\widetilde{A} = Q^{-1}AQ = Q^t AQ.$$

Proof of the spectral theorem

Step 2. Induct on n. The case $n = 1$ is trivial. Assume $n > 1$, and take an eigenvalue-eigenvector pair $\lambda_1 \in \mathbb{R}$ and $v_1 \in \mathbb{R}^n$. Complete to an ordered orthonormal basis for $\mathbb{R}^n$:

$$\langle v_1, \ldots, v_n \rangle.$$

Let Q be the orthogonal matrix with columns $v_1, \ldots, v_n$, and define

$$\widetilde{A} = Q^{-1} A Q = Q^t A Q.$$

Consider the structure of $\widetilde{A}$, and use induction.

Proof of the spectral theorem

- $\widetilde{A} = Q^t A Q$ is symmetric.

Proof of the spectral theorem

- $\widetilde{A} = Q^t A Q$ is symmetric.
- $$\widetilde{A} = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right)$$

Proof of the spectral theorem

- $\widetilde{A} = Q^t A Q$ is symmetric.
- $$\widetilde{A} = \left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right)$$
- By induction,

$$\widetilde{A} = \underbrace{\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & T & \\ 0 & & & \end{array}\right)}_{S} \underbrace{\left(\begin{array}{c|ccc} \lambda_1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & E & \\ 0 & & & \end{array}\right)}_{D} \underbrace{\left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & T^t & \\ 0 & & & \end{array}\right)}_{S^t}$$

Proof of the spectral theorem

We have $\widetilde{A} = Q^t AQ = SDS^t$ with Q and S orthogonal and D a real diagonal matrix.

Proof of the spectral theorem

We have $\widetilde{A} = Q^t A Q = S D S^t$ with Q and S orthogonal and D a real diagonal matrix.

Define $P = QS$.

Proof of the spectral theorem

We have $\widetilde{A} = Q^t A Q = SDS^t$ with Q and S orthogonal and D a real diagonal matrix.

Define $P = QS$. Then P is orthogonal, and

$$A = PDP^t.$$

□

Generalization

Definition. A matrix $A \in M_{n \times n}(\mathbb{C})$ is *Hermitian* if $A^* = A$ (so $A = \bar{A}^t$).

Generalization

Definition. A matrix $A \in M_{n\times n}(\mathbb{C})$ is *Hermitian* if $A^* = A$ (so $A = \bar{A}^t$).

A matrix $U \in M_{n\times n}(\mathbb{C})$ is *unitary* if its columns are orthonormal, or equivalently, if U is invertible with $U^{-1} = U^*$.

Generalization

Definition. A matrix $A \in M_{n\times n}(\mathbb{C})$ is *Hermitian* if $A^* = A$ (so $A = \bar{A}^t$).

A matrix $U \in M_{n\times n}(\mathbb{C})$ is *unitary* if its columns are orthonormal, or equivalently, if U is invertible with $U^{-1} = U^*$.

Theorem (Spectral theorem) Let A be an $n \times n$ Hermitian matrix. Then $A = UDU^*$ where U is unitary and D is a real diagonal matrix.

Applications

Statistics: least squares, singular value decomposition.

Applications

Statistics: least squares, singular value decomposition.

Multivariable calculus: optimization.

Applications

Statistics: least squares, singular value decomposition.

Multivariable calculus: optimization.

To find local minima and local maxima of $f : \mathbb{R}^n \to \mathbb{R}$, first set the derivative of f equal to zero:

Applications

Statistics: least squares, singular value decomposition.

Multivariable calculus: optimization.

To find local minima and local maxima of $f : \mathbb{R}^n \to \mathbb{R}$, first set the derivative of f equal to zero:

$$\frac{\partial f}{\partial x_1} = \cdots = \frac{\partial f}{\partial x_n} = 0$$

to find the critical points.

Applications

Statistics: least squares, singular value decomposition.

Multivariable calculus: optimization.

To find local minima and local maxima of $f : \mathbb{R}^n \to \mathbb{R}$, first set the derivative of f equal to zero:

$$\frac{\partial f}{\partial x_1} = \cdots = \frac{\partial f}{\partial x_n} = 0$$

to find the critical points.

Analyze these critical points: are they minima? maxima? neither?

Optimization

Let p be a critical point.

Optimization

Let p be a critical point. By applying a translation, we may assume that $p = 0 \in \mathbb{R}^n$.

Optimization

Let p be a critical point. By applying a translation, we may assume that $p = 0 \in \mathbb{R}^n$.

The Taylor series for f at p has no constant or linear term.

Optimization

Let p be a critical point. By applying a translation, we may assume that $p = 0 \in \mathbb{R}^n$.

The Taylor series for f at p has no constant or linear term. So for x near p, we have

$$f(x) \approx Q(x)$$

where Q consists of terms of degree 2.

Optimization

Let p be a critical point. By applying a translation, we may assume that $p = 0 \in \mathbb{R}^n$.

The Taylor series for f at p has no constant or linear term. So for x near p, we have

$$f(x) \approx Q(x)$$

where Q consists of terms of degree 2.

We then analyze Q using the spectral theorem.

Example

$$f(x, y) = x^2 - 3xy + y^3$$

Example

$$f(x, y) = x^2 - 3xy + y^3$$

$$\left.\begin{array}{rcrcl} \frac{\partial f}{\partial x} & = & 2x - 3y & = & 0 \\ \frac{\partial f}{\partial y} & = & -3x + 3y^2 & = & 0 \end{array}\right\}$$

Example

$$f(x,y) = x^2 - 3xy + y^3$$

$$\left.\begin{array}{rcccl} \frac{\partial f}{\partial x} & = & 2x - 3y & = & 0 \\ \frac{\partial f}{\partial y} & = & -3x + 3y^2 & = & 0 \end{array}\right\} \Rightarrow \begin{array}{rcl} y & = & \frac{2}{3}x \\ x & = & y^2 \end{array} \Bigg\}$$

Example

$$f(x,y) = x^2 - 3xy + y^3$$

$$\left.\begin{array}{rcl} \frac{\partial f}{\partial x} = & 2x - 3y & = 0 \\ \frac{\partial f}{\partial y} = & -3x + 3y^2 & = 0 \end{array}\right\} \Rightarrow \left.\begin{array}{rcl} y & = & \frac{2}{3}x \\ x & = & y^2 \end{array}\right\} \Rightarrow (x,y) = \begin{cases} (0,0) \\ \text{or} \\ \left(\frac{9}{4}, \frac{3}{2}\right) \end{cases}$$

Example

$$f(x,y) = x^2 - 3xy + y^3$$

$$\left.\begin{array}{rcrcl} \frac{\partial f}{\partial x} & = & 2x - 3y & = & 0 \\ \frac{\partial f}{\partial y} & = & -3x + 3y^2 & = & 0 \end{array}\right\} \Rightarrow \left.\begin{array}{rcl} y & = & \frac{2}{3}x \\ x & = & y^2 \end{array}\right\} \Rightarrow (x,y) = \begin{cases} (0,0) \\ \text{or} \\ \left(\frac{9}{4}, \frac{3}{2}\right) \end{cases}$$

Second order Taylor approximation at $p = \left(\frac{9}{4}, \frac{3}{2}\right)$:

$$f(x,y) \approx -\frac{27}{16} + \left(x - \frac{9}{4}\right)^2 - 3\left(x - \frac{9}{4}\right)\left(y - \frac{3}{2}\right) + \frac{9}{2}\left(y - \frac{3}{2}\right)^2$$

Example

$$f(x,y) = x^2 - 3xy + y^3$$

$$\left.\begin{array}{rcccl} \frac{\partial f}{\partial x} & = & 2x - 3y & = & 0 \\ \frac{\partial f}{\partial y} & = & -3x + 3y^2 & = & 0 \end{array}\right\} \Rightarrow \left.\begin{array}{rcl} y & = & \frac{2}{3}x \\ x & = & y^2 \end{array}\right\} \Rightarrow (x,y) = \begin{cases} (0,0) \\ \text{or} \\ \left(\frac{9}{4}, \frac{3}{2}\right) \end{cases}$$

Second order Taylor approximation at $p = \left(\frac{9}{4}, \frac{3}{2}\right)$:

$$f(x,y) \approx -\frac{27}{16} + \left(x - \frac{9}{4}\right)^2 - 3\left(x - \frac{9}{4}\right)\left(y - \frac{3}{2}\right) + \frac{9}{2}\left(y - \frac{3}{2}\right)^2$$

Translate and forget constant term:

$$Q(x,y) = x^2 - 3xy + \frac{9}{2}y^2$$

Example

$$Q(x, y) = x^2 - 3xy + \frac{9}{2}y^2$$

Example

$$Q(x,y) = x^2 - 3xy + \frac{9}{2}y^2 = \begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{3}{2} & \frac{9}{2} \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix}$$

Example

$$Q(x,y) = x^2 - 3xy + \frac{9}{2}y^2 = \begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{3}{2} & \frac{9}{2} \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$Q(x,y) = \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix}$$

Example

$$Q(x,y) = x^2 - 3xy + \frac{9}{2}y^2 = \begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{3}{2} & \frac{9}{2} \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$Q(x,y) = \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} P^t D P \begin{pmatrix} x \\ y \end{pmatrix}$$

Example

$$Q(x,y) = x^2 - 3xy + \frac{9}{2}y^2 = \begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{3}{2} & \frac{9}{2} \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$Q(x,y) = \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} P^t DP \begin{pmatrix} x \\ y \end{pmatrix}$$

Linear change of coordinates: $(u,v) = P\binom{x}{y}$:

$$Q(u,v) = (u,v)\, D \begin{pmatrix} u \\ v \end{pmatrix} = \lambda_1 u^2 + \lambda_2 v^2.$$

Example

$$Q(x,y) = x^2 - 3xy + \frac{9}{2}y^2 = \begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{3}{2} & \frac{9}{2} \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$Q(x,y) = \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} P^t DP \begin{pmatrix} x \\ y \end{pmatrix}$$

Linear change of coordinates: $(u,v) = P\binom{x}{y}$:

$$Q(u,v) = (u,v)\, D \begin{pmatrix} u \\ v \end{pmatrix} = \lambda_1 u^2 + \lambda_2 v^2.$$

$\lambda_1 = \frac{11-\sqrt{85}}{4} > 0, \lambda_2 = \frac{11+\sqrt{85}}{4} > 0$

Example

$$Q(x,y) = x^2 - 3xy + \frac{9}{2}y^2 = \begin{pmatrix} x & y \end{pmatrix} \underbrace{\begin{pmatrix} 1 & -\frac{3}{2} \\ -\frac{3}{2} & \frac{9}{2} \end{pmatrix}}_{A} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$Q(x,y) = \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x & y \end{pmatrix} P^t D P \begin{pmatrix} x \\ y \end{pmatrix}$$

Linear change of coordinates: $(u,v) = P\binom{x}{y}$:

$$Q(u,v) = (u,v)\, D \begin{pmatrix} u \\ v \end{pmatrix} = \lambda_1 u^2 + \lambda_2 v^2.$$

$\lambda_1 = \frac{11-\sqrt{85}}{4} > 0, \lambda_2 = \frac{11+\sqrt{85}}{4} > 0 \Rightarrow (\frac{9}{4}, \frac{3}{2})$ is a (local) minimum.