Rayleigh Quotients

February 19, 2025

- Quadratic forms

- Quadratic forms
- Positive (semi) definite matrices

- Quadratic forms
- Positive (semi) definite matrices
- Rayleigh quotients

- Quadratic forms
- Positive (semi) definite matrices
- Rayleigh quotients
- Generalized Rayleigh quotients

- Quadratic forms
- Positive (semi) definite matrices
- Rayleigh quotients
- Generalized Rayleigh quotients

Symmetric Matrices and their Properties

- A symmetric matrix is a square matrices A satisfying

$$A^T = A$$

.

Symmetric Matrices and their Properties

- A symmetric matrix is a square matrices A satisfying

$$A^T = A$$

- Real symmetric matrices $A \in \mathbb{R}_{n \times n}$ have many nice properties:

.

Symmetric Matrices and their Properties

- A symmetric matrix is a square matrices A satisfying

$$A^T = A$$

- Real symmetric matrices $A \in \mathbb{R}_{n \times n}$ have many nice properties:

  .

  - All their eigenvalues are real numbers (no complex eigenvalues)

Symmetric Matrices and their Properties

- A symmetric matrix is a square matrices A satisfying
$$A^T = A$$
- Real symmetric matrices $A \in \mathbb{R}_{n\times n}$ have many nice properties:
  .
  - All their eigenvalues are real numbers (no complex eigenvalues)
  - They are orthogonally diagonalizable, i.e., there exist an orthogonal matrix Q and a diagonal matrix Λ, both of the same size as A, such that
$$A = Q\Lambda Q^T$$
(This is known as the spectral decomposition of A.)

Symmetric Matrices and their Properties

- A symmetric matrix is a square matrices A satisfying

$$A^T = A$$

- Real symmetric matrices $A \in \mathbb{R}_{n \times n}$ have many nice properties:

.

  - All their eigenvalues are real numbers (no complex eigenvalues)
  - They are orthogonally diagonalizable, i.e., there exist an orthogonal matrix Q and a diagonal matrix Λ, both of the same size as A, such that

$$A = Q \Lambda Q^T$$

(This is known as the spectral decomposition of A.)

The matrix Λ consists of the eigenvalues of A along the diagonal, and Q has the corresponding orthonormal eigenvectors in its columns.

Quadratic Form

Definition Let $A \in \mathbb{R}_{n \times n}$ be a symmetric matrix. A quadratic form corresponding to A is a function $Q : \mathbb{R}^n \to \mathbb{R}$ with

Quadratic Form

Definition Let $A \in \mathbb{R}_{n\times n}$ be a symmetric matrix. A quadratic form corresponding to A is a function $Q : \mathbb{R}^n \to \mathbb{R}$ with

$$Q(x) = x^T Ax, \text{ for all } x \in \mathbb{R}^n.$$

Quadratic Form

Definition Let $A \in \mathbb{R}_{n\times n}$ be a symmetric matrix. A quadratic form corresponding to A is a function $Q : \mathbb{R}^n \to \mathbb{R}$ with

$$Q(x) = x^T Ax, \text{ for all } x \in \mathbb{R}^n.$$

Remark. A quadratic form is a polynomial with terms all of second order:

$$x^T Ax = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j.$$

Quadratic Form

Definition Let $A \in \mathbb{R}_{n \times n}$ be a symmetric matrix. A quadratic form corresponding to A is a function $Q : \mathbb{R}^n \to \mathbb{R}$ with

$$Q(x) = x^T A x, \text{ for all } x \in \mathbb{R}^n.$$

Remark. A quadratic form is a polynomial with terms all of second order:

$$x^T A x = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j.$$

Example. Let

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

Example. Let

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

then

$$Q(x) = x^T A x = x_1^2 + 2x_2^2 + 6x_1 x_2.$$

Definition A symmetric matrix $A \in \mathbb{R}_{n\times n}$ is said to be positive semidefinite if the corresponding quadratic form

$$Q(x) = x^T Ax \geq 0$$

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

Definition A symmetric matrix $A \in \mathbb{R}_{n \times n}$ is said to be positive semidefinite if the corresponding quadratic form

$$Q(x) = x^T A x \geq 0$$

for all $x \in \mathbb{R}^n$.
If the equality holds true only for $x = 0$ (i.e., $x^T A x > 0$ for all $x \neq 0$), then A is said to be positive definite.

Theorem A symmetric matrix is positive definite (semidefinite) if and only if all of its eigenvalues are positive (nonnegative).
Example Let

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 3 \\ 3 & 2 \end{pmatrix}.$$

According to the theorem, A is positive definite, B is positive semidefinite, and C is neither.

Problem: Let $A \in \mathbb{R}_{n \times n}$ be a positive semi definite matrix. Find another matrix B of the same size such that $A = B^2$. We call B the square root of A and denote it by $B = A^{1/2}$.

Solution: Since A is symmetric and positive semi definite matrix, there exists an orthogonal matrix $Q \in \mathbb{R}_{n\times n}$ and a diagonal matrix $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ with all $\lambda_i \geq 0$ such that $A = Q\Lambda Q^T$.

Solution: Since A is symmetric and positive semi definite matrix, there exists an orthogonal matrix $Q \in \mathbb{R}_{n\times n}$ and a diagonal matrix $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ with all $\lambda_i \geq 0$ such that $A = Q\Lambda Q^T$. Define $\Lambda^{1/2} = \text{diag}(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2})$. Clearly, $\Lambda^{1/2}\Lambda^{1/2} = \Lambda$.

Solution: Since A is symmetric and positive semi definite matrix, there exists an orthogonal matrix $Q \in \mathbb{R}_{n\times n}$ and a diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with all $\lambda_i \geq 0$ such that $A = Q\Lambda Q^T$.
Define $\Lambda^{1/2} = \mathrm{diag}(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2})$. Clearly, $\Lambda^{1/2}\Lambda^{1/2} = \Lambda$.
Let $B = Q\Lambda^{1/2}Q^T$. Then

$$B^2 = (Q\Lambda^{1/2}Q^T)(Q\Lambda^{1/2}Q^T) = Q\Lambda^{1/2}\Lambda^{1/2}Q^T = Q\Lambda Q^T = A.$$

Solution: Since A is symmetric and positive semi definite matrix, there exists an orthogonal matrix $Q \in \mathbb{R}_{n\times n}$ and a diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ with all $\lambda_i \geq 0$ such that $A = Q\Lambda Q^T$.
Define $\Lambda^{1/2} = \mathrm{diag}(\lambda_1^{1/2}, \ldots, \lambda_n^{1/2})$. Clearly, $\Lambda^{1/2}\Lambda^{1/2} = \Lambda$.
Let $B = Q\Lambda^{1/2}Q^T$. Then

$$B^2 = (Q\Lambda^{1/2}Q^T)(Q\Lambda^{1/2}Q^T) = Q\Lambda^{1/2}\Lambda^{1/2}Q^T = Q\Lambda Q^T = A.$$

Thus, $B = A^{1/2} = Q\Lambda^{1/2}Q^T$ is still a positive semi definite matrix matrix.

Example Let

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix},$$

which is positive semi definite matrix because it has two nonnegative eigenvalues $\lambda_1 = 5, \lambda_2 = 0$. To find the matrix square root of A, we need to find its orthogonal diagonalization.

Example Let

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix},$$

which is positive semi definite matrix because it has two nonnegative eigenvalues $\lambda_1 = 5, \lambda_2 = 0$. To find the matrix square root of A, we need to find its orthogonal diagonalization.

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}^T.$$

Example Let

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix},$$

which is positive semi definite matrix because it has two nonnegative eigenvalues $\lambda_1 = 5, \lambda_2 = 0$. To find the matrix square root of A, we need to find its orthogonal diagonalization.

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}^T.$$

It follows that

$$A^{1/2} = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{5} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}^T.$$

Rayleigh Quotient Rayleigh quotients are encountered in many statistical and machine learning problems. It is thus necessary to study them systematically.

Rayleigh Quotient Rayleigh quotients are encountered in many statistical and machine learning problems. It is thus necessary to study them systematically.

Rayleigh quotients have many applications:

- PCA: $\max_{v \neq 0} \dfrac{v^T \Sigma v}{v^T v}$ (where Σ is a covariance matrix)
- LDA: $\max_{v \neq 0} \dfrac{v^T S_b v}{v^T S_w v}$ (where S_b is the between-class scatter matrix and S_w is the within-class scatter matrix)
- Spectral clustering: $\max_{v \neq 0} \dfrac{v^T L v}{v^T D v}$ (where L is the graph Laplacian and D is the degree matrix)

Definition The Rayleigh quotient for a given symmetric matrix A is a multivariate function $f : \mathbb{R}^n - \{0\} \to \mathbb{R}$ defined by

Definition The Rayleigh quotient for a given symmetric matrix A is a multivariate function $f : \mathbb{R}^n - \{0\} \to \mathbb{R}$ defined by

$$f(x) = \frac{x^T A x}{x^T x}, \quad x \neq 0.$$

Remark. A Rayleigh quotient is always scaling invariant. That is, for any nonzero vector $x \in \mathbb{R}^n$,

$$f(kx) = \frac{(kx)^T A(kx)}{(kx)^T(kx)} = \frac{x^T Ax}{x^T x} = f(x), \quad \text{for all } k \neq 0.$$

Remark. A Rayleigh quotient is always scaling invariant. That is, for any nonzero vector $x \in \mathbb{R}^n$,

$$f(kx) = \frac{(kx)^T A(kx)}{(kx)^T(kx)} = \frac{x^T Ax}{x^T x} = f(x), \quad \text{for all } k \neq 0.$$

Another way to see this is by rewriting the Rayleigh quotient as follows:

$$f(x) = \frac{x^T Ax}{x^T x} = \frac{x^T Ax}{\|x\|^2} = \left(\frac{x}{\|x\|}\right)^T A \left(\frac{x}{\|x\|}\right), \quad x \neq 0.$$

Remark. A Rayleigh quotient is always scaling invariant. That is, for any nonzero vector $x \in \mathbb{R}^n$,

$$f(kx) = \frac{(kx)^T A(kx)}{(kx)^T(kx)} = \frac{x^T Ax}{x^T x} = f(x), \quad \text{for all } k \neq 0.$$

Another way to see this is by rewriting the Rayleigh quotient as follows:

$$f(x) = \frac{x^T Ax}{x^T x} = \frac{x^T Ax}{\|x\|^2} = \left(\frac{x}{\|x\|}\right)^T A \left(\frac{x}{\|x\|}\right), \quad x \neq 0.$$

Interpretation: The Rayleigh quotient is essentially a quadratic form over the unit sphere.

Example The Rayleigh quotient for

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

Example The Rayleigh quotient for

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

is

$$f(x) = \frac{x^T A x}{x^T x} = \frac{x_1^2 + 2x_2^2 + 6x_1 x_2}{x_1^2 + x_2^2}, \quad x \neq 0.$$

Example The Rayleigh quotient for

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

is

$$f(x) = \frac{x^T A x}{x^T x} = \frac{x_1^2 + 2x_2^2 + 6x_1 x_2}{x_1^2 + x_2^2}, \quad x \neq 0.$$

It is a function defined over $\mathbb{R}^2$ with the origin excluded.

Problem: Given symmetric matrix A, find the maximum (or minimum) of the associated Rayleigh quotient

$$\max_{x \neq 0 \in \mathbb{R}^n} \frac{x^T A x}{x^T x}.$$

Problem: Given symmetric matrix A, find the maximum (or minimum) of the associated Rayleigh quotient

$$\max_{x \neq 0 \in \mathbb{R}^n} \frac{x^T A x}{x^T x}.$$

Equivalent formulations:

$$\max_{x \in \mathbb{R}^n, \|x\|=1} x^T A x,$$

$$\max_{x \in \mathbb{R}^n} x^T A x \quad \text{subject to} \quad \|x\|^2 = 1.$$

Theorem For any given symmetric matrix $A \in \mathbb{R}_{n \times n}$,

$$\max_{x \in \mathbb{R}^n, x \neq 0} \frac{x^T A x}{x^T x} = \lambda_{\max},$$

Theorem For any given symmetric matrix $A \in \mathbb{R}_{n \times n}$,

$$\max_{x \in \mathbb{R}^n, x \neq 0} \frac{x^T A x}{x^T x} = \lambda_{\max},$$

$$\min_{x \in \mathbb{R}^n, x \neq 0} \frac{x^T A x}{x^T x} = \lambda_{\min}.$$

x is the eigenvector associated with the largest and smallest eigenvalue of A respectively.

Proof: Let $A = V \Lambda V^T$ be the spectral decomposition, where $V = [v_1, \ldots, v_n]$ is orthogonal and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is diagonal with sorted diagonals from large to small. Then for any unit vector x,

Proof: Let $A = V\Lambda V^T$ be the spectral decomposition, where $V = [v_1, \ldots, v_n]$ is orthogonal and $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ is diagonal with sorted diagonals from large to small. Then for any unit vector x,

$$x^T Ax = x^T(V\Lambda V^T)x = (x^T V)\Lambda(V^T x) = y^T \Lambda y,$$

where $y = V^T x$ is also a unit vector

Proof: Let $A = V\Lambda V^T$ be the spectral decomposition, where $V = [v_1, \ldots, v_n]$ is orthogonal and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is diagonal with sorted diagonals from large to small. Then for any unit vector x,

$$x^T Ax = x^T(V\Lambda V^T)x = (x^T V)\Lambda(V^T x) = y^T \Lambda y,$$

where $y = V^T x$ is also a unit vector

$$\|y\|^2 = y^T y = (V^T x)^T(V^T x) = x^T VV^T x = x^T x = 1.$$

Proof: Let $A = V\Lambda V^T$ be the spectral decomposition, where $V = [v_1, \ldots, v_n]$ is orthogonal and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ is diagonal with sorted diagonals from large to small. Then for any unit vector x,

$$x^T A x = x^T (V\Lambda V^T) x = (x^T V)\Lambda(V^T x) = y^T \Lambda y,$$

where $y = V^T x$ is also a unit vector

$$\|y\|^2 = y^T y = (V^T x)^T (V^T x) = x^T V V^T x = x^T x = 1.$$

So the original optimization problem becomes:

$$\max_{y \in \mathbb{R}^n, \|y\|=1} y^T \Lambda y.$$

$$y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2 \quad \text{subject to} \quad y_1^2 + y_2^2 + \cdots + y_n^2 = 1.$$

$$y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2 \quad \text{subject to} \quad y_1^2 + y_2^2 + \cdots + y_n^2 = 1.$$

Then clearly

$$\sum_{i=1}^{n} \lambda_i |y_i|^2 \leq \lambda_n \sum_{i=1}^{n} |y_i|^2 = \lambda_n \|y\|^2$$

$$y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2 \quad \text{subject to} \quad y_1^2 + y_2^2 + \cdots + y_n^2 = 1.$$

Then clearly

$$\sum_{i=1}^{n} \lambda_i |y_i|^2 \leq \lambda_n \sum_{i=1}^{n} |y_i|^2 = \lambda_n \|y\|^2$$

and

$$\sum_{i=1}^{n} \lambda_i |y_i|^2 \geq \lambda_1 \sum_{i=1}^{n} |y_i|^2 = \lambda_1 \|y\|^2$$

$$y^T \Lambda y = \sum_{i=1}^{n} \lambda_i y_i^2 \quad \text{subject to} \quad y_1^2 + y_2^2 + \cdots + y_n^2 = 1.$$

Then clearly

$$\sum_{i=1}^{n} \lambda_i |y_i|^2 \leq \lambda_n \sum_{i=1}^{n} |y_i|^2 = \lambda_n \|y\|^2$$

and

$$\sum_{i=1}^{n} \lambda_i |y_i|^2 \geq \lambda_1 \sum_{i=1}^{n} |y_i|^2 = \lambda_1 \|y\|^2$$

therefore

$$\lambda_1 \leq R(x) \leq \lambda_n \quad \text{for all } x \neq 0.$$

Example For the semi positive definite matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix},$$

Example For the semi positive definite matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix},$$

we have previously obtained its eigenvalues and eigenvectors:

$$\lambda_1 = 5, \lambda_2 = 0; v_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix}, v_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

The associated Rayleigh quotient

$$Q(x) = \frac{x_1^2 + 4x_2^2 + 4x_1x_2}{x_1^2 + x_2^2}$$

The associated Rayleigh quotient

$$Q(x) = \frac{x_1^2 + 4x_2^2 + 4x_1x_2}{x_1^2 + x_2^2}$$

has the following extreme values:

- The maximum value of $Q(x)$ is $\lambda_1 = 5$, achieved at $x = \pm v_1$;
- The minimum is $\lambda_2 = 0$, achieved at $x = \pm v_2$.

The overall range of the Rayleigh quotient is thus $[0, 5]$.