Chapter 6 : Symmetric Matrices and Quadratic Forms

ENSIA, May 2024

Contents

Symmetric Matrices

Definition 1

A symmetric matrix is a (square) matrix A such that ${}^tA = A$.

Example 1

The matrices $\begin{pmatrix} 1 & 3 \\ 3 & 0 \end{pmatrix}$ and $\begin{pmatrix} -1 & 2 & 0 \\ 2 & 5 & 4 \\ 0 & 4 & 3 \end{pmatrix}$ are symmetric.

The matrices $\begin{pmatrix} 1 & -3 \\ 3 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$ and $\begin{pmatrix} -1 & 2 & 0 \\ 2 & 5 & 4 \\ 0 & -4 & 3 \end{pmatrix}$ are not symmetric.

Orthogonality of Eigenvectors

Theorem 1

If A is symmetric, then any two eigenvectors from different eigenspaces of A are orthogonal.

Example 2

The matrix

$$A = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

has two distinct eigenvalues $\lambda_1 = -2$ and $\lambda_2 = 0$ with respective corresponding eigenvectors of the form $v_1 = (\alpha, -\alpha)$ and $v_2 = (\beta, \beta)$. Obviously, v_1 and v_2 are orthogonal.

Orthogonally Diagonalizable Matrix

Definition 2

An $n \times n$ matrix A is said to be orthogonally diagonalizable if there are an orthogonal matrix P (with $P^{-1} = {}^{t}P$) and a diagonal matrix D such that

$$A = PD\,{}^{t}P = PDP^{-1} \qquad (1)$$

Theorem 2

An $n \times n$ matrix A is orthogonally diagonalizable if and only if A is a symmetric matrix.

The Spectral Theorem

Theorem 3 (The Spectral Theorem for Symmetric Matrices)

An $n \times n$ symmetric matrix A has the following properties :

a) A has n real eigenvalues, counting multiplicities.

b) The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic polynomial.

c) The eigenspaces are mutually orthogonal, that is, the eigenvectors corresponding to different eigenvalues are orthogonal.

d) A is orthogonally diagonalizable.

Spectral Decomposition

Proposition 1

Suppose $A = PDP^{-1}$, where the columns of P are orthonormal eigenvectors $u_1, \cdots, u_n$ of A and the corresponding eigenvalues $\lambda_1, \cdots, \lambda_n$ are in the diagonal matrix D. Then we have

$$A = \lambda_1 u_1{}^t u_1 + \cdots + \lambda_n u_n{}^t u_n. \qquad (2)$$

Definition 3

1) The decomposition (2) of A is called a spectral decomposition of A.
2) Each matrix $u_j{}^t u_j$ is called a projection matrix.

Remark 1

For all x in $\mathbb{R}^n$, the vector $\left(u_j{}^t u_j\right)x$ is the orthogonal projection of x onto the subspace spanned by u_j.

Example

Example 4

The matrix

$$A = \begin{pmatrix} 7 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \begin{pmatrix} 8 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix}$$

has the spectral decomposition

$$A = 8u_1{}^t u_1 + 3u_2{}^t u_2,$$

where

$$u_1 = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \text{ and } u_2 = \begin{pmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix},$$

So

$$A = 8\begin{pmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{pmatrix} + 3\begin{pmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{pmatrix}.$$

Quadratic Forms

Definition 4

A quadratic form on $\mathbb{R}^n$ is a function Q defined on $\mathbb{R}^n$ whose value at a vector x in $\mathbb{R}^n$ can be computed by an expression of the form

$$Q(x) = {}^t x A x,$$

where A is an $n \times n$ symmetric matrix. The matrix A is called the matrix of the quadratic form.

Example

Example 5

Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$. Then we have

$$ {}^t x A x = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a x_1 \\ b x_2 \end{pmatrix} = a x_1^2 + b x_2^2 $$

Example

Example 6

Let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $A = \begin{pmatrix} 3 & -2 \\ -2 & 7 \end{pmatrix}$. Then we have

$$
{}^t x A x = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3 & -2 \\ -2 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3x_1 - 2x_2 \\ -2x_1 + 7x_2 \end{pmatrix}
$$

$$
= x_1(3x_1 - 2x_2) + x_2(-2x_1 + 7x_2)
$$

$$
= 3x_1^2 - 4x_1 x_2 + 7x_2^2
$$

The term $x_1 x_2$ is called cross-product term.

Change of Variable

If x represents a variable vector in $\mathbb{R}^n$, then a change of variable is an equation of the form

$$x = Py, \quad \text{or equivalently,} \quad y = P^{-1}x, \qquad (1)$$

where P is an invertible matrix and y is a new variable vector in $\mathbb{R}^n$. If the change of variable (1) is made in a quadratic form ${}^t x A x$, then

$${}^t x A x = {}^t(Py)A(Py) = {}^t y({}^t PAP)y \qquad (2)$$

and the new matrix of the quadratic form is ${}^t PAP$. Since A is symmetric, by Theorem 2, there exists an *orthogonal* matrix P such that ${}^t PAP$ is a diagonal matrix D, and the quadratic form in (2) becomes ${}^t yDy$.

The Principal Axes Theorem

Theorem 4 (The Principal Axes Theorem)

Let A be an $n \times n$ symmetric matrix. Then there exists an orthogonal matrix P so that the change of variable $x = Py$ transforms the quadratic form ${}^t xAx$ into a quadratic form ${}^t yDy$ with no cross-product term, where $D = {}^t PAP$ is diagonal.

Definition 5

The columns of P in Theorem 4 are called the principal axes of the quadratic form ${}^t xAx$.

Example

Example 7

The quadratic form corresponding to the matrix $A = \begin{pmatrix} 1 & -4 \\ -4 & -5 \end{pmatrix}$ is given by

$$Q(x) = {}^t x A x = x_1^2 - 8x_1 x_2 - 5x_2^2.$$

The matrix A has eigenvalues 3 and -7 with respective unit eigenvectors $\begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}$ and $\begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$.
Let

$$P = \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \text{ and } D = \begin{pmatrix} 3 & 0 \\ 0 & -7 \end{pmatrix}.$$

Now, by setting $x = Py$, we obtain

$$Q(x) = {}^t y({}^t PAP)y = {}^t y D y = 3y_1^2 - 7y_2^2.$$

Positive and Negative Definite Quadratic Forms

Definition 6

A quadratic form Q is :

a) positive definite if $Q(x) > 0$ for all $x \neq 0$.

b) positive semidefinite if $Q(x) \geq 0$ for all $x \in \mathbb{R}^n$.

c) negative definite if $Q(x) < 0$ for all $x \neq 0$.

d) negative semidefinite if $Q(x) \leq 0$ for all $x \in \mathbb{R}^n$.

e) indefinite if $Q(x)$ has both positive and negative values.

Positive and Negative Definite Quadratic Forms

Theorem 5 (Quadratic Forms and Eigenvalues)

Let A be an $n \times n$ symmetric matrix. Then a quadratic form ${}^t x A x$ is :

a) positive definite if and only if the eigenvalues of A are all positive,

b) negative definite if and only if the eigenvalues of A are all negative, or

c) indefinite if and only if A has both positive and negative eigenvalues.

Example

Example 8

Let the quadratic form

$$Q(x) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3.$$

Its matrix is given by

$$A = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix}.$$

The eigenvalues of A are $5, 2$, and -1. It follows that Q is an indefinite quadratic form.

Constrained Optimization

Theorem 6

Let A be a symmetric matrix, and let

$$m = min\{{}^{t}xAx : \|x\| = 1\}, \qquad M = max\{{}^{t}xAx : \|x\| = 1\}.$$

Then M is the greatest eigenvalue λ_1 of A and m is the least eigenvalue of A. The value of ${}^{t}xAx$ is M when x is a unit eigenvector u_1 corresponding to M. The value of ${}^{t}xAx$ is m when x is a unit eigenvector corresponding to m.

Example

Example 9

We have seen in Example 8 that the matrix

$$A = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 2 & 2 \\ 0 & 2 & 1 \end{pmatrix}.$$

of the quadratic form

$$Q(x) = 3x_1^2 + 2x_2^2 + x_3^2 + 4x_1x_2 + 4x_2x_3$$

has eigenvalues $5, 2,$ and -1. Then the maximum value of $Q(x)$ subject to the constraint $\|x\| = 1$ is 5, and this constrained maximum is attained when $x = \begin{pmatrix} 2/3 \\ 2/3 \\ 1/3 \end{pmatrix}$.

Maximum with Two Constraints

Theorem 7

Let A, λ_1 and u_1 be as in Theorem 6. Then the maximum value of $Q(x)$ subject to the constraints

$$\|x\| = 1, {}^t x\, u_1 = 0$$

is the second greatest eigenvalue, λ_2, and this maximum is attained when x is an eigenvector u_2 corresponding to λ_2.

Example

Example 10

We determine the maximum value of

$$Q(x) = 9x_1^2 + 4x_2^2 + 3x_3^2$$

subject to the constraints

$$\|x\| = 1 \text{ and } {}^t x\, u_1 = 0, \text{ where } u_1 = (1, 0, 0).$$

Note that u_1 is a unit eigenvector corresponding to the greatest eigenvalue $\lambda_1 = 9$ of the matrix of $Q(x)$. Thus the constrained maximum of the quadratic form is attained for $x = (0, 1, 0)$, which is an eigenvector for the second greatest eigenvalue $\lambda_2 = 4$ of the matrix of the quadratic form. Therefore, the constrained maximum value is $Q(0, 1, 0) = 4$.

Maximum with Several Constraints

Theorem 8

Let A be a symmetric $n \times n$ matrix with an orthogonal diagonalization $A = PDP^{-1}$, where the entries on the diagonal of D are arranged so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ and where the columns of P are corresponding unit eigenvectors $u_1, u_2, \cdots, u_n$. Then for $k = 2, \cdots, n$, the maximum value of $Q(x)$ subject to the constraints

$$\|x\| = 1, {}^t x\, u_1 = 0, \cdots, {}^t x\, u_{k-1} = 0$$

is the eigenvalue λ_k, and this maximum is attained at $x = u_k$.

The Singular Values of a Matrix

Proposition 2

Let A be an $m \times n$ matrix. Then the matrix ${}^tA\,A$ is symmetric and all its eigenvalues are nonnegative.

Definition 7

Let A be an $m \times n$ matrix and denote by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ the eigenvalues of ${}^tA\,A$. Then the singular values of A are given by

$$\sigma_i = \sqrt{\lambda_i} \quad \text{for } 1 \leq i \leq n.$$

Remark 2

From the proof of Proposition 2, we can see that the singular values of A are the lengths of the vectors $Av_1, \cdots, Av_n$, where $v_1, \cdots, v_n$ are unit eigenvectors corresponding to $\lambda_1, \cdots, \lambda_n$.

Example

Example 11

Let

$$A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}.$$

Then

$${}^{t}A\,A = \begin{pmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{pmatrix}.$$

The eigenvalues of ${}^{t}A\,A$ are $\lambda_1 = 360,\ \lambda_2 = 90$ and $\lambda_3 = 0$, so the singular values of A are

$$\sigma_1 = 6\sqrt{10}, \qquad \sigma_2 = 3\sqrt{10} \qquad \text{and } \sigma_3 = 0.$$

Singular Values and Orthogonal Basis for Col A

Theorem 9

Let A be an $m \times n$ matrix. Suppose $\{v_1, \cdots, v_n\}$ is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of the matrix ${}^tA\,A$, arranged so that the corresponding eigenvalues of ${}^tA\,A$ satisfy $\lambda_1 \geq \cdots \geq \lambda_n$, and suppose A has r nonzero singular values. Then $\{Av_1, \cdots, Av_r\}$ is an orthogonal basis for $\mathrm{Col}\,A$, and $\mathrm{rank}\,A = r$.

The Singular Value Decomposition

Theorem 10 (The Singular Value Decomposition)

Let A be an $m \times n$ matrix with rank r. Then there exists an $m \times n$ matrix of the form

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix},$$

where D is an $r \times r$ diagonal matrix whose diagonal entries are the first r singular values of A, $\sigma_1 \geq \cdots \geq \sigma_r > 0$, and there exist an $m \times m$ orthogonal matrix U and an $n \times n$ orthogonal matrix V such that

$$A = U\Sigma^t V.$$

Definition 8

The factorization given in Theorem 10 is called a singular value decomposition (or SVD) of A. The columns of U in such a decomposition are called left singular vectors of A, and the columns of V are called right singular vectors of A.

Example

Example 12

Consider again the matrix

$$A = \begin{pmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{pmatrix}.$$

Step 1 *Find an orthogonal diagonalization of* ${}^tA\,A$**.**

That is, find the eigenvalues of ${}^tA\,A$ and a corresponding orthonormal set of eigenvectors.

Example

Example 12 (continued)

Step 2. *Set up* V *and* Σ

Arrange the eigenvalues of ${}^{t}A\,A$ in decreasing order : $\lambda_1 = 360$, $\lambda_2 = 90$ and $\lambda_3 = 0$, then the corresponding unit eigenvectors v_1, v_2 and v_3 are the columns of the matrix

$$V = (v_1\ v_2\ v_3) = \begin{pmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{pmatrix}.$$

The singular nonzero values of A are $\sigma_1 = 6\sqrt{10}$ and $\sigma_2 = 3\sqrt{10}$, so

$$\Sigma = \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix}.$$

Example

Example 12 (continued)

Step 3. Construct U.

The first two columns of U are the normalized vectors obtained from Av_1 and Av_2.

Since $\|Av_i\| = \sigma_i$ for $1 \le i \le 2$, then we have

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{6\sqrt{10}} \begin{pmatrix} 18 \\ 6 \end{pmatrix} = \begin{pmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{pmatrix},$$

$$u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{3\sqrt{10}} \begin{pmatrix} 3 \\ -9 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{pmatrix}.$$

Example

Example 12 (continued)

Note that $\{u_1, u_2\}$ is a basis of $\mathbb{R}^2$, so no additional vectors are needed for U and $U = (u_1\ u_2)$.

Finally, the singular value decomposition of A is

$$A = \begin{pmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{pmatrix} \begin{pmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{pmatrix} \begin{pmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{pmatrix}.$$

Exercises

Exercise 1

1) Let $Q(x) = 3x_1^2 + 3x_2^2 + 2x_1x_2$. Find a change of variable that transforms Q into a quadratic form with no cross-product term, and give the new quadratic form.

2) Find the maximum value of $Q(x)$ subject to the constraint ${}^t xx = 1$, and find a unit vector at which the maximum is attained.

Exercise 2

Let λ be any eigenvalue of a symmetric matrix A. Show that $m \leq \lambda \leq M$.

Exercises

Exercise 3

Show that in the SVD of A, the columns of V are eigenvectors of ${}^tAA$, the columns of U are eigenvectors of A^tA , and the diagonal entries of Σ are the singular values of A.

Exercise 4

1) Given a singular value decomposition, $A = U\Sigma{}^tV$, find an SVD of ${}^tA$. How are the singular values of A and ${}^tA$ related?

2) For any $n \times n$ matrix A, use the SVD to show that there is an $n \times n$ orthogonal matrix Q such that $A^tA = {}^tQ({}^tAA)Q$.

Exercise 5

Show that if A is square, then $|\det(A)|$ is the product of the singular values of A.