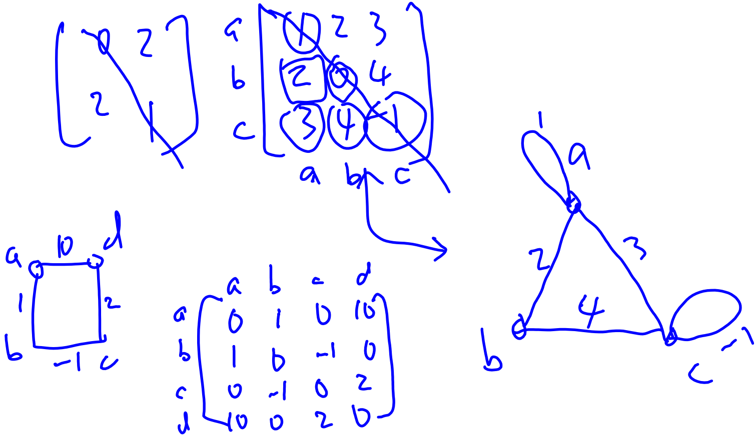
Section 7.1 Diagonalization of symmetric matrices and 7.2 Quadratic forms

Gexin Yu
gyu@wm.edu

College of William and Mary

Symmetric matrix

- A symmetric matrix is a matrix A such that $A^T = A$.



Symmetric matrix

- A symmetric matrix is a matrix A such that $A^T = A$.
- Such a matrix is necessarily square.

Symmetric matrix

- A symmetric matrix is a matrix A such that $A^T = A$.
- Such a matrix is necessarily square.
- Its main diagonal entries are arbitrary, but its other entries occur in pairs on opposite sides of the main diagonal.

- Theorem: If A is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.

- Theorem: If A is symmetric, then any two eigenvectors from different eigenspaces are orthogonal.

$Av_1 = \lambda_1 v_1$

Proof: Let v_1 and v_2 be eigenvectors that correspond to distinct eigenvalues, say λ_1 and λ_2. We show that $v_1 \cdot v_2 = 0$. Note that

$$\begin{aligned}\lambda_1 v_1 \cdot v_2 &= (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2 = (v_1^T A^T) v_2 \\ &= v_1^T (A^T v_2) = v_1^T (Av_2) = v_1^T (\lambda_2 v_2) = \lambda_2 (v_1^T v_2) \\ &= \lambda_2 v_1 \cdot v_2.\end{aligned}$$

It follows that $(\lambda_1 - \lambda_2)(v_1 \cdot v_2) = 0$. So $v_1 \cdot v_2 = 0$.

Orthogonally diagonalizable matrix

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

Orthogonally diagonalizable matrix

- An $n \times n$ matrix A is said to be orthogonally diagonzlizable if there are orthogonal matrix P (with $P^{-1} = P^T$) and a diagonal matrix D such that
$$A = PDP^T = PDP^{-1}$$
- Such a diagonalization requires n linearly independent and orthogonal eigenvectors.

Orthogonally diagonalizable matrix

- An $n \times n$ matrix A is said to be orthogonally diagonzlizable if there are orthogonal matrix P (with $P^{-1} = P^T$) and a diagonal matrix D such that
$$A = PDP^T = PDP^{-1}$$
- Such a diagonalization requires n linearly independent and orthogonal eigenvectors.
- When is this possible?

Orthogonally diagonalizable matrix

- An $n \times n$ matrix A is said to be orthogonally diagonzlizable if there are orthogonal matrix P (with $P^{-1} = P^T$) and a diagonal matrix D such that
$$A = PDP^T = PDP^{-1}$$
- Such a diagonalization requires n linearly independent and orthogonal eigenvectors.
- When is this possible?
- If A is orthogonally diagonalizable, then
$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$

Orthogonally diagonalizable matrix

- An $n \times n$ matrix A is said to be orthogonally diagonzlizable if there are orthogonal matrix P (with $P^{-1} = P^T$) and a diagonal matrix D such that
$$A = PDP^T = PDP^{-1}$$
- Such a diagonalization requires n linearly independent and orthogonal eigenvectors.
- When is this possible?
- If A is orthogonally diagonalizable, then
$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$
- So A should be symmetric.

Orthogonally diagonalizable matrix

- An $n \times n$ matrix A is said to be orthogonally diagonzlizable if there are orthogonal matrix P (with $P^{-1} = P^T$) and a diagonal matrix D such that
$$A = PDP^T = PDP^{-1}$$
- Such a diagonalization requires n linearly independent and orthogonal eigenvectors.
- When is this possible?
- If A is orthogonally diagonalizable, then
$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A$$
- So A should be symmetric.
- Theorem: An $n \times n$ matrix A is orthogonally diagonalizable if and only if A is symmetric.

Example

- Ex: Orthogonally diagonalize the matrix $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3. \end{bmatrix}$

Example

- Ex: Orthogonally diagonalize the matrix $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3. \end{bmatrix}$
- First of all, its charateristic equation is $det(A - \lambda I) = (\lambda - 7)^2(\lambda + 2)$. So its eigenvalues are 7 (with multiplicity 2) and -2.

$$det\left(\begin{bmatrix} 3-\lambda & -2 & 4 \\ -2 & 6-\lambda & 2 \\ 4 & 2 & 3-\lambda \end{bmatrix}\right)$$

Example

- Ex: Orthogonally diagonalize the matrix $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3. \end{bmatrix}$
- First of all, its charateristic equation is $det(A - \lambda I) = (\lambda - 7)^2(\lambda + 2)$. So its eigenvalues are 7 (with multiplicity 2) and -2.
- The bases for the eigenspaces are

$$\lambda = 7 : v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}; \qquad \lambda = -2 : v_3 = \begin{pmatrix} -1 \\ -\frac{1}{2} \\ 1 \end{pmatrix}$$

$$(A - 7I)x = 0$$

$$A - 7I = \begin{bmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

x_2 x_3

$$x_1 = -\frac{1}{2}x_2 + x_3 \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{bmatrix} -\frac{1}{2}x_2 + x_3 \\ x_2 \\ x_3 \end{bmatrix} =$$

Example

- Ex: Orthogonally diagonalize the matrix $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3. \end{bmatrix}$
- First of all, its charateristic equation is $det(A - \lambda I) = (\lambda - 7)^2(\lambda + 2)$. So its eigenvalues are 7 (with multiplicity 2) and -2.
- The bases for the eigenspaces are

$$\lambda = 7 : v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}; \qquad \lambda = -2 : v_3 = \begin{pmatrix} -1 \\ -\frac{1}{2} \\ 1 \end{pmatrix}$$

- The projection of v_2 onto v_1 is $\frac{v_2 \cdot v_1}{v_1 \cdot v_1} v_1$, so the component of v_2 orthogonal to v_1 is $z_2 = v_2 - \frac{v_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \begin{pmatrix} -\frac{1}{4} \\ 1 \\ \frac{1}{4} \end{pmatrix}$

Gram-Schmidt process (6.4)

Example

- Ex: Orthogonally diagonalize the matrix $\begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3. \end{bmatrix}$
- First of all, its charateristic equation is $det(A - \lambda I) = (\lambda - 7)^2(\lambda + 2)$. So its eigenvalues are 7 (with multiplicity 2) and -2.
- The bases for the eigenspaces are

$$\lambda = 7 : v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}; \qquad \lambda = -2 : v_3 = \begin{pmatrix} -1 \\ -\frac{1}{2} \\ 1 \end{pmatrix}$$

- The projection of v_2 onto v_1 is $\frac{v_2 \cdot v_1}{v_1 \cdot v_1} v_1$, so the component of v_2 orthogonal to v_1 is $z_2 = v_2 - \frac{v_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \begin{pmatrix} -\frac{1}{4} \\ 1 \\ \frac{1}{4} \end{pmatrix}$
- So $\{v_1, z_2\}$ is an orthogonal set in the eigenspace for $\lambda = 7$.

- Normalize v_1 and z_2 to get a orthonormal basis for the eigenspace for $\lambda = 7$:

$u_1 = \frac{v_1}{\|v_1\|}$

$$u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \quad u_2 = \begin{pmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{pmatrix}$$

- Normalize v_1 and z_2 to get a orthonormal basis for the eigenspace for $\lambda = 7$:

$$u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \quad u_2 = \begin{pmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{pmatrix}$$

- An orthonormal basis for the eigenspace for $\lambda = -2$ is

$$u_3 = \frac{v_3}{||v_3||} = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}$$

- Normalize v_1 and z_2 to get a orthonormal basis for the eigenspace for $\lambda = 7$:

$$u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \quad u_2 = \begin{pmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{pmatrix}$$

- An orthonormal basis for the eigenspace for $\lambda = -2$ is

$$u_3 = \frac{v_3}{||v_3||} = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}$$

- Now $\{u_1, u_2, u_3\}$ is an orthonormal set.

- Normalize v_1 and z_2 to get a orthonormal basis for the eigenspace for $\lambda = 7$:

$$u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \quad u_2 = \begin{pmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{pmatrix}$$

- An orthonormal basis for the eigenspace for $\lambda = -2$ is

$$u_3 = \frac{v_3}{||v_3||} = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}$$

- Now $\{u_1, u_2, u_3\}$ is an orthonormal set.
- Let

$$P = [u_1 \ u_2 \ u_3] = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{18}} & -\frac{2}{3} \\ 0 & \frac{4}{\sqrt{18}} & -\frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} & \frac{2}{3} \end{bmatrix}, \quad D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

- Normalize v_1 and z_2 to get a orthonormal basis for the eigenspace for $\lambda = 7$:

$$u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \quad u_2 = \begin{pmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{pmatrix}$$

- An orthonormal basis for the eigenspace for $\lambda = -2$ is

$$u_3 = \frac{v_3}{||v_3||} = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}$$

- Now $\{u_1, u_2, u_3\}$ is an orthonormal set.
- Let

$$P = [u_1 \ u_2 \ u_3] = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{18}} & -\frac{2}{3} \\ 0 & \frac{4}{\sqrt{18}} & -\frac{1}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} & \frac{2}{3} \end{bmatrix}, \quad D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

- then P orthogonally diagonalizes A, and $A = PDP^T$.

- The set of eigenvalues of a matrix A is sometimes called the spectrum of A.

the Spectral Theorem

- The set of eigenvalues of a matrix A is sometimes called the spectrum of A.
- Spectral Theorem: An $n \times n$ symmetric matrix A has the following properties:

the Spectral Theorem

- The set of eigenvalues of a matrix A is sometimes called the spectrum of A.
- Spectral Theorem: An $n \times n$ symmetric matrix A has the following properties:
  1. A has n real eigenvalues, counting multiplicities.

the Spectral Theorem

- The set of eigenvalues of a matrix A is sometimes called the spectrum of A.

- Spectral Theorem: An $n \times n$ symmetric matrix A has the following properties:
  1. A has n real eigenvalues, counting multiplicities.
  2. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation.

the Spectral Theorem

- The set of eigenvalues of a matrix A is sometimes called the spectrum of A.
- Spectral Theorem: An $n \times n$ symmetric matrix A has the following properties:
  1. A has n real eigenvalues, counting multiplicities.
  2. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation.
  3. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.

the Spectral Theorem

- The set of eigenvalues of a matrix A is sometimes called the spectrum of A.

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

  1. A has n real eigenvalues, counting multiplicities.

  2. The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation.

  3. The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.

  4. A is orthogonally diagonalizable.

Spectral decomposition

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

$$A = PDP^T = [u_1 \;\; u_2 \;\; \ldots \;\; u_n] \begin{bmatrix} \lambda_1 & 0 \ldots & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} \begin{pmatrix} u_1^T \\ u_2^T \\ \ldots \\ u_n^T \end{pmatrix}$$

$$= \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \ldots + \lambda_n u_n u_n^T$$

$n \times n$ matrix

$$u_1 u_1^T = \begin{bmatrix} u_{11} \\ \vdots \\ u_{1n} \end{bmatrix}_{n\times 1} \left[u_{11} \cdots u_{1n} \right]_{1 \times n}$$

$$= (\lambda_1 u_1 \;\; \lambda_2 u_2 \cdots \lambda_n u_n) \begin{pmatrix} u_1^T \\ \vdots \\ u_n^T \end{pmatrix}$$

Spectral decomposition

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

$$
\begin{aligned}
A =& PDP^T = [u_1 \ u_2 \ \ldots \ u_n] \begin{bmatrix} \lambda_1 & 0\ldots & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} \begin{pmatrix} u_1^T \\ u_2^T \\ \ldots \\ u_n^T \end{pmatrix} \\
=& \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \ldots + \lambda_n u_n u_n^T
\end{aligned}
$$

- The expression $A = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \ldots + \lambda_n u_n u_n^T$ is called a spectral decomposition of A, because if breaks A into pieces determined by the spectrum (eigenvalues) of A.

Spectral decomposition

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

$$A = PDP^T = [u_1 \ u_2 \ \ldots \ u_n] \begin{bmatrix} \lambda_1 & 0 \ldots & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} \begin{pmatrix} u_1^T \\ u_2^T \\ \ldots \\ u_n^T \end{pmatrix}$$
$$= \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \ldots + \lambda_n u_n u_n^T$$

- The expression $A = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \ldots + \lambda_n u_n u_n^T$ is called a spectral decomposition of A, because if breaks A into pieces determined by the spectrum (eigenvalues) of A.

- Each term in the decomposition is an $n \times n$ matrix of rank 1.

Example

- Ex:

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

u_1 u_2

Example

- Ex:

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

- Then

$$A = 8u_1u_1^T + 3u_2u_2^T = \begin{bmatrix} \frac{32}{5} & \frac{16}{5} \\ \frac{16}{5} & \frac{8}{5} \end{bmatrix} + \begin{bmatrix} \frac{3}{5} & -\frac{6}{5} \\ -\frac{6}{5} & \frac{12}{5} \end{bmatrix}$$

Quadratic forms

- A quadratic form is a function defined on $\mathbf{R}^n$ whose value at a vector x in $\mathbf{R}^n$ can be computed by an expression of the form $Q(x) = x^T Ax$, where A is an $n \times n$ symmetric matrix.

Quadratic forms

- A quadratic form is a function defined on $\mathbf{R}^n$ whose value at a vector x in $\mathbf{R}^n$ can be computed by an expression of the form $Q(x) = x^T A x$, where A is an $n \times n$ symmetric matrix.
- The matrix A is called the matrix of the quadratic form.

$$2x_1^2 + 3x_1x_2 - 2x_1x_3 + 4x_2x_3 + x_2^2 + 5x_3^2 = x^T A x$$

$$A = \begin{pmatrix} 2 & 3/2 & -1 \\ 3/2 & 1 & 2 \\ -1 & 2 & 5 \end{pmatrix} \qquad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Quadratic forms

- A quadratic form is a function defined on $\mathbf{R}^n$ whose value at a vector x in $\mathbf{R}^n$ can be computed by an expression of the form $Q(x) = x^T A x$, where A is an $n \times n$ symmetric matrix.
- The matrix A is called the matrix of the quadratic form.
- Ex: let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Compute $x^T A x$ for the following matrices:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$$

x^TAx

$= (x_1\ x_2)\begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ (1×2, 2×2)

$= (4x_1\ \ 3x_2)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 4x_1^2 + 3x_2^2$

Quadratic forms

- A quadratic form is a function defined on $\mathbf{R}^n$ whose value at a vector x in $\mathbf{R}^n$ can be computed by an expression of the form $Q(x) = x^T Ax$, where A is an $n \times n$ symmetric matrix.
- The matrix A is called the matrix of the quadratic form.
- Ex: let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Compute $x^T Ax$ for the following matrices:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$$

- $x^T Ax = 4x_1^2 + 3x_2^2$.

Quadratic forms

- A quadratic form is a function defined on $\mathbf{R}^n$ whose value at a vector x in $\mathbf{R}^n$ can be computed by an expression of the form $Q(x) = x^T Ax$, where A is an $n \times n$ symmetric matrix.
- The matrix A is called the matrix of the quadratic form.
- Ex: let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Compute $x^T Ax$ for the following matrices:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$$

- $x^T Ax = 4x_1^2 + 3x_2^2$.
- $x^T Bx = 3x_1^2 - 4x_1x_2 + 7x_2^2$.

Quadratic forms

- A quadratic form is a function defined on $\mathbf{R}^n$ whose value at a vector x in $\mathbf{R}^n$ can be computed by an expression of the form $Q(x) = x^T Ax$, where A is an $n \times n$ symmetric matrix.
- The matrix A is called the matrix of the quadratic form.
- Ex: let $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Compute $x^T Ax$ for the following matrices:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}, \qquad B = \begin{bmatrix} 3 & -2 \\ -2 & 7 \end{bmatrix}$$

- $x^T Ax = 4x_1^2 + 3x_2^2$.
- $x^T Bx = 3x_1^2 - 4x_1x_2 + 7x_2^2$.
- The term $-4x_1x_2$ is called a cross-product term.

Change of variable in a quadratic form

- If x represents a variable vector in $\mathbf{R}^n$, then a change of variable is an equation of the form $x = Py$, or $y = P^{-1}x$, where P is an invertible matrix and y is a new variable vector in $\mathbf{R}^n$.

Change of variable in a quadratic form

- If x represents a variable vector in $\mathbf{R}^n$, then a change of variable is an equation of the form $x = Py$, or $y = P^{-1}x$, where P is an invertible matrix and y is a new variable vector in $\mathbf{R}^n$.
- By make a change of variable on x^TAx, we have

$$x^TAx = (Py)^TA(Py) = y^TP^TAPy = y^T(P^TAP)y$$

we obtain a new matrix of the quadratic form P^TAP.

Change of variable in a quadratic form

- If x represents a variable vector in $\mathbf{R}^n$, then a change of variable is an equation of the form $x = Py$, or $y = P^{-1}x$, where P is an invertible matrix and y is a new variable vector in $\mathbf{R}^n$.
- By make a change of variable on $x^T Ax$, we have

$$x^T Ax = (Py)^T A(Py) = y^T P^T APy = y^T (P^T AP)y$$

we obtain a new matrix of the quadratic form $P^T AP$.

- Since A is symmetric, there is an orthogonal matrix P such that $P^T AP$ is a diagonal matrix D, and the quadratic form above becomes $y^T Dy$, which contains no cross-product terms.

Change of variable in a quadratic form

- If x represents a variable vector in $\mathbf{R}^n$, then a change of variable is an equation of the form $x = Py$, or $y = P^{-1}x$, where P is an invertible matrix and y is a new variable vector in $\mathbf{R}^n$.
- By make a change of variable on $x^T Ax$, we have

$$x^T Ax = (Py)^T A(Py) = y^T P^T APy = y^T(P^T AP)y$$

  we obtain a new matrix of the quadratic form $P^T AP$.
- Since A is symmetric, there is an orthogonal matrix P such that $P^T AP$ is a diagonal matrix D, and the quadratic form above becomes $y^T Dy$, which contains no cross-product terms.
- Theorem: Let A be an $n \times n$ symmetric matrix. Then there is an orthogonal change of variable, $x = Py$, that transforms the quadratic form $x^T Ax$ into a quadratic form $y^T Dy$ with no cross-product term.

Example

- Ex: Make a change of variable that transforms the quadratic form $Q(x) = x_1^2 - 8x_1x_2 - 5x_2^2$ into a quadratic form with no cross-product term.

Example

- Ex: Make a change of variable that transforms the quadratic form $Q(x) = x_1^2 - 8x_1x_2 - 5x_2^2$ into a quadratic form with no cross-product term.
- The matrix of the given quadratic form is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

Example

- Ex: Make a change of variable that transforms the quadratic form $Q(x) = x_1^2 - 8x_1x_2 - 5x_2^2$ into a quadratic form with no cross-product term.
- The matrix of the given quadratic form is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

- Orthogonally diagonalize A: the eigenvalues of A are 3 and -7, and the associated unit eigenvectors are $\lambda = 3 : [\frac{2}{\sqrt{5}} \;\; -\frac{1}{\sqrt{5}}]^T$ and $\lambda = -7 : [\frac{1}{\sqrt{5}} \;\; \frac{2}{\sqrt{5}}]^T$.

Example

- Ex: Make a change of variable that transforms the quadratic form $Q(x) = x_1^2 - 8x_1x_2 - 5x_2^2$ into a quadratic form with no cross-product term.
- The matrix of the given quadratic form is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

- Orthogonally diagonalize A: the eigenvalues of A are 3 and -7, and the associated unit eigenvectors are $\lambda = 3 : [\frac{2}{\sqrt{5}} \quad -\frac{1}{\sqrt{5}}]^T$ and $\lambda = -7 : [\frac{1}{\sqrt{5}} \quad \frac{2}{\sqrt{5}}]^T$.
- Let

$$P = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}, D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

Example

- Ex: Make a change of variable that transforms the quadratic form $Q(x) = x_1^2 - 8x_1x_2 - 5x_2^2$ into a quadratic form with no cross-product term.
- The matrix of the given quadratic form is

$$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$$

- Orthogonally diagonalize A: the eigenvalues of A are 3 and -7, and the associated unit eigenvectors are $\lambda = 3 : [\frac{2}{\sqrt{5}} \quad -\frac{1}{\sqrt{5}}]^T$ and $\lambda = -7 : [\frac{1}{\sqrt{5}} \quad \frac{2}{\sqrt{5}}]^T$.
- Let

$$P = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}, D = \begin{bmatrix} 3 & 0 \\ 0 & -7 \end{bmatrix}$$

- then $A = PDP^T$ and $D = P^TAP$. Let $x = Py$, then

$$x_1^2 - 8x_1x_2 - 5x_2^2 = y^TDy = 3y_1^2 - 7y_2^2$$

- The columns of P in the theorem are called the principal axes of the quadratic form $x^T Ax$.

The principal axes theorem

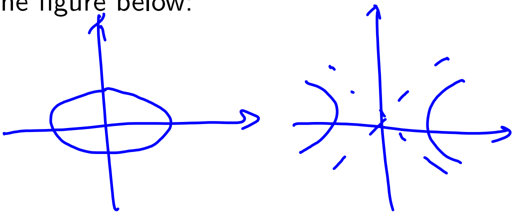
- The columns of P in the theorem are called the principal axes of the quadratic form $x^T Ax$.
- The vector y is the coordinate vector of x relative to the orthonormal basis of $\mathbf{R}^n$ given by these principal axes.

The principal axes theorem

- The columns of P in the theorem are called the principal axes of the quadratic form $x^T Ax$.
- The vector y is the coordinate vector of x relative to the orthonormal basis of $\mathbf{R}^n$ given by these principal axes.
- It can be shown that the set of all x in $\mathbf{R}^n$ that satisfy $x^T Ax = c$ either corresponds to an ellipse (or a circle), a hyperbola, two intersecting lines, or a single point, or contains no points at all.

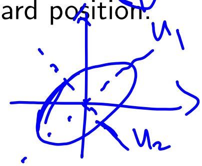
The principal axes theorem

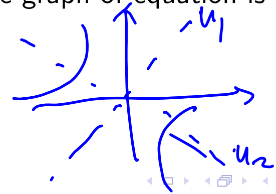
- The columns of P in the theorem are called the principal axes of the quadratic form $x^T Ax$.
- The vector y is the coordinate vector of x relative to the orthonormal basis of $\mathbf{R}^n$ given by these principal axes.
- It can be shown that the set of all x in $\mathbf{R}^n$ that satisfy $x^T Ax = c$ either corresponds to an ellipse (or a circle), a hyperbola, two intersecting lines, or a single point, or contains no points at all.
- If A is a diagonal matrix, the graph is in standard position, such as the figure below:



The principal axes theorem

- The columns of P in the theorem are called the principal axes of the quadratic form x^TAx.
- The vector y is the coordinate vector of x relative to the orthonormal basis of $\mathbf{R}^n$ given by these principal axes.
- It can be shown that the set of all x in $\mathbf{R}^n$ that satisfy $x^TAx = c$ either corresponds to an ellipse (or a circle), a hyperbola, two intersecting lines, or a single point, or contains no points at all.
- If A is a diagonal matrix, the graph is in standard position, such as the figure below:
- If A is not a diagonal matrix, the graph of equation is rotated out of the standard position.





The principal axes theorem

- The columns of P in the theorem are called the principal axes of the quadratic form $x^T Ax$.
- The vector y is the coordinate vector of x relative to the orthonormal basis of $\mathbf{R}^n$ given by these principal axes.
- It can be shown that the set of all x in $\mathbf{R}^n$ that satisfy $x^T Ax = c$ either corresponds to an ellipse (or a circle), a hyperbola, two intersecting lines, or a single point, or contains no points at all.
- If A is a diagonal matrix, the graph is in standard position, such as the figure below:
- If A is not a diagonal matrix, the graph of equation is rotated out of the standard position.
- Finding the principal axes (determined by the eigenvectors of A) amounts to finding a new coordinate system with respect to which the graph is in standard position.

Classifying quadratic forms

- Definition: A quadratic form $Q(x)$ is

Classifying quadratic forms

- Definition: A quadratic form $Q(x)$ is
  - positive definite if $Q(x) > 0$ for all $x \neq 0$;

Classifying quadratic forms

- Definition: A quadratic form $Q(x)$ is
  - positive definite if $Q(x) > 0$ for all $x \neq 0$;
  - negative definite if $Q(x) < 0$ for all $x \neq 0$;

Classifying quadratic forms

- Definition: A quadratic form $Q(x)$ is
  - positive definite if $Q(x) > 0$ for all $x \neq 0$;
  - negative definite if $Q(x) < 0$ for all $x \neq 0$;
  - indefinite if $Q(x)$ assumes both positive and negative values.

$x_1^2 + 2x_2^2 > 0$

$-x_1^2 - 4x_2^2 < 0$

$x_1^2 - 2x_2^2 \begin{cases} > 0 \\ < 0 \end{cases}$

Classifying quadratic forms

- Definition: A quadratic form $Q(x)$ is
  - positive definite if $Q(x) > 0$ for all $x \neq 0$;
  - negative definite if $Q(x) < 0$ for all $x \neq 0$;
  - indefinite if $Q(x)$ assumes both positive and negative values.

- Also, Q is said to be positive semidefinite if $Q(x) \geq 0$ for all x, and negative semidefinite if $Q(x) \leq 0$ for all x.

Classifying quadratic forms

- Definition: A quadratic form $Q(x)$ is
  - positive definite if $Q(x) > 0$ for all $x \neq 0$;
  - negative definite if $Q(x) < 0$ for all $x \neq 0$;
  - indefinite if $Q(x)$ assumes both positive and negative values.
- Also, Q is said to be positive semidefinite if $Q(x) \geq 0$ for all x, and negative semidefinite if $Q(x) \leq 0$ for all x.
- Theorem: Let A be an $n \times n$ symmetric matrix. Then a quadratic form $xTAx$ is:

Classifying quadratic forms

- Definition: A quadratic form $Q(x)$ is
  - positive definite if $Q(x) > 0$ for all $x \neq 0$;
  - negative definite if $Q(x) < 0$ for all $x \neq 0$;
  - indefinite if $Q(x)$ assumes both positive and negative values.

- Also, Q is said to be positive semidefinite if $Q(x) \geq 0$ for all x, and negative semidefinite if $Q(x) \leq 0$ for all x.

- Theorem: Let A be an $n \times n$ symmetric matrix. Then a quadratic form $xTAx$ is: $\longrightarrow \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$
  - positive definite if and only if the eigenvalues of A are all positive,

Classifying quadratic forms

- Definition: A quadratic form $Q(x)$ is
  - positive definite if $Q(x) > 0$ for all $x \neq 0$;
  - negative definite if $Q(x) < 0$ for all $x \neq 0$;
  - indefinite if $Q(x)$ assumes both positive and negative values.
- Also, Q is said to be positive semidefinite if $Q(x) \geq 0$ for all x, and negative semidefinite if $Q(x) \leq 0$ for all x.
- Theorem: Let A be an $n \times n$ symmetric matrix. Then a quadratic form $xTAx$ is:
  - positive definite if and only if the eigenvalues of A are all positive,
  - negative definite if and only if the eigenvalues of A are all negative, or

Classifying quadratic forms

- Definition: A quadratic form $Q(x)$ is
  - positive definite if $Q(x) > 0$ for all $x \neq 0$;
  - negative definite if $Q(x) < 0$ for all $x \neq 0$;
  - indefinite if $Q(x)$ assumes both positive and negative values.

- Also, Q is said to be positive semidefinite if $Q(x) \geq 0$ for all x, and negative semidefinite if $Q(x) \leq 0$ for all x.

- Theorem: Let A be an $n \times n$ symmetric matrix. Then a quadratic form $xTAx$ is:
  - positive definite if and only if the eigenvalues of A are all positive,
  - negative definite if and only if the eigenvalues of A are all negative, or
  - indefinite if and only if A has both positive and negative eigenvalues.