BITS Pilani
Pilani Campus

Mathematical Foundations for Data Science

MFDS Team

BITS Pilani
Pilani Campus

DSECL ZC416, MFDS

Lecture No. 3

Agenda

- Eigenvalues and eigenvectors
- Gerschgorin's theorem
- Similarity transformation
- Diagonalization of matrices
- Quadratic forms

Eigenvalue Problem

A matrix eigenvalue problem considers the vector equation

(1) $$\mathbf{A}\mathbf{x} = \lambda \mathbf{x},$$

where **A** is a given square matrix, λ an unknown scalar(real or complex), and **x** an unknown vector.

The task is to determine λ's and **x**'s (dependent on λ's)that satisfy (1).

Since $\mathbf{x} = \mathbf{0}$ is always a solution for any λ,we only admit solutions with $\mathbf{x} \neq \mathbf{0}$.

The solutions to (1) are given the following names: The λ's that satisfy (1) are called **eigenvalues of A** and the corresponding nonzero **x**'s that also satisfy (1) are called **eigenvectors of A**.

Spectrum

The set of all the eigenvalues of **A** is called the **spectrum** of **A**. We shall see that the spectrum consists of at least one eigenvalue and at most of n numerically different eigenvalues.

The largest of the absolute values of the eigenvalues of **A** is called the *spectral radius* of **A**, a name to be motivated later.

Determination of Eigen Value and Eigen Vector

$$\mathbf{A} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}.$$

Solution.

(a) ***Eigenvalues.*** These must be determined ***first***.

Equation (1) is in components

$$\mathbf{Ax} = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix};$$

$$\begin{aligned} -5x_1 + 2x_2 &= \lambda x_1 \\ 2x_1 - 2x_2 &= \lambda x_2. \end{aligned}$$

Example

Solution. ***(continued 1)***

(a) ***Eigenvalues.*** ***(continued 1)***

Transferring the terms on the right to the left, we get

$$(2^*)\qquad \begin{aligned} (-5-\lambda)x_1 + \qquad\quad 2x_2 &= 0 \\ 2x_1 + (-2-\lambda)x_2 &= 0 \end{aligned}$$

This can be written in matrix notation

$$(3^*)\qquad (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$$

Because (1) is $\mathbf{Ax} - \lambda\mathbf{x} = \mathbf{Ax} - \lambda\mathbf{Ix} = (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, which gives (3*).

Example

Solution. ***(continued 2)***

(a) ***Eigenvalues.*** ***(continued 2)***

We see that this is a ***homogeneous*** linear system. By Cramer's theorem in Sec. 7.7 it has a nontrivial solution (an eigenvector of **A** we are looking for) if and only if its coefficient determinant is zero, that is,

$$(4^*)\qquad \begin{aligned} D(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} -5-\lambda & 2 \\ 2 & -2-\lambda \end{vmatrix} \\ &= (-5-\lambda)(-2-\lambda) - 4 = \lambda^2 + 7\lambda + 6 = 0. \end{aligned}$$

Example

Solution. ***(continued 3)***

(a) ***Eigenvalues.*** ***(continued 3)***

We call $D(\lambda)$ the **characteristic determinant** or, if expanded, the **characteristic polynomial**, and $D(\lambda) = 0$ the **characteristic equation** of **A**. The solutions of this quadratic equation are $\lambda_1 = -1$ and $\lambda_2 = -6$. These are the eigenvalues of **A**.

(b_1) ***Eigenvector of* A *corresponding to*** $\boldsymbol{\lambda_1}$. This vector is obtained from (2*) with $\lambda = \lambda_1 = -1$, that is,

$$\begin{aligned} -4x_1 + 2x_2 &= 0 \\ 2x_1 - x_2 &= 0. \end{aligned}$$

Example

Solution. ***(continued 4)***

(b_1) ***Eigenvector of* A *corresponding to*** $\boldsymbol{\lambda_1}$. ***(continued)***

A solution is $x_2 = 2x_1$, as we see from either of the two equations, so that we need only one of them. This determines an eigenvector corresponding to $\lambda_1 = -1$ up to a scalar multiple. If we choose $x_1 = 1$, we obtain the eigenvector

$$v = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \text{Check:} \quad \mathbf{A}v = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix} = (-1)v = \lambda_1 v.$$

Example

Solution. *(continued 5)*

(b_2) ***Eigenvector of A corresponding to*** λ_2.

For $\lambda = \lambda_2 = -6$, equation (2*) becomes

$$x_1 + 2x_2 = 0$$

$$2x_1 + 4x_2 = 0.$$

A solution is $x_2 = -x_1/2$ with arbitrary x_1. If we choose $x_1 = 2$, we get $x_2 = -1$. Thus an eigenvector of **A** corresponding to $\lambda_2 = -6$ is

$$w = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \text{Check:} \quad \mathbf{A}w = \begin{bmatrix} -5 & 2 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -12 \\ 6 \end{bmatrix} = (-6)w = \lambda_2 w.$$

Eigen Value Analysis

This example illustrates the general case as follows. Equation (1) written in components is

$$
\begin{aligned}
a_{11}x_1 + \cdots + a_{1n}x_n &= \lambda x_1 \\
a_{21}x_1 + \cdots + a_{2n}x_n &= \lambda x_2 \\
\cdots\cdots\cdots\cdots&\cdots\cdots\cdots \\
a_{n1}x_1 + \cdots + a_{nn}x_n &= \lambda x_n .
\end{aligned}
$$

Transferring the terms on the right side to the left side, we have

$$
\begin{aligned}
(a_{11}-\lambda)x_1 + \quad a_{12}x_2 + \cdots + \quad a_{1n}x_n &= 0 \\
a_{21}x_1 + (a_{22}-\lambda)x_2 + \cdots + \quad a_{2n}x_n &= 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots&\cdots\cdots\cdots \\
a_{n1}x_1 + \quad a_{n2}x_2 + \cdots + (a_{nn}-\lambda)x_n &= 0.
\end{aligned} \tag{2}
$$

Eigen Value Analysis

In matrix notation,

(3) $$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}.$$

By Cramer's theorem in Sec. 7.7, this homogeneous linear system of equations has a nontrivial solution if and only if the corresponding determinant of the coefficients is zero:

(4) $$D(\lambda) = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} = 0.$$

Eigen Value Analysis

$\mathbf{A} - \lambda\mathbf{I}$ is called the **characteristic matrix** and $D(\lambda)$ the **characteristic determinant** of **A**. Equation (4) is called the **characteristic equation** of **A**. By developing $D(\lambda)$ we obtain a polynomial of nth degree in λ. This is called the **characteristic polynomial** of **A**.

Eigen Values

Theorem 1

Eigenvalues

The eigenvalues of a square matrix **A** *are the roots of the characteristic equation* (4) *of* **A**.
Hence an $n \times n$ *matrix has at least one eigenvalue and at most* n *numerically different eigenvalues.*

The eigen**values** ***must be determined first*****.**
Once these are known, corresponding eigen*vectors* are obtained from the system (2), for instance, by the Gauss elimination, where λ is the eigenvalue for which an eigenvector is wanted.

Eigen Space

Theorem 2

Eigenvectors, Eigenspace

If $\mathbf{w}$ *and* $\mathbf{x}$ *are eigenvectors of a matrix* $\mathbf{A}$ *corresponding to* ***the same*** *eigenvalue* λ, *so are* $\mathbf{w} + \mathbf{x}$ (provided $\mathbf{x} \neq -\mathbf{w}$) *and* $k\mathbf{x}$ *for any* $k \neq 0$.

Hence the eigenvectors corresponding to one and the same eigenvalue λ *of* $\mathbf{A}$, *together with* $\mathbf{0}$, *form a vector space called the* **eigenspace** *of* $\mathbf{A}$ *corresponding to that* λ.

Multiple Eigen Values

Example 2: Find the eigenvalues and eigenvectors of

$$\mathbf{A} = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}.$$

Solution.

For our matrix, the characteristic determinant gives the characteristic equation

$$-\lambda^3 - \lambda^2 + 21\lambda + 45 = 0.$$

The roots (eigenvalues of $\mathbf{A}$) are $\lambda_1 = 5$, $\lambda_2 = \lambda_3 = -3$.

Multiple Eigen Values

Solution. *(continued 1)*

To find eigenvectors, we apply the Gauss elimination (Sec. 7.3) to the system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, first with $\lambda = 5$ and then with $\lambda = -3$. For $\lambda = 5$ the characteristic matrix is

$$\mathbf{A} - \lambda\mathbf{I} = \mathbf{A} - 5\mathbf{I} = \begin{bmatrix} -7 & 2 & -3 \\ 2 & -4 & -6 \\ -1 & -2 & -5 \end{bmatrix}.$$

It row-reduces to
$$\begin{bmatrix} -7 & 2 & -3 \\ 0 & -24/7 & -48/7 \\ 0 & 0 & 0 \end{bmatrix}.$$

Multiple Eigen Values

Solution. *(continued 2)*

Hence it has rank 2. Choosing $x_3 = -1$ we have $x_2 = 2$ from $-\frac{24}{7}x_2 - \frac{48}{7}x_3 = 0$ and then $x_1 = 1$ from $-7x_1 + 2x_2 - 3x_3 = 0.$

Hence an eigenvector of $\mathbf{A}$ corresponding to $\lambda = 5$ is $\mathbf{x}_1 = [1 \quad 2 \quad -1]^T.$

For $\lambda = -3$ the characteristic matrix

$$\mathbf{A} - \lambda\mathbf{I} = \mathbf{A} + 3\mathbf{I} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{bmatrix}$$

row-reduces to

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

Multiple Eigen Values

Solution. *(continued 3)*

Hence it has rank 1.

From $x_1 + 2x_2 - 3x_3 = 0$ we have $x_1 = -2x_2 + 3x_3$. Choosing $x_2 = 1, x_3 = 0$ and $x_2 = 0, x_3 = 1$, we obtain two linearly independent eigenvectors of $\mathbf{A}$ corresponding to $\lambda = -3$ [as they must exist by (5), Sec. 7.5, with rank = 1 and $n = 3$],

$$\mathbf{x}_2 = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad \mathbf{x}_3 = \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}.$$

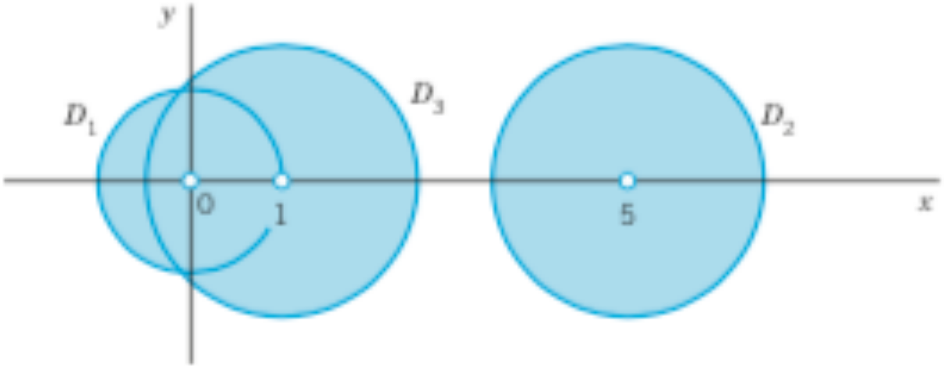
Gerschgorin's Theorem

Theorem gives the bound on Eigenvalues

Every eigenvalue of matrix A_{nxn} satisfies :

$$\lambda - \{a_{ii}\} \leq \sum_{j \neq i} |\{a_{ij}\}|, \quad i = 1,2,\ldots.n$$

Example : $A = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 5 & 1 \\ \frac{1}{2} & 1 & 1 \end{bmatrix}$



We get Gerschgorin disks D_1: Centre 0, radius 1

D_2 : Centre 5, radius 1.5

D_3 : Centre 1, radius 1.5

The centers are main diagonal entries of A. These would be the eigenvalues of A if A were diagonal

Algebraic Multiplicity &Geometric Multiplicity

The order M_λ of an eigenvalue λ as a root of the characteristic polynomial is called the **algebraic multiplicity** of λ. The number m_λ of linearly independent eigenvectors corresponding to λ is called the **geometric multiplicity** of λ. Thus m_λ is the dimension of the eigenspace corresponding to this λ.

Since the characteristic polynomial has degree n, the sum of all the algebraic multiplicities must equal n. In Example 2 for $\lambda = -3$ we have $m_\lambda = M_\lambda = 2$. In general, $m_\lambda \leq M_\lambda$, as can be shown. The difference $\Delta_\lambda = M_\lambda - m_\lambda$ is called the **defect** of λ. Thus $\Delta_{-3} = 0$ in Example 2, but positive defects Δ_λ can easily occur.

Special cases

innovate achieve lead

Theorem 3

Eigenvalues of the Transpose

The transpose $\mathbf{A}^{\mathrm{T}}$ *of a square matrix* $\mathbf{A}$ *has the same eigenvalues as* $\mathbf{A}$.

Basis of Eigenvectors

If an $n \times n$ *matrix* $\mathbf{A}$ *has n* ***distinct*** *eigenvalues, then* $\mathbf{A}$ *has a basis of eigenvectors* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *for* R^n.

Eigenvectors corresponding to Distinct Eigenvalues are Linearly Independent

Let k be the smallest positive integer such that $v_1, v_2, \ldots, v_k$ are linearly independent. If $k = p$, nothing is to be proved.

If $k < p$, then v_{k+1} is a linear combination of $v_1, \ldots, v_k$; that is, there exist constants $c_1, c_2, \ldots, c_k$ such that

$$v_{k+1} = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k.$$

Applying the matrix A to both sides, we have

$$\begin{aligned} Av_{k+1} &= \lambda_{k+1} v_{k+1} \\ &= \lambda_{k+1}(c_1 v_1 + c_2 v_2 + \cdots + c_k v_k) \\ &= c_1 \lambda_{k+1} v_1 + c_2 \lambda_{k+1} v_2 + \cdots + c_k \lambda_{k+1} v_k; \end{aligned}$$

$$\begin{aligned} Av_{k+1} &= A(c_1 v_1 + c_2 v_2 + \cdots + c_k v_k) \\ &= c_1 Av_1 + c_2 Av_2 + \cdots + c_k Av_k \\ &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_k \lambda_k v_k. \end{aligned}$$

Thus $c_1(\lambda_{k+1} - \lambda_1)v_1 + c_2(\lambda_{k+1} - \lambda_2)v_2 + \cdots + c_k(\lambda_{k+1} - \lambda_k)v_k = 0.$

Since $v_1, v_2, \ldots, v_k$ are linearly independent, we have

$c_1(\lambda_{k+1} - \lambda_1) = c_2(\lambda_{k+1} - \lambda_2) = \cdots = c_k(\lambda_{k+1} - \lambda_k) = 0.$

Note that the eigenvalues are distinct. Hence

$$c_1 = c_2 = \cdots = c_k = 0,$$

which implies that v_{k+1} is the zero vector. This is contradictory to $v_{k+1} \neq 0$.

Similarity of Matrices

Similar Matrices. Similarity Transformation

An $n \times n$ matrix $\hat{\mathbf{A}}$ is called **similar** to an $n \times n$ matrix $\mathbf{A}$ if

$$\hat{\mathbf{A}} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \tag{4}$$

for some (nonsingular!) $n \times n$ matrix $\mathbf{P}$. This transformation, which gives $\hat{\mathbf{A}}$ from $\mathbf{A}$, is called a **similarity transformation**.

Eigenvalues and Eigenvectors of Similar Matrices

If $\hat{\mathbf{A}}$ is similar to $\mathbf{A}$, then $\hat{\mathbf{A}}$ has the same eigenvalues as $\mathbf{A}$.
Furthermore, if $\mathbf{x}$ is an eigenvector of $\mathbf{A}$, then $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ is an eigenvector of $\hat{\mathbf{A}}$ corresponding to the same eigenvalue.

Diagonalization of a Matrix

Diagonalization of a Matrix

If an $n \times n$ matrix **A** *has a basis of eigenvectors, then*

(5) $$\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X}$$

is diagonal, with the eigenvalues of **A** *as the entries on the main diagonal. Here* **X** *is the matrix with these eigenvectors as column vectors. Also,*

(5*) $$\mathbf{D}^m = \mathbf{X}^{-1}\mathbf{A}^m\mathbf{X} \qquad (m = 2, 3, \dots).$$

Diagonalization of a Matrix

Diagonalize

$$\mathbf{A} = \begin{bmatrix} 7.3 & 0.2 & -3.7 \\ -11.5 & 1.0 & 5.5 \\ 17.7 & 1.8 & -9.3 \end{bmatrix}.$$

Solution.

The characteristic determinant gives the characteristic equation $-\lambda^3 - \lambda^2 + 12\lambda = 0$. The roots (eigenvalues of $\mathbf{A}$) are $\lambda_1 = 3$, $\lambda_2 = -4$, $\lambda_3 = 0$. By the Gauss elimination applied to $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ with $\lambda = \lambda_1, \lambda_2, \lambda_3$ we find eigenvectors and then $\mathbf{X}^{-1}$ by the Gauss–Jordan elimination

Diagonalization of a Matrix

Solution. (continued 1)

The results are

$$\begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} -1 & 1 & 2 \\ 3 & -1 & 1 \\ -1 & 3 & 4 \end{bmatrix},$$

$$\mathbf{X}^{-1} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix}.$$

Diagonalization of a Matrix

Solution. *(continued 2)*
Calculating **AX** and multiplying by $\mathbf{X}^{-1}$ from the left, we thus obtain

$$\mathbf{D} = \mathbf{X}^{-1}\mathbf{A}\mathbf{X} = \begin{bmatrix} -0.7 & 0.2 & 0.3 \\ -1.3 & -0.2 & 0.7 \\ 0.8 & 0.2 & -0.2 \end{bmatrix} \begin{bmatrix} -3 & -4 & 0 \\ 9 & 4 & 0 \\ -3 & -12 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & -4 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Quadratic Forms
Transformation to Principle Axes

By definition, a **quadratic form** Q in the components $x_1, \ldots, x_n$ of a vector **x** is a sum n^2 of terms, namely,

$$Q = \mathbf{x}^{\mathrm{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k$$

(7)

$$\begin{aligned} &= a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{1n}x_1x_n \\ &\quad + a_{21}x_2x_1 + a_{22}x_2^2 + \cdots + a_{2n}x_2x_n \\ &\quad + \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\quad + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \cdots + a_{nn}x_n^2. \end{aligned}$$

A = $[a_{jk}]$ is called the **coefficient matrix** of the form. We may assume that **A** is ***symmetric***, because we can take off-diagonal terms together in pairs and write the result as a sum of two equal terms; see the following example.

Quadratic Form
Symmetric Coefficient Matrix

Let

$$\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 3x_1^{\,2} + 4x_1x_2 + 6x_2x_1 + 2x_2^{\,2}$$

$$= 3x_1^{\,2} + 10x_1x_2 + 2x_2^{\,2}.$$

Here $4 + 6 = 10 = 5 + 5$.

From the corresponding ***symmetric*** matrix $\mathbf{C} = [c_{jk}]$ where $c_{jk} = \frac{1}{2}(a_{jk} + a_{kj})$, thus $c_{11} = 3$, $c_{12} = c_{21} = 5$, $c_{22} = 2$, we get the same result; indeed,

$$\mathbf{x}^{\mathsf{T}}\mathbf{C}\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= 3x_1^{\,2} + 5x_1x_2 + 5x_2x_1 + 2x_2^{\,2}$$

$$= 3x_1^{\,2} + 10x_1x_2 + 2x_2^{\,2}.$$

Quadratic Form
Symmetric Coefficient Matrix

A *symmetric* matrix **A** of (example 7) has an orthonormal basis of eigenvectors.($v_i.v_j = 0$ if $i \neq j$ and $v_i.v_j = 1$ if $i = j$)
Hence if we take these as column vectors, we obtain a matrix **X** that is orthogonal ($\mathbf{X}^{-1} = \mathbf{X}^{\mathsf{T}}$).
Thus $\mathbf{A} = \mathbf{XDX}^{-1} = \mathbf{XDX}^{\mathsf{T}}$. Substitution into (7) gives

(8) $$Q = \mathbf{x}^{\mathsf{T}}\mathbf{XDX}^{\mathsf{T}}\mathbf{x}.$$

If we set $\mathbf{X}^{\mathsf{T}}\mathbf{x} = \mathbf{y}$, then, since $\mathbf{X}^{-1} = \mathbf{X}^{\mathsf{T}}$, we have $\mathbf{X}^{-1}\mathbf{x} = \mathbf{y}$ and thus obtain

(9) $$\mathbf{x} = \mathbf{Xy}.$$

Furthermore, in (8) we have $\mathbf{x}^{\mathsf{T}}\mathbf{X} = (\mathbf{X}^{\mathsf{T}}\mathbf{x})^{\mathsf{T}} = \mathbf{y}^{\mathsf{T}}$ and $\mathbf{X}^{\mathsf{T}}\mathbf{x} = \mathbf{y}$, so that Q becomes simply

(10) $$Q = \mathbf{y}^{\mathsf{T}}\mathbf{Dy} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \dots + \lambda_n y_n^2.$$

Principal Axes Theorem

Theorem 5

Principal Axes Theorem

The substitution (9) *transforms a quadratic form*

$$Q = \mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}x_j x_k \qquad (a_{kj} = a_{jk})$$

to the principal axes form or **canonical form** (10), *where* $\lambda_1, \ldots,$ λ_n *are the (not necessarily distinct) eigenvalues of the (symmetric!) matrix* **A**, *and* **X** *is an orthogonal matrix with corresponding eigenvectors* $\mathbf{x}_1, \ldots, \mathbf{x}_n$, *respectively, as column vectors.*

Transformation to Principal Axes Conic Sections

innovate achieve lead

Find out what type of conic section the following quadratic form represents and transform it to principal axes:

$$Q = 17x_1^2 - 30x_1x_2 + 17x_2^2 = 128.$$

Solution. We have $Q = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}$, where

$$\mathbf{A} = \begin{bmatrix} 17 & -15 \\ -15 & 17 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Transformation to Principal Axes
Conic Sections

Solution. **(continued 1)**

This gives the characteristic equation $(17 - \lambda)^2 - 15^2 = 0$. It has the roots $\lambda_1 = 2$, $\lambda_2 = 32$. Hence (10) becomes

$$Q = 2{y_1}^2 + 32{y_2}^2.$$

We see that $Q = 128$ represents the ellipse $2{y_1}^2 + 32{y_2}^2 = 128$, that is,

$$\frac{{y_1}^2}{8^2} + \frac{{y_2}^2}{2^2} = 1.$$

Transformation to Principal Axes Conic Sections

Solution. **(continued 2)**

If we want to know the direction of the principal axes in the x_1x_2-coordinates, we have to determine normalized eigenvectors from $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ with $\lambda = \lambda_1 = 2$ and $\lambda = \lambda_2 = 32$ and then use (9). We get

$$\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Observe that the above vectors are orthonormal

Transformation to Principal Axes
Conic Sections

Solution. **(continued 3)**

hence

$$\mathbf{x} = \mathbf{X}\mathbf{y} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \qquad \begin{aligned} x_1 &= y_1/\sqrt{2} - y_2/\sqrt{2} \\ x_2 &= y_1/\sqrt{2} + y_2/\sqrt{2}. \end{aligned}$$

This is a 45° rotation.