Lecture 10

- Eigenvalues, eigenvectors and diagonalisation
- Orthogonal diagonalisation of symmetric matrices
- Spectral theorem for (square) symmetric matrices
- Quadratic Forms

Eigenvectors and Eigenvalues

We have seen that the linear transformation $T(\mathbf{x}) : \mathbf{x} \to \mathbf{A}\mathbf{x}$ changes the direction and the length of the vector $\mathbf{x}$. However there are cases when the action of the mapping is simple. Let

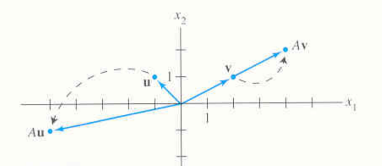
$$\mathbf{A} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}, \ \mathbf{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \ \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

The image of the vector $\mathbf{u}$

$$\mathbf{A}\mathbf{u} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \end{bmatrix}.$$

The image of the vector $\mathbf{v}$ is

$$\mathbf{A}\mathbf{v} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2\mathbf{v}.$$



Eigenvectors and Eigenvalues

Definition:

- An **eigenvector** of an $n \times n$ matrix $\mathbf{A}$ is nonzero vector $\mathbf{x}$ such that $\mathbf{Ax} = \lambda\mathbf{x}$ for some scalar λ.
- A scalar λ is called an **eigenvalue** of $\mathbf{A}$ if there is a nontrivial solution $\mathbf{x}$ of $\mathbf{Ax} = \lambda\mathbf{x}$.
- Such an $\mathbf{x}$ is called an **eigenvector corresponding to** λ.

  λ is an eigenvalue for $\mathbf{A}$ if and only if the equation

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

  has a nontrivial solution for $\mathbf{x}$. This happens if and only if

  $$\det(\mathbf{A} - \lambda\mathbf{I}) = 0 \quad \text{(Characteristic equation)}$$

- The characteristic equation is a polynomial equation (it can have, repeated roots and complex roots).
- The set of all solutions for $\mathbf{x}$ is the null space of the matrix $(\mathbf{A} - \lambda\mathbf{I})$.
- This set is a subspace of $\mathbb{R}^n$ and is called the **eigenspace** of $\mathbf{A}$ corresponding to λ.

Finding Eigenvectors and Eigenvalues

- Calculate the determinant of matrix $\mathbf{A} - \lambda\mathbf{I}$ and equate it to zero
$$\det[\mathbf{A} - \lambda\mathbf{I}] = 0$$
and obtain characteristic equation, which is a polynomial of degree n
- Matrix $\mathbf{I}$ is an Identity matrix with the same sizes as matrix $\mathbf{A}[n \times n]$
- Find the n roots of this polynomial. These roots λ_i are eigenvalues of matrix $\mathbf{A}$
- For each eigenvalue find the solution of homogeneous system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$
- The determinant of matrix $\mathbf{A} - \lambda\mathbf{I}$ is zero, so there are non-trivial solutions for linear system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$. These vectors are eigenvectors for matrix $\mathbf{A}$

Eigenvectors and Eigenvalues

Example: Let $\mathbf{A} = \begin{pmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{pmatrix}$. Given that

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (9-\lambda)(2-\lambda)(2-\lambda),$$

find a basis for the eigenspace corresponding to the smallest eigenvalue.

Solution: $\lambda = 2$ is eigenvalue of *multiplicity* 2. Solving $(\mathbf{A} - 2\mathbf{I})\mathbf{v} = \mathbf{0}$ yields the general solution:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$$

A basis for the eigenspace is

$$\left\{ \begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Eigenvectors and Eigenvalues

Theorem: The eigenvalues of a triangular matrix are the entries on its main diagonal.

- **Example:**

$$\mathbf{A} - \lambda \mathbf{I} = \begin{pmatrix} a_{11} - \lambda & a_{12} & a_{13} \\ 0 & a_{22} - \lambda & a_{23} \\ 0 & 0 & a_{33} - \lambda \end{pmatrix}$$

$$\det(\mathbf{A} - \lambda \mathbf{I}) = (a_{11} - \lambda)(a_{22} - \lambda)(a_{33} - \lambda) = 0$$

  So the eigenvalues are $\lambda_1 = a_{11}$ $\lambda_2 = a_{22}$, $\lambda_3 = a_{33}$.
- **Question:** What does it mean when the eigenvalue is 0? This can happen if

$$\mathbf{Ax} = 0\mathbf{x} = \mathbf{0}$$

  has a nontrivial solution, which happens if and only if **A** is not invertible. Thus 0 is an eigenvalue of **A** if and only if **A** is not invertible, so $\det \mathbf{A} = 0$.
- $\det \mathbf{A} = \prod_{\mathbf{i=1}}^{\mathbf{n}} \lambda_{\mathbf{i}} = \lambda_{\mathbf{1}} \lambda_{\mathbf{2}} \ldots \lambda_{\mathbf{n}}$
- $\lambda_1 + \ldots + \lambda_n = a_{11} + \ldots + a_{nn} = tr(\mathbf{A})$, $tr\mathbf{A}$ is trace of **A**.

Eigenvectors and Eigenvalues

Theorem: If $\mathbf{v}_1, \ldots, \mathbf{v}_r$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of an $n \times n$ matrix $\mathbf{A}$, then $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly independent.

Proof: (by contradiction.) Suppose $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ is linearly dependent. Choose p such that $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is linearly independent, but $\{\mathbf{v}_1, \ldots, \mathbf{v}_{p+1}\}$ is linearly dependent. Then

$$\mathbf{v}_{p+1} = c_1\mathbf{v}_1 + \ldots + c_p\mathbf{v}_p. \tag{1}$$

Multiplying both sides by $\mathbf{A}$ and using $\mathbf{A}\mathbf{v}_\mathbf{k} = \lambda_k\mathbf{v}_k$ we obtain

$$\mathbf{A}\mathbf{v}_{\mathbf{p+1}} = c_1\mathbf{A}\mathbf{v}_\mathbf{1} + \ldots + c_p\mathbf{A}\mathbf{v}_\mathbf{p}, \tag{2}$$

$$\text{or} \qquad \lambda_{p+1}\mathbf{v}_{p+1} = c_1\lambda_1\mathbf{v}_1 + \ldots + c_p\lambda_p\mathbf{v}_p. \tag{3}$$

We multiply the relation (1) by λ_{p+1}:

$$\lambda_{p+1}\mathbf{v}_{p+1} = c_1\lambda_{p+1}\mathbf{v}_1 + \ldots + c_p\lambda_{p+1}\mathbf{v}_p. \tag{4}$$

Subtracting (4) from (3) we have

$$c_1(\lambda_1 - \lambda_{p+1})\mathbf{v}_1 + \ldots + c_p(\lambda_p - \lambda_{p+1})\mathbf{v}_p = \mathbf{0}.$$

None of the $\lambda_i - \lambda_{p+1} = 0$ since the λ_i are distinct. Hence $c_i = 0$ for $i = 1, \ldots, p$ and therefore $\mathbf{v}_{p+1} = \mathbf{0}$. This is impossible therefore $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ must be linearly independent.

Similarity

If $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices then $\mathbf{A}$ is **similar** to $\mathbf{B}$ if there is an invertible matrix $\mathbf{P}$ such that

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P} \text{ or } \mathbf{A} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1}.$$

We say $\mathbf{A}$ and $\mathbf{B}$ are **similar** matrices and changing $\mathbf{A}$ into $\mathbf{B} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ is called a **similarity transformation**.

Theorem : If matrices $\mathbf{A}$ and $\mathbf{B}$ are similar then they have the same characteristic polynomial and the same eigenvalues.
Proof: Let $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. So $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{P}^{-1}\mathbf{x} = \lambda\mathbf{P}^{-1}\mathbf{x}$ and $\mathbf{B}\mathbf{P}^{-1}\mathbf{x} = \lambda\mathbf{P}^{-1}\mathbf{x}$, therefore $\mathbf{B}\mathbf{y} = \lambda\mathbf{y}$, where $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$.

A square matrix $\mathbf{A}$ is said to be **diagonalizable** if $\mathbf{A}$ is similar to a diagonal matrix $\mathbf{D}$.

This means $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ for some invertible matrix $\mathbf{P}$ and some diagonal matrix $\mathbf{D}$.

Diagonalizing Matrices

Theorem: The Diagonalization Theorem

An $n \times n$ matrix $\mathbf{A}$ is diagonalizable if and only if $\mathbf{A}$ has n linearly independent eigenvectors.

In particular, $\mathbf{A} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is diagonal matrix, if and only if the columns of $\mathbf{P}$ are n linearly independent eigenvectors of $\mathbf{A}$.

The diagonal entries of $\mathbf{D}$ are eigenvalues of $\mathbf{A}$ that correspond, respectively, to the eigenvectors in $\mathbf{P}$.

Proof: Assume $\mathbf{A}$ is diagonalizable, therefore $\mathbf{AP} = \mathbf{PD}$ for some matrix $\mathbf{P}$ with independent columns $\mathbf{v_1}, \mathbf{v_2}, \ldots \mathbf{v_n}$ and a diagonal matrix $\mathbf{D}$. Therefore $\mathbf{AP} = [\mathbf{Av_1}\ \mathbf{Av_2}\ \ldots \mathbf{Av_n}]$ and $\mathbf{PD} = [\lambda_\mathbf{1}\mathbf{v_1}\ \lambda_\mathbf{2}\mathbf{v_2} \ldots \lambda_\mathbf{n}\mathbf{v_n}]$. So $\mathbf{Av_1} = \lambda_1\mathbf{v_1}$, $\mathbf{Av_2} = \lambda_2\mathbf{v_2}$, $\ldots$, $\mathbf{Av_n} = \lambda_n\mathbf{v_n}$. Therefore λ_i and $\mathbf{v_i}$ are eigenvalues and eigenvectors of $\mathbf{A}$.

Notes

- We can also say that $\mathbf{A}$ is diagonalizable if and only if there are enough eigenvectors to form a basis of $\mathbb{R}^n$.
- We call such a basis an **eigenvector basis**.

Diagonalizing Matrices

Theorem: An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

Note: This is only a sufficient condition. It is not necessary for an $n \times n$ matrix to have n distinct eigenvalues in order to be diagonalizable.

Example: Determine whether the following matrix is diagonalizable:

$$\mathbf{A} = \begin{pmatrix} 5 & -8 & 1 \\ 0 & 0 & 7 \\ 0 & 0 & -2 \end{pmatrix}$$

Solution: This is a triangular matrix and the eigenvalues are $5, 0, -2$. The eigenvalues are distinct and $\mathbf{A}$ is diagonalizable.

Diagonalizing Matrices

Theorem: Let **A** be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \ldots, \lambda_p$.
(a) For $1 \leq k \leq p$ the dimension of the eigenspace for λ_k is less than or equal to the multiplicity of the eigenvalue λ_k.

(b) The matrix **A** is diagonalizable if and only if the sum of the dimensions of the eigenspaces for distinct eigenvalues equals n.

This happens only if the dimension of the eigenspace for each λ_k equals the multiplicity of λ_k. *Algebraic multiplicity of an eigenvalue λ equals to geometric multiplicity of its eigenvector*.

(c) If **A** is diagonalizable and $\mathcal{B}_k$ is a basis for the eigenspace corresponding to λ_k for each k then the total collection of vectors in the sets $\mathcal{B}_1, \ldots, \mathcal{B}_p$ forms an eigenvector basis for $\mathbb{R}^n$.

Symmetric Matrices

Symmetric Matrices

A matrix $\mathbf{A}$ is symmetric if $\mathbf{A}^{\mathsf{T}} = \mathbf{A}$. Such a matrix is always a square matrix $n \times n$. The main diagonal entries are arbitrary but its other entries occur in pairs: $a_{ij} = a_{ji}$.

Example: Symmetric matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & -3 \end{pmatrix}, \begin{pmatrix} 0 & -1 & 0 \\ -1 & 5 & 8 \\ 0 & 8 & 7 \end{pmatrix}, \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}.$$

Nonsymmetric matrices

$$\begin{pmatrix} 1 & -3 \\ 3 & -3 \end{pmatrix}, \begin{pmatrix} 1 & -4 & 0 \\ -6 & 1 & -4 \\ 0 & -6 & 1 \end{pmatrix}, \begin{pmatrix} 5 & 4 & 3 & 2 \\ 4 & 3 & 2 & 1 \\ 3 & 2 & 1 & 0 \end{pmatrix}.$$

Example: If possible, diagonalize the matrix

$$\mathbf{A} = \begin{pmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{pmatrix}.$$

Solution: First, we find eigenvalues and corresponding eigenvectors. The characteristic equation for $\mathbf{A}$ is

$$0 = \det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 6-\lambda & -2 & -1 \\ -2 & 6-\lambda & -1 \\ -1 & -1 & 5-\lambda \end{vmatrix}$$
$$= -\lambda^3 + 17\lambda^2 - 90\lambda + 144$$
$$= -(\lambda - 8)(\lambda - 6)(\lambda - 3),$$

so the eigenvalues are $\lambda_1 = 8, \lambda_2 = 6, \lambda_3 = 3$. For $\lambda_1 = 8$ we have

$$(\mathbf{A} - \lambda_1 I)\mathbf{x} = \mathbf{0}$$

$$\Leftrightarrow \begin{pmatrix} 6-\lambda_1 & -2 & -1 \\ -2 & 6-\lambda_1 & -1 \\ -1 & -1 & 5-\lambda_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\text{or} \quad \begin{pmatrix} 6-8 & -2 & -1 \\ -2 & 6-8 & -1 \\ -1 & -1 & 5-8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The augmented matrix for this system is

$$\begin{pmatrix} -2 & -2 & -1 & 0 \\ -2 & -2 & -1 & 0 \\ -1 & -1 & -3 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 3 & 0 \\ 0 & 0 & -5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ (\textit{echelon form})}$$

or

$$\begin{pmatrix} 1 & 1 & 3 \\ 0 & 0 & -5 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

It follows that $x_3 = 0$ and $x_1 + x_2 + 3x_3 = 0$ and the solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_2 \\ 0 \end{pmatrix} = x_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

For simplicity choose $x_2 = 1$ and the first eigenvector is then

$$\mathbf{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Symmetric Matrices

Using the same approach for the rest of the eigenvectors we obtain

$$\lambda_2 = 6, \ \ \mathbf{v_2} = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}, \ \ \lambda_3 = 3, \ \ \mathbf{v_3} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Recall for completeness that $\lambda_1 = 8, \ \ \mathbf{v_1} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$. These three vectors are linearly independent and orthogonal $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$, $\mathbf{v}_1 \cdot \mathbf{v}_3 = 0$, $\mathbf{v}_3 \cdot \mathbf{v}_2 = 0$. They are a basis for $\mathbb{R}^3$ and are columns for the matrix $\mathbf{P} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3]$ that diagonalizes $\mathbf{A}$.
It will often be more useful to normalize the $\mathbf{v}_i$ to have an orthonormal basis. We scale the eigenvectors $\mathbf{v}_i$ by the inverses of:

$$\begin{aligned} \|\mathbf{v}_1\| &= \sqrt{1^2 + 1^2} = \sqrt{2}, \\ \|\mathbf{v}_2\| &= \sqrt{(-1)^2 + (-1)^2 + 4} = \sqrt{6}, \\ \|\mathbf{v}_3\| &= \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}, \ldots \end{aligned}$$

Symmetric Matrices

...and obtain

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix},$$

$$\mathbf{u}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \begin{pmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix}, \ \mathbf{u}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}.$$

The corresponding matrices $\mathbf{P}$ and $\mathbf{D}$ are:

$$P = \begin{pmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}, \ D = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

and $\mathbf{A} = \mathbf{PDP^{-1}}$ as usual.

Note: here $\mathbf{P}$ is square and has orthonormal columns, that is, $\mathbf{P}$ is orthogonal matrix and so $\mathbf{P^{-1}} = \mathbf{P^T}$ and in fact

$$\mathbf{A} = \mathbf{PDP^T}.$$

Symmetric Matrices

Theorem: If $\mathbf{A}$ is symmetric, then any two eigenvectors from different eigenspaces (that is, associated with distinct eigenvalues) are orthogonal.

Proof: Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be eigenvectors that correspond to distinct eigenvalues λ_1, λ_2. We must prove that $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. Now,

$$\begin{aligned}
\lambda_1 \mathbf{v}_1 \cdot \mathbf{v}_2 &= (\lambda_1 \mathbf{v}_1)^T \mathbf{v}_2 \\
&= (A\mathbf{v}_1)^T \mathbf{v}_2 \\
&= (\mathbf{v}_1^T A^T)\mathbf{v}_2 \\
&= \mathbf{v}_1{}^T (A\mathbf{v}_2) \\
&= \mathbf{v}_1{}^T \lambda_2 \mathbf{v}_2 \\
&= \lambda_2 \mathbf{v}_1^T \mathbf{v}_2 \\
&= \lambda_2 \mathbf{v}_1 \cdot \mathbf{v}_2.
\end{aligned}$$

Therefore $(\lambda_1 - \lambda_2)\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$. But $\lambda_1 - \lambda_2 \neq 0$ hence $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

Symmetric Matrices

Definition: A matrix $\mathbf{A}$ is said to be **orthogonally diagonalizable** if there is an orthogonal matrix $\mathbf{P}$ ($\mathbf{P}^{-1} = \mathbf{P}^{\mathsf{T}}$) and a diagonal matrix $\mathbf{D}$ such that

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{\mathsf{T}} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}.$$

Note: To orthogonally diagonalize an $n \times n$ matrix we must find n linearly independent and orthonormal eigenvectors.

If $\mathbf{A}$ is orthogonally diagonalizable then

$$\mathbf{A}^{\mathsf{T}} = (\mathbf{P}\mathbf{D}\mathbf{P}^{\mathsf{T}})^{\mathsf{T}} = (\mathbf{P}^{\mathsf{T}})^{\mathsf{T}}\mathbf{D}^{\mathsf{T}}\mathbf{P}^{\mathsf{T}} = \mathbf{P}\mathbf{D}\mathbf{P}^{\mathsf{T}} = \mathbf{A}.$$

Thus $\mathbf{A}$ is symmetric.

Theorem: An $n \times n$ matrix $\mathbf{A}$ is orthogonally diagonalizable if and only if $\mathbf{A}$ is a symmetric matrix.

Note: In general it is impossible to say if a matrix is diagonalizable. A symmetric matrix is always diagonalizable.

Symmetric Matrices

Example: Orthogonally diagonalize the matrix

$$\mathbf{A} = \begin{pmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{pmatrix}.$$

Solution: The characteristic equation of this matrix is

$$0 = -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda - 7)^2(\lambda + 2).$$

Eigenvalues are $\lambda_{1,2} = 7$ (with multiplicity 2) and $\lambda_3 = -2$.
Solving $(A - \lambda_1 I)\mathbf{x} = (A - 7I)\mathbf{x} = 0$,

$$\begin{pmatrix} 3-\lambda_1 & -2 & 4 \\ -2 & 6-\lambda_1 & 2 \\ 4 & 2 & 3-\lambda_1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

the augmented matrix is

$$\begin{pmatrix} -4 & -2 & 4 & 0 \\ -2 & -1 & 2 & 0 \\ 4 & 2 & -4 & 0 \end{pmatrix} \sim \begin{pmatrix} -2 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \ldots$$

Symmetric Matrices

There are two free variables x_2, x_3 and the solution is obtained from $-2x_1 - x_2 + 2x_3 = 0$, or $x_1 = -x_2/2 + x_3$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2/2 + x_3 \\ x_2 \\ x_3 \end{pmatrix} = x_2 \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Thus there are two linearly independent eigenvectors associated with $\lambda_{1,2} = 7$:

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \ \mathbf{v}_2 = \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix}.$$

These eigenvectors are independent but not orthogonal. We use the Gram-Schmidt process to orthogonalize them.

$$\mathbf{z}_2 = \mathbf{v}_2 - \frac{\mathbf{v}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix} - \frac{-1/2}{2} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/4 \\ 1 \\ 1/4 \end{pmatrix}.$$

The set $\{\mathbf{v}_1, \mathbf{z}_2\}$ is an orthogonal basis for the eigenspace associated with $\lambda = 7$.
Now we normalize the vectors $\{\mathbf{v}_1, \mathbf{z}_2\}$:

$$\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} = \begin{pmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \mathbf{u}_2 = \frac{\mathbf{z}_2}{\|\mathbf{z}_2\|} = \begin{pmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{pmatrix}.$$

The eigenvector associated with $\lambda_3 = -2$ is

$$\mathbf{v}_3 = \begin{pmatrix} -1 \\ -1/2 \\ 1 \end{pmatrix}, \ \mathbf{u}_3 = \frac{2\mathbf{v}_3}{\|2\mathbf{v}_3\|} = \begin{pmatrix} -2/3 \\ -1/3 \\ 2/3 \end{pmatrix}.$$

Hence $\mathbf{P}$ and $\mathbf{D}$ are:

$$\mathbf{P} = (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3) = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{pmatrix}, \quad D = \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

and $\mathbf{P}$ orthogonally diagonalizes A: $\mathbf{A} = \mathbf{PDP}^{-1} = \mathbf{PDP}^{\mathsf{T}}$.

The Spectral Theorem

The set of eigenvalues of a matrix $\mathbf{A}$ is sometimes called the **spectrum** of $\mathbf{A}$.

Theorem: The Spectrum Theorem for Symmetric Matrices

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

(a) $\mathbf{A}$ 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 equation;

(c) The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal;

(d) $\mathbf{A}$ is orthogonally diagonalizable.

Spectral Decomposition

Let $\mathbf{A} = \mathbf{PDP^{-1}}$, where the columns of $\mathbf{P}$ are orthonormal eigenvectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ and $\mathbf{D}$ is the diagonal matrix of eigenvalues $\lambda_1, \ldots, \lambda_n$. Since $\mathbf{P^{-1}} = \mathbf{P^T}$ we have

$$\begin{aligned}
\mathbf{A} = \mathbf{PDP^T} &= \begin{pmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \begin{pmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{pmatrix} \\
&= \begin{pmatrix} \lambda_1 \mathbf{u}_1 & \ldots & \lambda_n \mathbf{u}_n \end{pmatrix} \begin{pmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T \end{pmatrix} \\
&= \lambda_1 \mathbf{u}_1 \mathbf{u}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^T + \ldots + \lambda_n \mathbf{u}_n \mathbf{u}_n^T.
\end{aligned}$$

This representation of $\mathbf{A}$ is called a **spectral decomposition** of $\mathbf{A}$. Each term in the decomposition is an $n \times n$ matrix with rank 1. Each matrix $\mathbf{u}_j \mathbf{u}_j^T$ is a projection matrix given that for each $\mathbf{x} \in \mathbb{R}^n$ the vector $\mathbf{u}_j \mathbf{u}_j^T \mathbf{x}$ is the orthogonal projection of $\mathbf{x}$ into the subspace spanned by $\mathbf{u}_j$.

Spectral Decomposition

Example: Construct a spectral decomposition of the matrix $\mathbf{A}$ with the orthogonal diagonalization

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

Solution: We denote the columns of P by $\mathbf{u}_1$ and $\mathbf{u}_2$. Then

$$\mathbf{A} = 8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T.$$

To verify this decomposition we calculate

$$\mathbf{u}_1\mathbf{u}_1^T = \begin{pmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{pmatrix} \begin{pmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{pmatrix} = \begin{pmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{pmatrix}$$

$$\mathbf{u}_2\mathbf{u}_2^T = \begin{pmatrix} \frac{-1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{pmatrix}$$

and . . .

$$8\mathbf{u}_1\mathbf{u}_1^T + 3\mathbf{u}_2\mathbf{u}_2^T = \begin{pmatrix} 32/5 & 16/5 \\ 16/5 & 8/5 \end{pmatrix} + \begin{pmatrix} 3/5 & -6/5 \\ -6/5 & 12/5 \end{pmatrix} = \begin{pmatrix} 7 & 2 \\ 2 & 4 \end{pmatrix}.$$

Quadratic Forms

Quadratic Forms

Definition: A **quadratic form** on $\mathbb{R}^n$ is a function Q defined on $\mathbb{R}^n$ whose values at $\mathbf{x}$ can be computed from

$$Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x},$$

where $\mathbf{A}$ is an $n \times n$ symmetric matrix called the **matrix of the quadratic form**.

The simplest quadratic form is

$$Q(\mathbf{x}) = \mathbf{x}^T\mathbf{I}\mathbf{x} = \mathbf{x}^{\mathbf{T}}\mathbf{x} = \|\mathbf{x}\|^2.$$

Example: Let

$$\mathbf{A} = \begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 3 & -2 \\ -2 & 7 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Calculate $\mathbf{x}^T\mathbf{A}\mathbf{x}$ and $\mathbf{x}^T\mathbf{B}\mathbf{x}$.

Solution:

$$\begin{aligned} \mathbf{x}^T\mathbf{A}\mathbf{x} &= (x_1 \quad x_2)\begin{pmatrix} 4 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 \quad x_2)\begin{pmatrix} 4x_1 \\ 3x_2 \end{pmatrix} \\ &= 4x_1^2 + 3x_2^2. \end{aligned}$$

Quadratic Forms

$$\mathbf{B} = \begin{pmatrix} 3 & -2 \\ -2 & 7 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$\begin{aligned} \mathbf{x}^{\mathbf{T}}\mathbf{B}\mathbf{x} &= (x_1 \quad x_2) \begin{pmatrix} 3 & -2 \\ -2 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= (x_1 \quad x_2) \begin{pmatrix} 3x_1 - 2x_2 \\ -2x_1 + 7x_2 \end{pmatrix} \\ &= x_1(3x_1 - 2x_2) + x_2(-2x_1 + 7x_2) \\ &= 3x_1^2 - 4x_1x_2 + 7x_2^2. \end{aligned}$$

Quadratic Forms

Example: For $\mathbf{x}$ in $\mathbb{R}^3$ let

$$Q(x) = 5x_1^2 + 3x_2^2 + 2x_3^2 - x_1x_2 + 8x_2x_3.$$

Write this quadratic form as $\mathbf{x}^T\mathbf{A}\mathbf{x}$.

Solution: The coefficients of x_1^2, x_2^2, x_3^2 go on the diagonal of $\mathbf{A}$. To make $\mathbf{A}$ symmetric we split the coefficient of x_ix_j between the i, j and j, i elements.

$$\begin{aligned} Q(\mathbf{x}) &= \mathbf{x}^T\mathbf{A}\mathbf{x} \\ &= (x_1\ x_2\ x_3)\begin{pmatrix} 5 & -1/2 & 0 \\ -1/2 & 3 & 4 \\ 0 & 4 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \end{aligned}$$

Quadratic Forms

Note: since $\mathbf{A}$ is symmetric, $\mathbf{A} = \mathbf{PDP^T}$ (for suitable $\mathbf{D}$, $\mathbf{P}$) is an orthogonal diagonalisation of $\mathbf{A}$. Now introduce a change of basis (variable) given by

$$\mathbf{x} = \mathbf{Py}, \;\; \mathbf{y} = \mathbf{P}^{-\mathbf{1}}\mathbf{x} = \mathbf{P^T x}.$$

Then $\mathbf{y}$ is the coordinate vector of $\mathbf{x}$ relative to the basis given by the columns of $\mathbf{P}$, and

$$\mathbf{x}^T\mathbf{Ax} = (\mathbf{Py})^T\mathbf{A}(\mathbf{Py}) = \mathbf{y}^T\mathbf{P^T}\mathbf{APy} = \mathbf{y}^T\mathbf{Dy}$$

and the matrix in the new quadratic form is diagonal.

Example: Make a change of variables in the quadratic form

$$Q(\mathbf{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$$

to eliminate the cross term.

Solution: The matrix of the quadratic form is $\mathbf{A} = \begin{pmatrix} 1 & -4 \\ -4 & -5 \end{pmatrix}$.
First orthogonally diagonalize $\mathbf{A}$. The eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = -7$. The unit eigenvectors are:

$$\lambda_1 = 3, \quad \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}; \quad \lambda_2 = -7, \quad \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}.$$

These vectors are orthogonal because $\mathbf{A}$ is symmetric and they correspond to distinct eigenvalues. Then

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

The change of variable is $\mathbf{x} = \mathbf{P}\mathbf{y}$. Then

$$\begin{aligned} x_1^2 - 8x_1x_2 - 5x_2^2 &= \mathbf{x}^{\mathbf{T}}\mathbf{A}\mathbf{x} = (\mathbf{P}\mathbf{y})^{\mathbf{T}}\mathbf{A}(\mathbf{P}\mathbf{y}) \\ &= \mathbf{y}^T\mathbf{D}\mathbf{y} = 3y_1^2 - 7y_2^2, \end{aligned}$$

where $\mathbf{x} = \mathbf{P}\mathbf{y}$, $\mathbf{y} = \mathbf{P}^{-1}\mathbf{x}$ and $\mathbf{D} = \mathbf{P}^{\mathbf{T}}\mathbf{A}\mathbf{P}$.

Theorem: The Principal Axes Theorem
Let $\mathbf{A}$ be an $n \times n$ matrix. Then there is an orthogonal change of variable $\mathbf{x} = \mathbf{P}\mathbf{y}$ that transforms the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$ into a quadratic form $\mathbf{y}^T\mathbf{D}\mathbf{y}$ with no cross-product terms.

The columns of $\mathbf{P}$ are called the **principal axes** of the quadratic form $\mathbf{x}^T\mathbf{A}\mathbf{x}$.

The vector $\mathbf{y}$ is the coordinate vector of $\mathbf{x}$ relative to the orthonormal basis of $\mathbb{R}^n$ given by these principal axes.

It can be shown that the set of all $\mathbf{x}$ in $\mathbb{R}^2$ that satisfy

$$\mathbf{x}^T\mathbf{A}\mathbf{x} = c$$

corresponds to an ellipse, hyperbola, parabola, two intersecting lines, a single point or no point at all.

Quadratic Forms

If **A** is diagonal then the graph is in the standard position

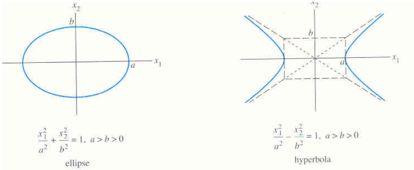


FIGURE 2 An ellipse and a hyperbola in standard position.

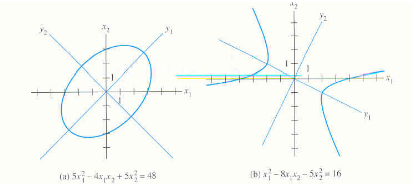
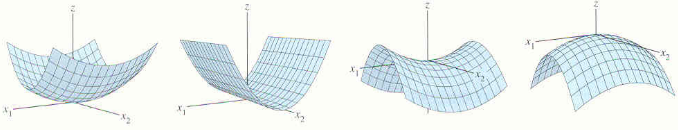


FIGURE 3 An ellipse and a hyperbola *not* in standard position.

$$\mathbf{A} = \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 5 & -2 \\ -2 & -5 \end{pmatrix}$$

Quadratic Forms

Here we plot function $z = Q(\mathbf{x})$ for some typical example matrices $\mathbf{A}$.



(a) $z = 3x_1^2 + 7x_2^2$ (b) $z = 3x_1^2$ (c) $z = 3x_1^2 - 7x_2^2$ (d) $z = -3x_1^2 - 7x_2^2$

FIGURE 4 Graphs of quadratic forms.

The simple 2×2 examples in Fig. 4 illustrate the following definitions.

Definition: A quadratic form Q is

(a) **positive definite** if $Q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$
(b) **negative definite** if $Q(\mathbf{x}) < 0$ for all $\mathbf{x} \neq 0$
(c) **indefinite** if $Q(\mathbf{x})$ assumes positive and negative values
(d) **positive semidefinite** if $Q(\mathbf{x}) \geq 0$ for all $\mathbf{x}$
(e) **negative semidefinite** if $Q(\mathbf{x}) \leq 0$ for all $\mathbf{x}$

Quadratic Forms

Theorem: Quadratic Forms and Eigenvalues
Let $\mathbf{A}$ be an $n \times n$ symmetric matrix. Then a quadratic form $Q(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ is

(a) positive definite if and only if the eigenvalues of $\mathbf{A}$ are all positive.

(b) negative definite if and only if the eigenvalues of $\mathbf{A}$ are all negative.

(c) indefinite if and only if $\mathbf{A}$ has both positive and negative eigenvalues.

Constrained Optimization

It is often necessary to find the maximum or minimum values of a quadratic form $Q(\mathbf{x})$ for the set of $\mathbf{x}$ given by $\|\mathbf{x}\| = 1$.

Example: Find maximum and minimum values of $Q(\mathbf{x}) = 9x_1^2 + 4x_2^2 + 2x_3^2$ subject to the condition $\mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2 = 1$.

Solution:

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 2x_3^2 \\ &\leq 9x_1^2 + 9x_2^2 + 9x_3^2 \\ &= 9(x_1^2 + x_2^2 + x_3^2) \\ &= 9 \end{aligned}$$

So the maximum value of $Q(\mathbf{x}) = 9$. This is the case when $\mathbf{x} = (1, 0, 0)$.

Constrained Optimization

The matrix for this quadratic form is $\mathbf{A} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$. The largest eigenvalue is 9 and the corresponding eigenvector is $\mathbf{v}_1 = \mathbf{x} = (1, 0, 0)$.

Thus the quadratic form is maximal along the eigenvector with largest eigenvalue given the constraint $\|\mathbf{x}\| = 1$.

Constrained Optimization

To find the minimum we consider

$$\begin{aligned} Q(\mathbf{x}) &= 9x_1^2 + 4x_2^2 + 2x_3^2 \\ &\geq 2x_1^2 + 2x_2^2 + 2x_3^2 \\ &= 2(x_1^2 + x_2^2 + x_3^2) \\ &= 2, \end{aligned}$$

so the minimum value of $Q(\mathbf{x}) = 2$. This value is achieved when $\mathbf{x} = (0, 0, 1)$. Recall:

$$\mathbf{A} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

The smallest eigenvalue is 2 and the corresponding eigenvector is $\mathbf{v}_3 = \mathbf{x} = (0, 0, 1)$.

Thus the quadratic form is minimal along the eigenvector with smallest eigenvalue given the constraint $\|\mathbf{x}\| = 1$.

Constrained Optimization

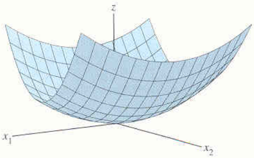


FIGURE 1 $z = 3x_1^2 + 7x_2^2$.

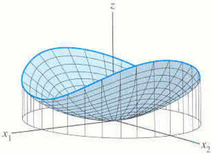


FIGURE 2 The intersection of $z = 3x_1^2 + 7x_2^2$ and the cylinder $x_1^2 + x_2^2 = 1$.

$$\|\mathbf{x}\| = 1$$

The values of $Q(\mathbf{x})$ satisfy $2 \leq Q(\mathbf{x}) \leq 7$.
The set of all possible values of $\mathbf{x}^T A\mathbf{x}$ for $\|\mathbf{x}\| = 1$ is a closed set.

Constrained Optimization

Theorem: Let $\mathbf{A}$ be an $n \times n$ symmetric matrix and let

$$m = \min\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\}$$
$$M = \max\{\mathbf{x}^T A\mathbf{x} : \|\mathbf{x}\| = 1\}.$$

Then M is the greatest eigenvalue λ_1 of $\mathbf{A}$ and m is the smallest eigenvalue λ_n of $\mathbf{A}$.

The value of $\mathbf{x}^T\mathbf{A}\mathbf{x}$ is $M = \lambda_1$ when $\mathbf{x}$ is oriented along the corresponding eigenvector $\mathbf{u}_1$.

The value of $\mathbf{x}^T\mathbf{A}\mathbf{x}$ is $m = \lambda_n$ when $\mathbf{x}$ is oriented along the corresponding eigenvector $\mathbf{u}_n$.

Theorem: Let $\mathbf{A}$, λ_1 and $\mathbf{u}_1$ be as in Theorem (see above). The maximum value of $\mathbf{x}^T A\mathbf{x}$ subject to the constraints

$$\mathbf{x}^T\mathbf{x} = 1, \;\; \mathbf{x}^T\mathbf{u}_1 = 0$$

is the second greatest eigenvalue λ_2 and this maximum is attained when $\mathbf{x}$ is an eigenvector $\mathbf{u}_2$ corresponding to λ_2.

Constrained Optimization

Example: Find the maximum value of

$$9x_1^2 + 4x_2^2 + 3x_3^2$$

subject to the constraints $\mathbf{x}^T\mathbf{x} = 1$ and $\mathbf{x}^T\mathbf{u}_1 = 0$, where $\mathbf{u}_1 = (1, 0, 0)$.
Solution: If $\mathbf{x} = (x_1, x_2, x_3)$ then the constraint $\mathbf{x}^T\mathbf{u}_1 = 0$ means $x_1 = 0$.
For such unit vectors we have $x_2^2 + x_3^2 = 1$ and

$$\begin{aligned} 9x_1^2 + 4x_2^2 + 3x_3^2 &= 4x_2^2 + 3x_3^2 \\ &\leq 4x_2^2 + 4x_3^2 \\ &= 4(x_2^2 + x_3^2) \\ &= 4. \end{aligned}$$

Therefore the constrained maximum does not exceed 4.
This value is attained for $\mathbf{x} = (0, 1, 0)$ which is the eigenvector for the second greatest eigenvalue of the matrix of the quadratic form.