Chapter 8 : Eigenvalues and eigenvectors

1. **Definition.** Let V be a vector space over $\mathbb{F}$ and let $T : V \to V$ be a linear operator. A scalar $\lambda \in \mathbb{F}$ is said to be an **eigenvalue** of T if there is a nonzero vector $v \in V$ such that $T(v) = \lambda v$.
2. We say that v is an **eigenvector** of T with eigenvalue λ.
3. Let A be a $n \times n$ matrix over $\mathbb{F}$. An eigenvalue and eigenvector of A are an eigenvalue and eigenvector of the linear map $T_A : \mathbb{F}^n \to \mathbb{F}^n$. given by $T_A(x) = Ax,\ x \in \mathbb{F}^n$, i.e., $\lambda \in \mathbb{F}$ is an eigenvalue of A if there exists a nonzero vector $x \in \mathbb{F}^n$ with $Ax = \lambda x$.
4. **Example.** Let V be the real vector space of all smooth real valued functions on $\mathbb{R}$. Let $D : V \to V$ be the derivative map. The function $f(x) = e^{\lambda x}$ is an eigenvector with eigenvalue λ since $D(e^{\lambda x}) = \lambda e^{\lambda x}$.
5. **Example.** Let A be a diagonal matrix with scalars $\mu_1, \ldots, \mu_n$ on the diagonal. We write this as $A = \operatorname{diag}(\mu_1, \ldots, \mu_n)$.
6. Then $Ae_i = \mu_i e_i$ and so $\{e_1, \ldots, e_n\}$ are eigenvectors of A with (corresponding) eigenvalues $\mu_1, \ldots, \mu_n$.

Eigenvalues and eigenvectors of linear operators

1. Let $T : V \to V$ be linear and let $\lambda \in \mathbb{F}$. It can be checked that
$$E_\lambda = \{v \in V : T(v) = \lambda v\}$$
is a subspace of V. If $V_\lambda \neq \{0\}$, then λ is an eigenvalue of T.
2. Any nonzero vector in V_λ is an eigenvector with eigenvalue λ.
3. In this case we say that E_λ is the **eigenspace** of the eigenvalue λ.
4. **Theorem.** Let $T : V \to V$ be a linear operator. Let $\lambda_1, \dots, \lambda_n \in \mathbb{F}$ be distinct eigenvalues of T and let $v_1, \dots, v_n$ be corresponding eigenvectors. Then $v_1, v_2, \dots, v_n$ are linearly independent.
5. **Proof.** Use induction on n., The case $n = 1$ is clear.
6. Let $n > 1$. Assume $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$.
7. Apply T to get $a_1 \lambda_1 v_1 + \cdots + a_n \lambda_n v_n = 0. \quad (1)$
8. Multiply $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$ by λ_1 to get
$$a_1 \lambda_1 v_1 + a_2 \lambda_1 v_2 + \cdots + a_n \lambda_1 v_n = 0. \quad (2)$$
9. Hence $a_2(\lambda_2 - \lambda_1) v_2 + \cdots + a_n(\lambda_n - \lambda_1) v_n = 0$. Since $\lambda_1, \lambda_2, \dots, \lambda_n$ are distinct we get by induction, that $a_2 = \cdots = a_n = 0$. And now we get $a_1 = 0$.
10. **Example.** The functions $\{e^{\lambda_1 x}, \dots, e^{\lambda_n x}\}$ where $\lambda_1, \dots, \lambda_n$ are distinct real numbers, are linearly independent as $D(e^{ax}) = ae^{ax}$ for all $a \in \mathbb{R}$.

Diagonalizable matrices and linear operators

1. **Definition.** Let V be a f.d.v.s. over $\mathbb{F}$ and let $T : V \to V$ be a linear operator. We say that T is **diagonalizable** if there exists a basis of V consisting of eigenvectors of T.
2. If $B = (v_1, \dots, v_n)$ is an ordered basis with $T(v_i) = \lambda_i v_i,\ \lambda_i \in \mathbb{F}$ then
$$M_B^B(T) = \operatorname{diag}(\lambda_1, \dots, \lambda_n).$$
3. **Definition.** An $n \times n$ matrix A over $\mathbb{F}$ is said to be diagonalizable if $T_A : \mathbb{F}^n \to \mathbb{F}^n$, given by $T_A(x) = Ax,\ x \in \mathbb{F}^n$, is diagonalizable.
4. **Proposition.** An $n \times n$ matrix A over $\mathbb{F}$ is diagonalizable if and only if $P^{-1}AP$ is a diagonal matrix, for some invertible matrix P over $\mathbb{F}$.
5. In that case, the columns of P are eigenvectors of A and the ith diagonal entry of $P^{-1}AP$ is the eigenvalue associated with the ith column of P.
6. **Proof.** Let A be diagonalizable and let $\{v_1, \dots .v_n\}$ be a basis of $\mathbb{F}^n$ with $T_A(v_i) = Av_i = \lambda_i v_i$.
7. Let $P = [v_1\ v_2 \dots v_n]$ and $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Then
$$AP = A[v_1\ v_2 \dots v_n] = [\lambda_1 v_1\ \lambda_2 v_2 \dots \lambda_n v_n] = PD \implies P^{-1}AP = D.$$

Diagonalizable matrices and linear operators

1. Suppose $P^{-1}AP = D$, where D is diagonal. Then $AP = PD$.
2. Therefore the ith column of P is an eigenvector with eigenvalue λ_i.
3. **Definition.** Let A be a $n \times n$ matrix over $\mathbb{F}$. We define the **characteristic polynomial** P_A of A to be $P_A(t) = \det(tI - A)$.
4. $P_A(t)$ is a monic polynomial of degree n, i.e., the coefficient of t^n is 1.
5. **Proposition.** If $A = PBP^{-1}$ then $P_A(t) = P_B(t)$.
6. **Proof.** We have

$$\begin{aligned} P_A(t) &= \det(tI - PBP^{-1}) = \det(P(tI - B)P^{-1}) \\ &= \det(P)\det(tI - B)\det(P^{-1}) = P_B(t). \end{aligned}$$

7. **Proposition.** (1) Eigenvalues of a square matrix A are the roots of $P_A(t)$ lying in $\mathbb{F}$. (2) For a scalar $\lambda \in \mathbb{F}$, $V_\lambda =$ nullspace of $A - \lambda I$.
8. **Proof.** (1) $\lambda \in \mathbb{F}$ is an eigenvalue of A
   $\iff Av = \lambda v$ for some nonzero $v \iff (A - \lambda I)v = 0$ for some nonzero v.
9. $\iff \mathit{nullity}(A - \lambda I) > 0$
   $\iff \operatorname{rank}(A - \lambda I) < n \iff \det(A - \lambda I) = 0$.
10. (2) $V_\lambda = \{v \mid Av = \lambda v\} = \{v \mid (A - \lambda I)v = 0\} = \mathcal{N}(A - \lambda I)$.

Computation of eigenvalues and eigenspaces

1. **Example.** Let $A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$. To find the eigenvalues of A we solve:

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 1 & -2 \\ 0 & \lambda - 3 \end{bmatrix} = (\lambda - 1)(\lambda - 3) = 0.$$

2. Hence the eigenvalues of A are 1 and 3.
3. Let us calculate the eigenspaces V_1 and V_3. By definition

$$V_1 = \{v \mid (A - I)v = 0\} \text{ and } V_3 = \{v \mid (A - 3I)v = 0\}.$$

4. $A - I = \begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}$, $\begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2y \\ 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ So, $E_1 = L\{(1, 0)\}$.
5. $A - 3I = \begin{bmatrix} 1 - 3 & 2 \\ 0 & 3 - 3 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} -2 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.
6. Then $\begin{bmatrix} -2x + 2y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Hence $x = y$. Thus $E_3 = L(\{(1, 1)\})$.

Eigenvalues and eigenspaces of the rotation matrix

1. **Example.** We use the notation $i = \sqrt{-1}$.
2. Let $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, where $\theta \neq 0, 2\pi$. Now

$$\begin{aligned} P_A(t) &= \det \begin{bmatrix} t-\cos\theta & \sin\theta \\ -\sin\theta & t-\cos\theta \end{bmatrix} \\ &= (t-\cos\theta)^2 + \sin^2\theta \\ &= (t-e^{i\theta})(t-e^{-i\theta}), \end{aligned}$$

3. So, the real matrix A has no eigenvalues and thus no eigenvectors.
4. Note that A represents counter clockwise rotation by θ.
5. But as a complex matrix A has two distinct eigenvalues $e^{i\theta}$ and $e^{-i\theta}$.
6. An eigenvector corresponding to $e^{i\theta}$ is $(1, -i)^t$ and an eigenvector corresponding to $e^{-i\theta}$ is $(-i, 1)^t$.

Computation of powers of a matrix using eigenvalues

1. **Example.** Find A^8, where $A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$. The eigenvalues of A are $2, 1$.
2. The corresponding eigenvectors are $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.
3. Set $P = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$.
4. Then $P^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$ and $A = PDP^{-1}$.
5. We find A^8 using the eigenvalues.

$$\begin{aligned}
A^8 &= (PDP^{-1})^8 = (PDP^{-1}) \cdots (PDP^{-1}) = PD^8P^{-1} \\
&= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2^8 & 0 \\ 0 & 1^8 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 256 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\
&= \begin{bmatrix} 766 & -765 \\ 510 & -509 \end{bmatrix}.
\end{aligned}$$

Algebraic and geometric multiplicity of eigenvalues

1. Let $T : V \to V$ be a linear transformation of a fdvs over $\mathbb{F}$.
2. We define the **characteristic polynomial** $P_T(t)$ of T to be $P_A(t)$, where $A = M_B^B(T)$ wrt an ordered basis of V.
3. as similar matrices have same characteristic polynomials it is immaterial which ordered basis B we take.
4. Let $f(x)$ be a polynomial with coefficients in $\mathbb{F}$.
5. Let $\mu \in \mathbb{F}$ be a root of $f(x)$. Then $(x - \mu)$ divides $f(x)$.
6. The **multiplicity** of the root μ is the largest positive integer k such that $(x - \mu)^k$ divides $f(x)$ but $(x - \mu)^{k+1}$ does not.
7. Let V be a fdvs over $\mathbb{F}$ and let $T : V \to V$ be a linear operator.
8. Let μ be an eigenvalue of T, **geometric multiplicity** of $\mu := \dim V_\mu$.
9. **algebraic multiplicity** of $\mu :=$ multiplicity of μ as a root of $P_T(t)$.

Geometric multiplicity $\leq$ algebraic multiplicity

1. **Theorem.** Let T be a fdvs over $\mathbb{F}$. Then the geometric multiplicity of an eigenvalue $\mu \in \mathbb{F}$ of T is less than or equal to the algebraic multiplicity of μ.
2. **Proof.** Suppose that the algebraic multiplicity of μ is k and the geometric multiplicity of μ is g. Hence V_μ has a basis of g eigenvectors $v_1, v_2, \ldots, v_g$. We can extend this basis of V_μ to an ordered basis of V say $B = (v_1, v_2, \ldots, v_g, \ldots, v_n)$. Now

$$M_B^B(T) = \left[\begin{array}{c|c} \mu I_g & D \\ \hline 0 & C \end{array}\right]$$

3. D is an $g \times (n-g)$ matrix and C is an $(n-g) \times (n-g)$ matrix.
4. From the form of $M_B^B(T)$, $(\lambda - \mu)^g$ divides $\det(A - \lambda I)$. Thus $g \leq k$.

Criterion for diagonalizability

1. **Theorem.** Let $T : V \to V$ be a linear operator, where V is a n-dimensional vector space over $\mathbb{F}$. Then (1) T is diagonalizable $\iff \sum_\lambda \dim V_\lambda = \dim V$. (2) Assume $\mathbb{F} = \mathbb{C}$. Then T is diagonalizable iff the algebraic and geometric multiplicities are equal for each eigenvalue of T.
2. **Proof.** (1) Suppose that T is diagonalizable. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T. Let B_i be a basis of V_{λ_i} for $i = 1, 2, \ldots, k$.
3. Note that $V_\lambda \cap V_\mu = (o)$ for $\lambda \neq \mu$.
4. Therefore $B_1 \cup B_2 \cup \cdots \cup B_k$ is a basis of V having eigenvectors of T.
5. (2) Let $\mathbb{F} = \mathbb{C}$. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of T.
6. By the Fundamental theorem of Algebra, $P_T(t) = \prod_{i=1}^k (t - \lambda_i)^{m_i}$, where m_i is the algebraic multiplicity of λ_i.
7. Since $\sum_i m_i = n$, and m_i=geometric multiplicity of λ_i, the result follows.
8. **Example.** $A = \begin{bmatrix} 3 & 0 & 0 \\ -2 & 4 & 2 \\ -2 & 1 & 5 \end{bmatrix}$, $\det(\lambda I - A) = (\lambda - 3)^2(\lambda - 6)$.
9. Hence eigenvalues of A are 3 and 6. The eigenvalue $\lambda = 3$ has **algebraic multiplicity** 2 and the algebraic multiplcity of 6 is one.

Geometric and algebraic multiplicity of eigenvalues

1. Let us find the eigenspaces V_3 and V_6.

$$\lambda = 3 : A - 3I = \begin{bmatrix} 0 & 0 & 0 \\ -2 & 1 & 2 \\ -2 & 1 & 2 \end{bmatrix}. \quad \text{Hence } \operatorname{rank}(A - 3I) = 1.$$

2. Therefore nullity $(A - 3I) = 2$. By solving the system $(A - 3I)v = 0$, we find that $\mathcal{N}(A - 3I) = V_3 = L(\{(1,0,1)^t, (1,2,0)^t\})$.
3. So the geometric multiplicity of $\lambda = 3$ is 2.
4. $\lambda = 6 : A - 6I = \begin{bmatrix} -3 & 0 & 0 \\ -2 & -2 & 2 \\ -2 & 1 & -1 \end{bmatrix}$. Hence $\operatorname{rank}(A - 6I) = 2$.
5. Therefore $\dim V_6 = 1$. We can show that $\{(0,1,1)^t\}$ is a basis of V_6.
6. Therefore the algebraic and geometric multiplicities of $\lambda = 6$ are one.
7. Let $P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$, then $P^{-1}AP = \operatorname{diag}(3,3,6)$.
8. **Example.** Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Then $\det(\lambda I - A) = (\lambda - 1)^2$.
9. Show that $\dim E_1 = 1$. Hence A is not diagonalizable.

Orthogonally and unitarily diagonalizable matrices

1. Recall that a complex $n \times n$ matrix A is **diagonalizable** if ithere is an invertible matrix $P \in \mathbb{C}^{n \times n}$ so that $P^{-1}AP = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.
2. An $n \times n$ real matrix is called **orthogonal** if the column vectors of A form an orthonormal basis of $\mathbb{R}^n$. Equivalently $A^t A = I$.
3. A complex $n \times n$ matrix is called **unitary** if the column vectors of A form an orthonormal basis of $\mathbb{C}^n$. Equivalently $A^* A = I$.
4. **Definition.** A matrix $A \in \mathbb{C}^{n \times n}$ is called **unitarily diagonalizable** if there is an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors of A.
5. **Definition.** A real $n \times n$ matrix A is called **orthogonally diagonalizable** if there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A.
6. **Spectral Theorem for real matrices.** $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix $\iff$ A is orthogonally diagonalizable.
7. **Theorem.** (a) $A \in \mathbb{R}^{n \times n}$ is orthogonally diagonalizable $\implies$ $A = A^t$.
8. (b) $A \in \mathbb{C}^{n \times n}$ is unitarily diagonalizable $\implies$ $A^* A = AA^*$.
9. **Proof.** (a) Let A be a real $n \times n$ orthogonally diagonalizable matrix.
10. Let $v_1, v_2, \ldots, v_n$ be an orthonormal basis of $\mathbb{R}^n$ with $Av_i = \lambda_i v_i$, $\lambda_i \in \mathbb{R}$ and let $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

Normal and symmetric matrices

1. Let P be the $n \times n$ matrix with i^{th} column v_i. Then $AP = PD$.
2. Since the v_i are orthonormal we have $P^t P = I$. Therefore
$$A = PDP^t \text{ and } A^t = PD^t P^t.$$
3. Since D is a diagonal matrix we have $D = D^t$ and hence $A = A^t$.
4. (b) Let A be a complex $n \times n$ unitarily diagonalizable matrix.
5. Let $v_1, v_2, \ldots, v_n$ be an orthonormal basis of $\mathbb{C}^n$ with $Av_i = \lambda_i v_i$, $\lambda_i \in \mathbb{C}$ and let $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.
6. Let P be the $n \times n$ matrix with i^{th} column v_i. Then $AP = PD$.
7. Since the v_i are orthonormal we have $P^* P = I$. Thus
$$A = PDP^* \text{ and } A^* = PD^* P^*.$$
8. Therefore $AA^* = (PDP^*)(PD^*P^*) = PDD^*P^*$ and $A^*A = PD^*DP^*$.
9. Since D is a diagonal matrix, $D^*D = DD^*$ and therefore $AA^* = A^*A$.
10. A square complex matrix A is called **normal** if $A^*A = AA^*$.
11. A square complex matrix is called Hermitian (resp. skew Hermitioan) if $A^* = A$ $(resp. A^* = -A$. Note that a real symmetric matrix is Hermitian and Hermitian matrices are normal.

Statement of the Spectral Theorems

1. **Spectral Theorem for real symmetric matrices.** Any symmetric real $n \times n$ matrix A is orthogonally diagonalizable. In other words. there is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A.
2. **Spectral Theorem for normal matrices.** Let A be an $n \times n$ complex normal matrix. Then there is an orthonormal basis of $\mathbb{C}^n$ consisting of eigenvectors of A. In other words, A is unitarily diagonalizable.
3. We shall prove the Spectral Theorem for Hermitian matrices first and then deduce the one for normal matrices.
4. **Theorem.** The eigenvalues of a Hermitian matrix are real.
5. Let A be a Hermitian matrix. Then for any $v \in \mathbb{C}^n$
$$(v^*Av)^* = v^*A^*v = v^*Av.$$
6. Therefore v^*Av is a real number. Let λ be an eigenvalue of A with eigenvector v. Then $v^*Av = v^*(\lambda v) = \lambda(v^*v) = \lambda\|v\|^2 \implies \lambda \in \mathbb{R}$.

Self-adjoint operators on inner product spaces

1. Though a proof of the spectral theorem for self-adjoint matrices can be given working only with matrices, a coordinate free approach is more intuitive.
2. We now develop a coordinate free version of the concept of a self-adjoint matrix. The following definition covers both the real and complex cases.
3. **Definition.** Let V be a finite dimensional inner product space over $\mathbb{F}$. A linear operator $T : V \to V$ is said to be **self-adjoint** if

$$\langle x,\ T(y)\rangle = \langle T(x),\ y\rangle,\ \ x, y \in V.$$

4. **Example.** Let $A \in \mathbb{R}^{n\times n}$ be symmetric and $T_A : \mathbb{R}^n \to \mathbb{R}^n$ be the associated linear operator. Then T_A is self-adjoint.
5. **Proof.** Let $x, y \in \mathbb{R}^n$. Then $\langle x,\ Ay\rangle = x^t Ay = x^t A^t y = \langle Ax,\ y\rangle$.
6. **Exercise.** Prove: if A is Hermitian then $T_A : \mathbb{C}^n \to \mathbb{C}^n$ is self-adjoint.

Characterization of self-adjoint operators

1. **Theorem.** Let V be a finite dimensional inner product space over $\mathbb{F}$ and let $T : V \to V$ be a linear operator. Then T is self-adjoint iff $M_B^B(T)$ is self-adjoint for every ordered orthonormal basis B of V.
2. **Proof.** Let $B = (v_1, \ldots, v_n)$ be an ordered orthonormal basis of V
3. Suppose that T is self-adjoint and $A = (a_{ij}) = M_B^B(T)$.. Then $T(v_j) = \sum_{k=1}^n a_{kj} v_k$. So $\langle T(v_j), v_i \rangle = \langle \sum_{k=1}^n a_{kj} v_k, v_i \rangle = a_{ij}$.
4. Therefore $a_{ij} = \langle T(v_j), v_i \rangle = \langle v_j, T(v_i) \rangle = \langle v_j, \sum_{k=1}^n a_{ki} v_k \rangle = \overline{a_{ji}}$.
5. Conversely suppose that $A = (a_{ij}) = M_B^B(T)$ is self-adjoint. Then $\langle T(v_j), v_i \rangle = \langle v_j, T(v_i) \rangle$. Let $x = \sum_{j=1}^n a_j v_j$ and $y = \sum_{i=1}^n b_i v_i$.

$$\begin{aligned} \langle x, T(y) \rangle &= \langle \sum_j a_j v_j, \sum_i b_i T(v_i) \rangle = \sum_{j,i} \overline{a_j} b_i \langle v_j, T(v_i) \rangle, \\ \langle T(x), y \rangle &= \langle \sum_j a_j T(v_j), \sum_i b_i v_i \rangle = \sum_{j,i} \overline{a_j} b_i \langle T(v_j), v_i \rangle. \end{aligned}$$

6. Therefore T is self-adjoint.

Spectral Theorem for self-adjoint operators

1. **(Spectral Theorem for Self-Adjoint Operators)** Let V be a finite dimensional inner product space over $\mathbb{F}$ and let $T : V \to V$ be a self-adjoint linear operator. Then there exists an orthonormal basis of V consisting of eigenvectors of T.
2. **Proof.** By the fundamental theorem of algebra and the fact that Hermitian matrices have only real eigenvalues, there exists $\lambda \in \mathbb{R}$ and a unit vector $v \in V$ with $T(v) = \lambda v$. Put $W = L(\{v\})^{\perp}$.
3. **Claim.** $w \in W$ implies $T(w) \in W$, and $T : W \to W$ is self-adjoint.
4. **Proof.** $\langle T(w), v\rangle = \langle w, T(v)\rangle = \langle w, \lambda v\rangle = \lambda\langle w, v\rangle = 0$, since $w \in W$. Therefore $T(w) \in W$.
5. By induction on dimension, there is an orthonormal basis B of W consisting of eigenvectors of $T : W \to W$. Hence $\{v\} \cup B$ is an orthonormal basis of V.
6. **Spectral Theorem for Hermitian matrices** Let A be an $n \times n$ Hermitian matrix with eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. Set $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$. Then there exists an $n \times n$ unitary matrix U such that $U^* AU = D$.
7. **Spectral Theorem for Real Symmetric matrices** Let A be a $n \times n$ real symmetric matrix with eigenvalues $\lambda_1, \ldots, \lambda_n$. Set $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.
8. Then there exists an $n \times n$ real orthogonal matrix S such that $S^t AS = D$.

Eigenspaces of self-adjoint matrices are mutually $\perp$

1. **Proposition.** Let T be a self-adjoint operator on a finite-dimensional inner product space V. Let u, v be eigenvectors of T with distinct eigenvalues λ and μ respectively. Then $u \perp v$.
2. **Proof.** As T is self-adjoint, $\lambda, \mu \in \mathbb{R}$. Therefore,

$$\begin{aligned}(\lambda - \mu)\langle u,\, v\rangle &= \langle \lambda u,\, v\rangle - \langle u,\, \mu v\rangle \\ &= \langle Tu,\, v\rangle - \langle u,\, Tv\rangle \\ &= \langle u,\, Tv\rangle - \langle u,\, Tv\rangle = 0.\end{aligned}$$

3. Since $\lambda \neq \mu$, u and v are mutually perpendicular.
4. **Theorem.** Let T be a self-adjoint linear operator on a finite dim inner product space V. Let $\lambda_1, \ldots, \lambda_r$ be the distinct eigenvalues of T. Then
$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r} \text{ and } \dim V = \textstyle\sum_{i=1}^r \dim E_{\lambda_i}.$$
5. **Proof.** As T is self-adjoint, V has an orthonormal basis of eigenvectors of T.
6. Let $B_i = \{v_{i1}, \ldots, v_{in_r}\}$ be an orthonormal basis for the eigenspace V_{λ_i}.
7. Thus $V = V_{\lambda_1} + \cdots + V_{\lambda_r}$. Let $v_1 + v_2 + \cdots + v_r = 0$ where $v_i \in V_{\lambda_i}$.
8. Since eigenvectors corresponding to distinct eigenvalues are linearly independent, each $v_i = 0$. Therefore
$$V = V_{\lambda_1} \oplus V_{\lambda_2} \oplus \cdots \oplus V_{\lambda_r} \text{ and } \dim V = \textstyle\sum_{i=1}^r \dim E_{\lambda_i}.$$

Diagonalization of a real symmetric matrix

1. **Example.** Consider the real symmetric matrix $A = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 1 & 2 \\ -2 & 2 & 1 \end{bmatrix}$.
2. Solve $\det(\lambda I - A) = 0$. Check that the eigenvalues of A are $3, 3, -3$.
3. The eigenvectors for $\lambda = 3$ are in the null space of $\mathcal{N}(A - 3I)$.
4. They are are the nonzero solutions of
$$\begin{bmatrix} -2 & 2 & -2 \\ 2 & -2 & 2 \\ -2 & 2 & -2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$
5. Hence we obtain the single equation $x - y + z = 0$.
$$E_3 = \{(y - z, y, z) \mid y, z \in \mathbb{R}\} = L(\{u_1 = (0, 1, 1)^t,\ u_2 = (-1, 1, 2)^t\}).$$
6. Apply Gram-Schmidt process to get an orthonormal basis of E_3:
$$v_1 = \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^t \quad \text{and} \quad v_2 = \left(-\sqrt{\frac{2}{3}}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)^t.$$
7. Check that $\{v_3 = (1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})^t\}$ is an orthonormal basis of V_{-3}.
8. Set $S = [v_1, v_2, v_3]$ and $D = \operatorname{diag}(3, 3, -3)$. Then $S^t A S = D$.

Applications of spectral theorem to geometry

1. **Definition.** Let V be a real inner product space with orthonormal basis $\{e_1, e_2, \ldots, e_n\}$. Let $A = (a_{ij})$ be an $n \times n$ real matrix.
2. The **quadratic form** associated with A is $Q : V \to \mathbb{R}$ defined by :
$$Q_A(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \quad x_1, x_2, \ldots, x_n \in \mathbb{R}$$
3. If $A = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ then $Q_A(x) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$ is called a **diagonal form**.
4. **Proposition.** Let $X = (x_1, x_2, \ldots, x_n)^t$. Then $Q_A(x) = X^t A X$.
5. **Example.** (1) $A = \begin{bmatrix} 1 & 1 \\ 3 & 5 \end{bmatrix}, \ X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then
$$X^t A X = [x_1, x_2] A [x_1, x_2]^t = x_1^2 + 4x_1x_2 + 5x_2^2.$$
6. **Example.** (2) Let $B = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \ X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then
$$X^t B X = [x_1, x_2] \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} = x_1^2 + 4x_1x_2 + 5x_2^2.$$
7. A and B give rise to same $Q(x)$ and $B = \frac{1}{2}(A + A^t)$ is a *symmetric matrix*.

Examples of quadratic forms

1. **Proposition.** For any $n \times n$ matrix A and $X = (x_1, x_2, \ldots, x_n)^t$,
$$X^tAX = X^tBX \quad \text{where } B = \frac{1}{2}(A + A^t).$$
2. So every quadratic form is associated with a symmetric matrix.
3. **Proof.** X^tAX is a 1×1 matrix. Hence $(X^tAX)^t = X^tA^tX = X^tAX$.
4. Therefore
$$X^tAX = \frac{1}{2}X^tAX + \frac{1}{2}X^tA^tX = X^t\frac{1}{2}(A + A^t)X = X^tBX.$$
5. **Theorem.** Let X^tAX be the quadratic form associated to a real symmetric matrix A. Let U be an orthogonal matrix so that $U^tAU = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Then $X^tAX = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$.
6. **Proof.** Since $X = UY$, $X^tAX = (UY)^tA(UY) = Y^t(U^tAU)Y$.
7. Since $U^tAU = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, we get
$$X^tAX = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2.$$

Diagonalization quadratic forms

1. **Example.** Let us determine the orthogonal matrix U which reduces the quadratic form $Q(x) = 2x_1^2 + 4x_1x_2 + 5x_2^2$ to a diagonal form.
2. We write $Q(x) = [x_1, x_2]\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = X^t AX$.
3. The eigenvalues of $A = \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix}$ are $\lambda_1 = 1$ and $\lambda_2 = 6$.
4. An orthonormal set of eigenvectors for λ_1 and λ_2 is
$$u_1 = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ -1 \end{bmatrix} \quad \text{and} \quad u_2 = \frac{1}{\sqrt{5}}\begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$
5. Hence $U = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$. Check that $U^t AU = \mathrm{diag}(1, 6)$.
6. Now use $X = UY$ to get the diagonal form $Y^t \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix} Y = y_1^2 + 6y_2^2$.

Identification of conic sections

1. A conic section is the locus in the Cartesian plane $\mathbb{R}^2$ of an equation
$$ax^2 + bxy + cy^2 + dx + ey + f = 0.$$
2. It can be proved that this equation represents one of the following:
3. (i) the empty set (ii) single point (iii) one or two straight lines
4. (iv) ellipse (v) hyperbola (vi) parabola.
5. We consider the second degree part $Q(x, y) = ax^2 + bxy + cy^2$
6. This is a quadratic form. This determines the type of the conic.
7. We can write the matrix form after setting $x = x_1, y = x_2$:
$$[x_1, x_2] \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [d, e] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + f \quad = \quad 0$$
8. Write $A = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}$. Let $U = [u_1, u_2]$ be an orthogonal matrix where u_1 and u_2 are eigenvectors of A with eigenvalues λ_1 and λ_2.
9. Apply the change of variables $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = U \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ to diagonalize the quadratic form $Q(x_1, x_2)$ to the diagonal form $\lambda_1 y_1^2 + \lambda_2 y_2^2$.

Identification of conic sections

1. The orthonormal basis $\{u_1, u_2\}$ determines new coordinate axes.
2. The locus of the equation $X^t AX + BX + f = 0$
3. where $B = [d, e]$ is same as the locus of the equation

$$\begin{aligned} 0 &= Y^t \operatorname{diag}(\lambda_1, \lambda_2) Y + (BU)Y + f \\ &= Y^t \operatorname{diag}(\lambda_1, \lambda_2) Y + (BU)Y + f \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + [d, e][u_1, u_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + f. \end{aligned}$$

4. **Example of an ellipse.** We shall identify the conic section represented by

$$2x_1^2 + 4x_1x_2 + 5x_2^2 + 4x_1 + 13x_2 - 1/4 = 0.$$

5. We have earlier diagonalized the quadratic form $2x_1^2 + 4x_1x_2 + 5x_2^2$.
6. The associated symmetric matrix, the eigenvectors and eigenvalues are displayed in the equation of diagonalization :

$$U^t AU = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix}.$$

An Ellipse

1. Set $t = 1/\sqrt{5}$. Then the new coordinates are defined by
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2t & t \\ -t & 2t \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix},$$
2. This means $x_1 = t(2y_1 + y_2)$ and $x_2 = t(-y_1 + 2y_2)$.
3. Substitute these into the original equation to get
$$y_1^2 + 6y_2^2 - \sqrt{5}y_1 + 6\sqrt{5}y_2 - \frac{1}{4} = 0.$$
4. Complete the square to write this as
$$(y_1 - \frac{1}{2}\sqrt{5})^2 + 6(y_2 + \frac{1}{2}\sqrt{5})^2 = 9.$$
5. This represents an ellipse with center $(\frac{1}{2}\sqrt{5}, -\frac{1}{2}\sqrt{5})$ in the y_1y_2-plane.
6. The y_1 and y_2 axes are determined by the eigenvectors u_1 and u_2.
7. **Example.** Let us identify the locus of the equation
$$2x_1^2 - 4x_1x_2 - x_2^2 - 4x_1 + 10x_2 - 13 = 0.$$
8. We write the equation in matrix form as
$$[x_1, x_2] \begin{bmatrix} 2 & -2 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [-4, 10] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - 13 = 0.$$

A hyperbola

1. Let $t = 1/\sqrt{5}$. The eigenvalues of A are $\lambda_1 = 3, \lambda_2 = -2$.
2. An orthonormal set of eigenvectors is $\{u_1 = t(2, -1)^t, u_2 = t(1, 2)^t\}$.
3. Now write $U = t \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$ and $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = U \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$.
4. The transformed equation becomes

$$3y_1^2 - 2y_2^2 - 4t(2y_1 + y_2) + 10t(-y_1 + 2y_2) - 13 = 0$$

$$\implies 3y_1^2 - 2y_2^2 - 18ty_1 + 16ty_2 - 13 = 0.$$

5. Complete the square to get $3(y_1 - 3t)^2 - 2(y_2 - 4t)^2 = 12$. Therefore

$$\frac{(y_1 - 3t)^2}{4} - \frac{(y_2 - 4t)^2}{6} = 1.$$

6. This represents a hyperbola with center $(3t, 4t)$ in the y_1y_2-plane.
7. The vectors u_1 and u_2 are the directions of positive y_1 and y_2 axes.

A parabola

1. **Example.** Consider $9x_1^2 + 24x_1x_2 + 16x_2^2 - 20x_1 + 15x_2 = 0$.
2. The symmetric matrix for the quadratic part is $A = \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$.
3. The eigenvalues are $\lambda_1 = 25, \lambda_2 = 0$.
4. Put $a = 1/5$. An orthonormal set of eigenvectors is $\{u_1 = a(3,4)^t, u_2 = a(-4,3)^t\}$
5. An orthogonal diagonalizing matrix is $U = a\begin{bmatrix} 3 & -4 \\ 4 & 3 \end{bmatrix}$.
6. The equations of change of coordinates are

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = U \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \implies x_1 = a(3y_1 - 4y_2),\ x_2 = a(4y_1 + 3y_2).$$

7. The equation in y_1y_2-plane is $y_1^2 + y_2 = 0$.
8. This is an equation of parabola with its vertex at the origin.