$\S$6.1 Eigenvalues and Eigenvectors

We are interested in solving equations of the form

$$A\mathbf{x} = \lambda\mathbf{x}$$

for a square matrix A, a vector $\mathbf{x}$, and a scalar λ.
In other words, we're looking at the seemingly rare situation where multiplying the vector by the matrix gives exactly a multiple of the vector.
It turns out that in real life this is actually not that rare: it occurs in...

- steady-state vector problems
- vibrations in structures
- boundary value problems
- energy states of atoms...the list goes on and on.

To study this, first recall the definition of the Complex Numbers, denoted $\mathbb{C}$...
$\mathbb{C} = \{a + bi | a, b \in \mathbb{R}\}$, where $i = \sqrt{-1}$.

We need our scalar λ to be in $\mathbb{C}$, not just in $\mathbb{R}$.

Let $A \in \mathbb{C}^{n \times n}$. The vector $\mathbf{x} \in \mathbb{C}^n$ is called an **eigenvector** (or **characteristic vector**) if $\mathbf{x} \neq 0$ and there exists a scalar $\lambda \in \mathbb{C}$ such that $A\mathbf{x} = \lambda\mathbf{x}$. Then λ is called an **eigenvalue** (or **characteristic value**) of A.

We say $\mathbf{x}$ is the eigenvector belonging to the eigenvalue λ.

Example...

Equivalent statements about eigenvalues

- λ is an eigenvalue of A.
- $(A - \lambda I)\mathbf{x} = \nvdash$ has a nontrivial solution.
- The **eigenspace** $N(A - \lambda I) \neq \{\nvdash\}$.
- $A - \lambda I$ is singular.
- $\det\,(\mathrm{A} - \lambda\mathrm{I}) = 0$.

Now let's consider λ not as a specific value, but as a variable. Since A is an $n \times n$ matrix, we can expand $\det\,(\mathrm{A} - \lambda\mathrm{I})$ into an nth degree polynomial over λ:

$$p(\lambda) = \det\,(\mathrm{A} - \lambda\mathrm{I})$$

which is called the *characteristic polynomial*.

The equation

$$\det\,(\mathrm{A} - \lambda\mathrm{I}) = 0$$

is called the *characteristic equation* for the matrix A.

Why we need our scalars to be in $\mathbb{C}$

The fundamental theorem of algebra states that in $\mathbb{C}$, $p(\lambda)$ has exactly n roots, which are our eigenvalues

$$\lambda_1, \lambda_2, \ldots, \lambda_n.$$

This is because $p(\lambda)$ can be factored completely into

$$p(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda).$$

The λ_i's are specific values that solve the equation, while λ is just the dummy variable for this polynomial (usually we use x, but λ is standard for this study). Some of our λ_i's might have the same value (called multiplicities), but that's OK.

If $A \in \mathbb{R}^{n \times n}$, then the characteristic polynomial of A will have coefficients in $\mathbb{R}$, and so all its complex roots (which are all the complex eigenvalues) will occur in pairs λ, $\overline{\lambda}$. $\overline{\lambda}$ is called the **conjugate** of λ:

$$\lambda = a + bi \leftrightarrow \overline{\lambda} = a - bi.$$

The eigenvectors will also occur in conjugate pairs.

Products and Sums of Eigenvalues

When we factor $p(\lambda) = \det\,(\mathrm{A} - \lambda \mathrm{I})$, we get

$$p(\lambda) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n).$$

Regardless of whether n is even or odd, this is equal to

$$p(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda).$$

This is a polynomial in λ. If we plug in 0 for λ, we get

$$p(0) = \lambda_1 \lambda_2 \dots \lambda_n = \det\,(\mathrm{A}).$$

Recall that the sum of the diagonal elements of A is called the **trace** of A, denoted $\mathrm{tr}(A)$. (In other words, $\mathrm{tr}(A) = a_{11} + a_{22} + \ldots + a_{nn}$.)

So by the above factorization of the characteristic polynomial, we now also know that the trace is the sum of the eigenvalues:

$$\mathrm{tr}(A) = \lambda_1 + \lambda_2 + \ldots + \lambda_n.$$

Example...

Recall from Chapter 4 the definition of **Similar matrices**:
$A, B \in \mathbb{R}^{n\times n}$ are called **similar** if $B = QAQ^{-1}$ for some invertible matrix $Q \in \mathbb{R}^{n\times n}$.

Theorem: Similar matrices have the same char. poly.

Proof: $B = QAQ^{-1} \Rightarrow B - \lambda I = Q(A - \lambda I)Q^{-1}$.

Thus $\det\ (\mathrm{B} - \lambda \mathrm{I}) = \det\ \mathrm{Q} \det\ (\mathrm{A} - \lambda \mathrm{I}) \det\ \mathrm{Q}^{-1}$

$= \det\ (\mathrm{A} - \lambda \mathrm{I})$.

Since A and B have the same characteristic polynomial, they must also have the same eigenvalues.

§6.2 Systems of linear (ordinary) differential equations (SOLODE)

$$\begin{array}{ccccccc}
y_1' & = & a_{11}y_1 & + & a_{12}y_2 & + \ldots + & a_{1n}y_n \\
y_2' & = & a_{21}y_1 & + & a_{22}y_2 & + \ldots + & a_{2n}y_n \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots\vdots & \\
y_n' & = & a_{n1}y_1 & + & a_{n2}y_2 & + \ldots + & a_{1n}y_n
\end{array}$$

In matrix terms we write:

$$\mathbb{Y}' = A\mathbb{Y}$$

where $\mathbb{Y} = \mathbb{Y}(t) = (y_1(t), y_2(t), ..., y_n(t))^{\mathrm{T}}$ and $A \in \mathbb{C}^{n\times n}$ a constant matrix.

If $\mathbf{x}(\neq 0)$ is an eigenvector of A corresponding to the eigenvalue λ, then

$$\mathbb{Y}(t) = e^{\lambda t}\mathbf{x}$$

is a nontrivial solution of the given SOLODE.

Initial Value Problem

Let $A \in \mathbb{C}^{n \times n}$ and assume that

$$\det\,(\mathrm{A} - \lambda \mathrm{I}) = (\lambda_1 - \lambda)^{\mathrm{m}_1}(\lambda_2 - \lambda)^{\mathrm{m}_2}\ldots(\lambda_\mathrm{k} - \lambda)^{\mathrm{m}_\mathrm{k}},$$

where $\lambda_i \neq \lambda_j$ for $i \neq j$

and $1 \leq m_i$ (the multiplicity of λ_i).

Assume that $\dim \mathrm{N}(A - \lambda_i I) = m_i$ and

$\mathrm{N}(A - \lambda_i) = \mathrm{span}(\mathbf{x}_{i1}, \ldots, \mathbf{x}_{im_i})$ for $i = 1, \ldots, k$.

Then the general solution of SOLODE is:

$$\mathbb{Y}(t) = \sum_{i=1,j=1}^{k,m_i} c_{ij} e^{\lambda_i (t - t_0)} \mathbf{x}_{ij}.$$

$\mathbb{Y}(t)$ is determined by the initial condition $\mathbb{Y}(t_0) = \mathbf{c}$.

If $\mathbb{Y}_1$ and $\mathbb{Y}_2$ are both solutions to $\mathbb{Y}' = A\mathbb{Y}$, then $\alpha\mathbb{Y}_1 + \beta\mathbb{Y}_2$ is also a solution.

Note: a **vector function** is a function that has vectors as its domain and range.

Complex eigenvalues of real matrices

Recall Euler's formula for e^z where $z = a + ib,\ a, b \in \mathbb{R}$:

$e^z = e^{a+ib} = e^a e^{ib} = e^a(\cos b + i \sin b)$.
$\Rightarrow e^{zt} = e^{at}(\cos(bt) + i \sin(bt))$.

Claim: Let $A \in \mathbb{R}^{n\times n}$ and assume $\lambda := \alpha + i\beta,\ \alpha, \beta \in \mathbb{R}$ is non-real eigenvalue ($\beta \neq 0$). Then the corresponding eigenvector $\mathbf{x} = \mathbf{u} + i\mathbf{v},\ \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ ($A\mathbf{u} = \lambda\mathbf{u}$) is non-real ($\mathbf{v} \neq 0$). Furthermore $\overline{\lambda} = \alpha - i\beta \neq \lambda$ is another eigenvalue of A with the corresponding eigenvector $\overline{\mathbf{x}} = \mathbf{u} - i\mathbf{v}$. So $e^{\lambda t}\mathbf{x}$ and $e^{\overline{\lambda}t}\overline{\mathbf{x}}$ are solutions to the first-order system $\mathbb{Y}' = A\mathbb{Y}$.

Now let $\mathbb{Y}_1 = Re(e^{\lambda t}\mathbf{x})$ and $\mathbb{Y}_2 = Im(e^{\overline{\lambda}t}\overline{\mathbf{x}})$. So $\mathbb{Y}_1$ and $\mathbb{Y}_2$ are vector functions with real-valued solutions to $\mathbb{Y}' = A\mathbb{Y}$. Using Euler's formula, we can calculate

$Y_1 = e^{at}[(\cos bt)Re\mathbf{x} - (\sin bt)Im\mathbf{x}]$

$Y_2 = e^{at}[(\cos bt)Im\mathbf{x} + (\sin bt)Re\mathbf{x}]$.

Second Order Linear Differential Systems

$\mathbb{Y}'' = A_1\mathbb{Y} + A_2\mathbb{Y}'$,

$A_1, A_2 \in \mathbb{C}^{n\times n}, \ \mathbb{Y} = (y_1, \ldots, y_n)^{\mathrm{T}}$.

Let $\mathbf{z} = (y_1, \ldots, y_n, y_1', \ldots, y_n')^{\mathrm{T}}$. Then

$$\mathbf{z}' = A\mathbf{z}, \text{ where } A = \begin{pmatrix} 0_n & I_n \\ A_1 & A_2 \end{pmatrix} \in \mathbb{C}^{2n\times 2n}.$$

Here 0_n is $n \times n$ zero matrix and I_n is $n \times n$ identity matrix.

The initial conditions are
$\mathbb{Y}(t_0) = \mathbf{a} \in \mathbb{C}^n, \ \mathbb{Y}'(t_0) = \mathbf{b} \in \mathbb{C}^n$ which are equivalent to the initial conditions $\mathbf{z}(t_0) = \mathbf{c} \in \mathbb{C}^{2n}$.

The solution of the second order differential system with n unknown functions can be solved by converting this system to the first order system with $2n$ unknown functions.

Example...

§6.3 Diagonalization

Theorem. Let $A \in \mathbb{C}^{n\times n}$ and assume that
$\det\,(\mathrm{A}-\lambda\mathrm{I}) = (\lambda_1-\lambda)^{\mathrm{m}_1}(\lambda_2-\lambda)^{\mathrm{m}_2}\ldots(\lambda_\mathrm{k}-\lambda)^{\mathrm{m}_\mathrm{k}}$,
where $\lambda_i \neq \lambda_j$ for $i \neq j$ and $1 \leq m_i$ (the multiplicity of λ_i). Assume that $\dim \mathrm{N}(A-\lambda_i I) = m_i$ and
$\mathrm{N}(A-\lambda_i I) = \mathrm{span}(\mathbf{x}_{i1}, \ldots, \mathbf{x}_{im_i})$ for $i = 1, \ldots, k$.
Form the matrix whose columns are the vectors which span the null space
$X = (\mathbf{x}_{11} \ldots \mathbf{x}_{1m_1}\ \mathbf{x}_{21} \ldots \mathbf{x}_{2m_2} \ldots \mathbf{x}_{km_k}) \in \mathbb{C}^{n\times n}$
and the diagonal matrix whose entries are the eigenvalues of A: $\quad D = \mathrm{diag}(\lambda_1\ \ \ldots \lambda_k)$, where the diagonal entry λ_i repeats m_i times for $i = 1, ..., k$.

Then X is an invertible matrix and $A = XDX^{-1}$, i.e. A is similar to D.

Definition. $A \in \mathbb{C}^{n\times n}$ is called **diagonalizable** if A is similar to a diagonal matrix $D \in \mathbb{C}^{n\times n}$. (The diagonal entries of D are the eigenvalues of A counted with multiplicities.)

Lemma. Let $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_p$ be p eigenvectors of A corresponding to p distinct eigenvalues. Then $\mathbf{y}_1, ..., \mathbf{y}_p$ are linearly independent.

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

Observations

- If A is diagonalizable, then the column vectors of the diagonalizing matrix X are eigenvectors of A, and the diagonal elements of D are the corresponding eigenvalues of A.

- The diagonalizing matrix X is not unique. Given a diagonalizing matrix X, if we reorder the columns of X or multiply the columns by nonzero scalars, we will get a new diagonalizing matrix.

- If A is $n \times n$ and A has n distinct eigenvalues, then A is diagonalizable. If the eigenvalues are not distinct, then A may or may not be diagonalizable. (This depends on whether or not A has n linearly independent eigenvectors.)

- It's easy to compute powers of diagonalizable matrices:

$A = XDX^{-1}$

$\Rightarrow A^2 = (XDX^{-1})(XDX^{-1}) = XD^2X^{-1}$

$\Rightarrow A^m = XD^mX^{-1}$, where $D^m = \mathrm{diag}(\lambda_1^m \ldots \lambda_n^m)$

Example...

Defective matrices

$A \in \mathbb{C}^{n\times n}$ is called **defective** if it is not diagonalizable.

Thus a matrix A is defective if it does not have n linearly independent eigenvectors.

Equivalently, this means there exists an eigenvalue λ_i of A of multiplicity $m_i > 1$ such that $\dim \mathrm{N}(A - \lambda_i I) < m_i$.

In other words, A is defective if one of its eigenvalues has multiplicity that exceeds the dimension of the corresponding eigenspace.

Example: $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

$\det\,(\mathrm{A} - \lambda\mathrm{I}) = (\lambda - 0)^2 = \lambda^2$.

Thus $\lambda = 0$ is an eigenvalue with multiplicity 2. But $\dim \mathrm{N}(A - \lambda I) = \dim \mathrm{N}(A) = 1$. Thus A is defective. Thus A cannot be diagonalized.

A Taste of Markov Chains

In many applied problems, we come across an **iteration process**: given an initial vector $\mathbf{x}_0$, this vector is repeatedly multiplied by the matrix A. So we define

$$\mathbf{x}_1 = A\mathbf{x}_0, \mathbf{x}_2 = A\mathbf{x}_1, \mathbf{x}_3 = A\mathbf{x}_2, \ldots$$

This sequence of vectors is called a Markov Chain.
For any m, we have $\mathbf{x}_m = A\mathbf{x}_{m-1}$.
In other words, $\mathbf{x}_m = A^m\mathbf{x}_0$.
Question: ***Under what conditions will*** $\mathbf{x}_m$ ***converge to a steady-state vector*** $\mathbf{x}$***?***

Definition. A **stochastic process** means a process that involves some randomness. In other words, the outcome at any stage depends on chance, or probability.

A **Markov process** is a stochastic process with the following additional properties:

1. The set of possible outcomes or states is finite.
2. The probability of the next outcome depends only on the previous outcome.
3. The probabilities are constant over time.

$A \in \mathbb{R}^{n\times n}$ is called column stochastic (respectively, row stochastic) if all entries of A are nonnegative and the sum of each column (row) is 1.

That is, $A^{\mathrm{T}}\mathbf{e} = \mathbf{e}$, ($A\mathbf{e} = \mathbf{e}$), where $\mathbf{e} = (1, 1, \ldots, 1)^{\mathrm{T}}$.

Under mild assumptions, (namely, all entries of A are positive), we have that $\lim_{m\to\infty} A^m \mathbf{x}_0 = \mathbf{x}$. If A is column stochastic and $\mathbf{e}^{\mathrm{T}}\mathbf{x}_0 = 1$ then the limit vector is a unique probability eigenvector of A: $A\mathbf{x} = \mathbf{x}$, where $\mathbf{x} = (x_1, \ldots, x_n)^{\mathrm{T}}$,

$$0 < x_1, \ldots, x_n, \quad x_1 + x_2 + \ldots x_n = 1.$$

<u>Answer</u>:

If A is diagonalizable, then $\mathbf{x}_m$ converges to $\mathbf{x}$ for all $\mathbf{x}_0$

$$\Leftrightarrow$$

The first eigenvalue $\lambda_1 = 1$ and for each eigenvalue of A, either $|\lambda| < 1$ or $\lambda = 1$.

Exponential of a Matrix

For a scalar a, we can express e^a in terms of a power series: $e^a = 1 + a + \frac{1}{2!}a^2 + \frac{1}{3!}a^3 + \ldots$ In many applications, it is useful to define the **matrix exponential**:

For $A \in \mathbb{C}^{n\times n}$, let $e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \ldots$

For a diagonal matrix $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$, then $e^D = \mathrm{diag}(e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_n})$.

If A is diagonalizable, i.e. $A = XDX^{-1}$, then $e^A = Xe^DX^{-1}$. We can apply the matrix exponential to the initial value problem $\mathbb{Y}' = A\mathbb{Y}, \quad \mathbb{Y}(0) = \mathbb{Y}_0$.

Example

$$A = \begin{pmatrix} 3 & 4 \\ 3 & 2 \end{pmatrix}, \mathbb{Y}_0 = \begin{pmatrix} 6 \\ 1 \end{pmatrix}.$$

We did this example in the last section: the eigenvalues of A are $\lambda_1 = 6$ and $\lambda_2 = -1$ with corresponding eigenvectors $\mathbf{x}_1 = (4,3)^T$ and $\mathbf{x}_2 = (1,-1)^T$. So we diagonalize A:

$$A = XDX^{-1} = \begin{pmatrix} 4 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 6 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{7} & \frac{1}{7} \\ \frac{3}{7} & -\frac{4}{7} \end{pmatrix}$$

Now the solution is given by

$$\mathbb{Y} = e^{tA}\mathbb{Y}_0$$

$$= Xe^{tD}X^{-1}\mathbb{Y}_0$$

$$= \begin{pmatrix} 4 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} 6 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{7} & \frac{1}{7} \\ \frac{3}{7} & -\frac{4}{7} \end{pmatrix} \begin{pmatrix} 6 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 4e^{6t} + 2e^{-t} \\ 3e^{6t} - 2e^{-t} \end{pmatrix}.$$

Note that this is an alternative method to the same solution as Example 1 of §6.2.

§6.4 Hermitian Matrices 12/4/06

Let $\alpha \in \mathbb{C}$. Then $\alpha = a + bi$, where $a, b \in \mathbb{R}$. The **length** of α is given by: $|\alpha| = \sqrt{\overline{\alpha}\alpha} = \sqrt{a^2 + b^2}$.

Let $\mathbb{C}^n$ denote the vector space of all n-tuples of complex numbers. Then for a vector $\mathbf{z} = (z_1, \ldots, z_n)^T$ in $\mathbb{C}^n$, the length of $\mathbf{z}$ is

$$||\mathbf{z}|| = \sqrt{(|z_1|^2 + \cdots + |z_n|^2)}$$
$$= \sqrt{(\overline{z_1}z_1 + \cdots + \overline{z_n}z_n)} = \sqrt{(\overline{\mathbf{z}}^T\mathbf{z})}$$

Notation: Conjugate Transpose. Write $\mathbf{z}^H$ for $\overline{\mathbf{z}}^T$, the transpose of $\overline{\mathbf{z}}$. Then $||\mathbf{z}|| = \sqrt{(\mathbf{z}^H\mathbf{z})}$.

Definition

Let $\mathbf{V}$ be a vector space over $\mathbb{C}$. For $\mathbf{z}, \mathbf{w} \in V$, the **inner product** $\langle \mathbf{z}, \mathbf{w} \rangle$ is a complex number satisfying the following conditions:

- $\langle \mathbf{z}, \mathbf{z} \rangle \geq 0$ with equality if and only if $\mathbf{z} = 0$. (Positive)
- $\langle \mathbf{z}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{z} \rangle}$ for all $\mathbf{z}, \mathbf{w} \in V$. (***Conjugate*** Symmetric)
- $\langle \alpha\mathbf{z} + \mathbf{w}, \mathbf{u} \rangle = \alpha\langle \mathbf{z}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$ for any scalar $\alpha \in \mathbb{C}$. (Linear)

Just as we defined $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y}$ as a the standard inner product for the vector space $\mathbb{R}^n$, we can define an inner product on $\mathbb{C}^n$ by $\langle \mathbf{z}, \mathbf{w} \rangle = \mathbf{w}^H \mathbf{z}$. This satisfies the inner product axioms.

Example: Find $\mathbf{w}^H \mathbf{z}$, $\mathbf{z}^H \mathbf{z}$, and $\mathbf{w}^H \mathbf{w}$ for

$$\mathbf{z} = \begin{pmatrix} 5+i \\ 1-3i \end{pmatrix}, \ \mathbf{w} = \begin{pmatrix} 2+i \\ -2+3i \end{pmatrix}.$$

Complex Matrix Rules

Let $A, B, M \in \mathbb{C}^{m \times n}$, $C \in \mathbb{C}^{n \times r}$.

- Definition: if $M = A + iB$, then $\overline{M} = A - iB$.
- $(A^H)^H = A$.
- $(\alpha A + \beta B)^H = \overline{\alpha} A^H + \overline{\beta} B^H$.
- $(AC)^H = C^H A^H$.

Definition

A matrix M is **Hermitian** if $M = M^H$.

Example: $M = \begin{pmatrix} 3 & 2-i \\ 2+i & 4 \end{pmatrix}$. Check...

Note: If $M \in \mathbb{R}^{m\times n}$, then $\overline{M} = M$, and so in this case $M^H = M^T$. So if M is both real and also *symmetric* ($M^T = M$), then M is always Hermitian. We can think of Hermitian matrices as the complex analog of real symmetric matrices, and they have lots of convenient properties...

Theorem

The eigenvalues of a Hermitian matrix are all real. Furthermore, eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof: Later if we have time. (See p. 347)

Definition

An $n \times n$ matrix U is **unitary** if its column vectors form an orthonormal set in $\mathbb{C}^n$.

Lemma: U is unitary $\iff U^H U = I$.

Proof

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ be the column vectors of U. Since U is unitary, then $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{ij}$. But remember, this inner product is defined in $\mathbb{C}^n$, so $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \mathbf{u}_i^H \mathbf{u}_j$.
Thus $U^H U = I$. (Note: this is equivalent to $U^{-1} = U^H$.)

Corollary If the eigenvalues of a Hermitian matrix A are distinct, then there exists a unitary matrix U that diagonalizes A.

Example...

The above corollary is actually true even if the eigenvalues of A are not distinct. This is called the **Spectral Theorem**. We will not prove this in full detail here, but it can be proved by factoring A into UTU^H, where U is unitary and T is an upper triangular matrix. The factorization $A = UTU^H$ is called the *Schur decomposition* of A. It comes from the following theorem.

Schur's Theorem

For any $n \times n$ matrix A, there exists a unitary matrix U such that the matrix $T = U^H AU$ is upper triangular.

In the case that A is Hermitian, the matrix T will actually be **diagonal**, which is proved by the Spectral Theorem.

Spectral Theorem

If A is Hermitian, then there exists a unitary matrix U that diagonalizes A.

In the case that A is real and symmetric, then its eigenvalues and eigenvectors must be real. Thus the diagonalizing matrix U must be orthogonal.

Corollary

If A is a real symmetric matrix, then there is an orthogonal matrix U that diagonalizes A, that is, $U^T AU = D$, where D is diagonal.

Example...

There are non-Hermitian matrices that have complete orthonormal sets of eigenvectors. For example...

Definition. A Matrix A is **skew Hermitian** if $A^H = -A$.

Skew Hermitian matrices have complete orthonormal sets of eigenvectors. But other sets of matrices do also.

Question: What is the most general way to describe matrices with complete orthonormal sets of eigenvectors?

Definition. A Matrix A is **normal** if $AA^H = A^H A$.

Theorem

A matrix A is normal if and only if A has a complete orthonormal set of eigenvectors.

Brief glimpse at some stuff we didn't have time to cover

We know by now that matrices are crucial to the study of linear equations. But they also play an important role in quadratic equations.

Definition

A **quadratic equation** in two variables x and y is an equation of the form

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

which we can rewrite as

$$\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} +$$

$$\begin{pmatrix} d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + f = 0.$$

Let $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $\mathbf{x}^T A\mathbf{x} = ax^2 + 2bxy + cy^2$.

The term $\mathbf{x}^T A\mathbf{x}$ is called the **quadratic form** associated with the original quadratic equation.

We can use the quadratic form to solve for example:

- Conic sections
- Max/Min optimizations
- Stationary points of curves and surfaces.

Definition

A matrix is **positive definite** $\Leftrightarrow$ its eigenvalues are all positive. These types of matrices occur in many applications:

- Numerical solutions of boundary value problems
- Using finite difference methods
- Using finite element methods.

These are just a couple out of many, many examples. Keep your textbook and use it as a reference for later courses or applications. There are 140 pages we didn't get to!