Mathematical Techniques 3
Vector spaces, linear operators, and matrices

Alston J. Misquitta

Queen Mary University of London

September 25, 2018

Read this! I

These are not lecture notes!

These slides are merely an outline of what we will cover in the lectures. Use them as a guide, solve the problems indicated here, and follow-up on reading the material highlighted in the reference texts.

Outline of the Talk

1. Goals of MT3
2. Vector Spaces
3. Norm of a vector
4. Gram–Schmidt Orthogonalization
5. Linear Operators
6. Matrices
7. Operations on matrices
8. Function Spaces & Fourier Transforms

Goals of MT3 I

- Vector spaces
  1. Vector spaces
  2. Linear operators
  3. Matrices
  4. Basis functions
  5. Function spaces
  6. Fourier expansions (an example of a function space)
- Differential equations
  1. Ordinary differential equations
  2. Green's function methods
  3. Partial differential equations
- The Variational Principle

Outline of the Talk

1. Goals of MT3
2. Vector Spaces
3. Norm of a vector
4. Gram–Schmidt Orthogonalization
5. Linear Operators
6. Matrices
7. Operations on matrices
8. Function Spaces & Fourier Transforms

Vector Spaces I

Defn: A set of objects $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, etc. is said to form a linear vector space $\mathcal{V}$ if:

- The set is closed under commutative and associative addition:

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$$
$$\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$$

- The set is closed under multplication by a scalar, i.e., $\lambda\mathbf{a} \in \mathcal{V} \; \forall \lambda \in \mathbb{C}$.
- Multplication by a scalar is both distributive and associative:

$$\lambda(\mathbf{a} + \mathbf{b}) = \lambda\mathbf{a} + \lambda\mathbf{b}$$
$$\lambda(\mu\mathbf{a}) = (\lambda\mu)\mathbf{a}.$$

- There exists a null vector $\mathbf{0}$ s.t. $\mathbf{a} + \mathbf{0} = \mathbf{a}$.

Vector Spaces II

- Multplication by the unit scalar leaves any vector unchanged: $1\mathbf{a} = \mathbf{a}$.
- $\forall \mathbf{a} \in \mathcal{V} \ \exists \ -\mathbf{a}$ s.t. $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$.

See RHB §8.1 for more details.

Vector Spaces III

Linear dependencies
If $\nexists \ \alpha_i \neq 0$ s.t. $\sum_i^N \alpha_i \mathbf{a}_i = \mathbf{0}$ then the set $\{\mathbf{a}_i\}$ of N vectors is said to from a *linearly independent* set.

Dimension
In a space $\mathcal{V}$, if there are no more than N linearly independent vectors $\{\mathbf{a}_i\}$ then the space is said to have dimension N.

Vector Spaces IV

Basis sets

If $\mathcal{V}$ is an N-dimensional vector space then *any* set of N linearly independent vectors $\{\mathbf{e}_i\}$ forms a basis for $\mathcal{V}$.

If $\mathbf{x}$ is an arbitrary vector in $\mathcal{V}$ than the set $\{\mathbf{x}, \{\mathbf{e}_i\}\}$ *must be linearly dependent*. I.e., we must have

$$\sum_i^N \alpha_i \mathbf{e}_i + \beta \mathbf{x} = \mathbf{0},$$

where not all $\alpha_i = 0$, and $\beta \neq 0$.

Vector Spaces V

Since $\beta \neq 0$, we can define $x_i = -\alpha_i/\beta$ giving

$$\mathbf{x} = \sum_i^N x_i \mathbf{e}_i.$$

Q: Show that given a basis $\{\mathbf{e}_i\}$, the coefficients $\{x_i\}$ are unique.

See RHB §8.1.1 for more details.

Outline of the Talk

1 Goals of MT3
2 Vector Spaces
3 Norm of a vector
4 Gram–Schmidt Orthogonalization
5 Linear Operators
6 Matrices
7 Operations on matrices
8 Function Spaces & Fourier Transforms

The Norm I

Inner product

The inner product of two vectors results in a scalar $\langle \mathbf{a}|\mathbf{b}\rangle$ with the properties

- $\langle \mathbf{a}|\mathbf{b}\rangle = \langle \mathbf{b}|\mathbf{a}\rangle^*$, and
- $\langle \mathbf{a}|\lambda\mathbf{b} + \mu\mathbf{c}\rangle = \lambda\langle \mathbf{a}|\mathbf{b}\rangle + \mu\langle \mathbf{a}|\mathbf{c}\rangle$.

Q: Show that:
$\langle \lambda\mathbf{a} + \mu b|\mathbf{c}\rangle = \lambda^*\langle \mathbf{a}|\mathbf{c}\rangle + \mu^*\langle \mathbf{b}|\mathbf{c}\rangle$

Q: Show that:
$\langle \lambda\mathbf{a}|\mu\mathbf{b}\rangle = \lambda^*\mu\langle \mathbf{a}|\mathbf{b}\rangle$

The Norm II

Orthogonal vectors

$\mathbf{a}, \mathbf{b} \in \mathcal{V}$ are orthogonal iff $\langle \mathbf{a} | \mathbf{b} \rangle = 0$.

iff = if and only if

Norm of a vector

$\|\mathbf{a}\| = \sqrt{\langle \mathbf{a} | \mathbf{a} \rangle}$.

A *normalized* vector is one with has a unit norm. Any vector can be normalized as follows:

$$\mathbf{a} \to \frac{\mathbf{a}}{\|\mathbf{a}\|}.$$

The inner product $\langle a | a \rangle$ can have any sign. If we restrict it, as we will now do, to have $\langle a | a \rangle \geq 0$, then we get the Euclidean, or positive semi-definite norm.

The Norm III

Orthonormal basis
This is a basis set of orthogonal and normalised basis functions $\{\hat{\mathrm{e}}_i\}$ that satisfies

$$\langle \hat{\mathbf{e}}_i | \hat{\mathbf{e}}_j \rangle = \delta_{ij},$$

where δ_{ij} is the *Kronecker delta function* that is defined as

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

The Norm IV

Components of a vector:
Any vector $\mathbf{a} \in \mathcal{V}$ can be written as

$$\mathbf{a} = \sum_{i}^{N} a_i \hat{\mathbf{e}}_i,$$

where the components of $\mathbf{a}$ are the $\{a_i\}$ which are defined as

$$a_i = \langle \hat{\mathbf{e}}_i | \mathbf{a} \rangle$$

Q: Demonstrate this!

Q: Show that: $\langle a|b \rangle = \sum_{i=1}^{N} a_i^* b_i$.

Compare this defn of the inner product with the dot product.

The Norm V

What is the basis vectors are normalized but not orthogonal?
I.e., $\langle \hat{\mathbf{e}}_i | \hat{\mathbf{e}}_j \rangle = 1$, but

$$\langle \hat{\mathbf{e}}_i | \hat{\mathbf{e}}_j \rangle = \begin{cases} 1, & \text{if } i = j, \\ G_{ij}, & \text{if } i \neq j. \end{cases}$$

Q: Show that $\langle a | b \rangle = \sum_{ij}^{N} a_i^* G_{ij} a_j$.

Q: Show that if the norm of a vector is real, i.e, if $\|\mathbf{a}\| \in \mathbb{R}$, then $G_{ij} = G_{ji}^*$.

See RHB §8.1.2 for more details.

Outline of the Talk

1. Goals of MT3
2. Vector Spaces
3. Norm of a vector
4. Gram–Schmidt Orthogonalization
5. Linear Operators
6. Matrices
7. Operations on matrices
8. Function Spaces & Fourier Transforms

Orthogonalization I

Given a basis $\{\mathrm{e}_i\}$ of not necessarily normalization or orthogonalized vectors, we can create an orthogonalized basis $\{\hat{\mathrm{e}}'_i\}$ as follows:

$$\hat{\mathbf{e}}'_1 = \frac{\mathbf{e}_1}{\|\mathbf{e}_1\|}$$

$$\hat{\mathbf{e}}'_2 = \frac{\mathbf{e}_2 - \langle\hat{\mathbf{e}}_1|\mathbf{e}_2\rangle\hat{\mathbf{e}}_1}{\|\mathbf{e}_2 - \langle\hat{\mathbf{e}}_1|\mathbf{e}_2\rangle\hat{\mathbf{e}}_1\|}$$

$$\hat{\mathbf{e}}'_3 = \frac{\mathbf{e}_3 - \langle\hat{\mathbf{e}}_1|\mathbf{e}_3\rangle\hat{\mathbf{e}}_1 - \langle\hat{\mathbf{e}}_2|\mathbf{e}_3\rangle\hat{\mathbf{e}}_2}{\|\mathbf{e}_3 - \langle\hat{\mathbf{e}}_1|\mathbf{e}_3\rangle\hat{\mathbf{e}}_1 - \langle\hat{\mathbf{e}}_2|\mathbf{e}_3\rangle\hat{\mathbf{e}}_2\|}$$

$$\cdots = \cdots$$

See 2016 lecture notes for more details.

Outline of the Talk

1. Goals of MT3
2. Vector Spaces
3. Norm of a vector
4. Gram–Schmidt Orthogonalization
5. Linear Operators
6. Matrices
7. Operations on matrices
8. Function Spaces & Fourier Transforms

Linear Operators (RHB §8.2) I

A linear operator $\hat{A}$ on a vector space $\mathcal{V}$ associates every vector $\mathbf{x} \in \mathcal{V}$ with another vector $\mathbf{y} \in \mathcal{V}'$:

$$\mathbf{y} = \hat{A}\mathbf{x},$$

such that

- For $\mathbf{a}, \mathbf{b} \in \mathcal{V}$ and scalars $\lambda, \mu \in \mathbb{R}$,

$$\hat{A}(\lambda\mathbf{a} + \mu\mathbf{b}) = \lambda\hat{A}\mathbf{a} + \mu\hat{A}\mathbf{b}.$$

- $(\hat{A} + \hat{B})\mathbf{a} = \hat{A}\mathbf{a} + \hat{B}\mathbf{b}$.
- $(\hat{A}\hat{B})\mathbf{a} = \hat{A}(\hat{B}\mathbf{a})$.
- Null operator: $\hat{\mathcal{O}}\mathbf{a} = \mathbf{0}$.
- Identity: $\hat{\mathcal{I}}\mathbf{a} = \mathbf{a}$.
- If $\exists\hat{A}^{-1}$ s.t. $\hat{A}\hat{A}^{-1} = \hat{\mathcal{I}} = \hat{A}^{-1}\hat{A}$, then $\hat{A}^{-1}$ is the inverse of $\hat{A}$ and $\hat{A}$ is *non-singular*.

Linear Operators (RHB §8.2) II

$$\mathcal{V} \xrightarrow{\hat{A}} \mathcal{V}'$$

$$\{\hat{\mathbf{e}}_i\} \qquad \{\hat{\mathbf{f}}_j\}$$

$$N \qquad\qquad M$$

$$\mathbf{x} \xrightarrow{\hat{A}} \mathbf{y}$$

What is the action of $\hat{A}$ on a *basis function* of $\mathcal{V}$?
$\hat{A}$ transforms $\hat{\mathrm{e}}_i$ into a linear combination of basis functions $\{\hat{\mathbf{f}}_j\}$ that span space $\mathcal{V}'$:

$$\hat{A}\hat{\mathbf{e}}_i = \sum_{j=1}^{M} A_{ji}\hat{\mathbf{f}}_j, \quad i \in [1, N].$$

Here the A_{ji} are the scalars that determine the transformation.

Linear Operators (RHB §8.2) III

What is the action of $\hat{A}$ on a *vector* of $\mathcal{V}$?

$$\mathbf{x} = \sum_{i=1}^{N} x_i \hat{\mathbf{e}}_i \quad \in \mathcal{V}$$

$$\mathbf{y} = \sum_{j=1}^{M} y_j \hat{\mathbf{f}}_j \quad \in \mathcal{V}',$$

such that

$$\mathbf{y} = \hat{A}\mathbf{x}.$$

Q: Show that: $y_j = \sum_i^N A_{ji} x_i$.

Linear Operators (RHB §8.2) IV

$$y_j = \sum_{i=1}^{N} A_{ji} x_i$$

This can be represented as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1N} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & A_{M3} & \dots & A_{MN} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_N \end{bmatrix}$$

Linear Operators (RHB §8.2) V

If we use the notation $\mathbf{A}$ to denote the $M \times N$ object, then

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

This looks very similar to the operator form:

$$\mathbf{y} = \hat{A}\mathbf{x}.$$

But $\mathbf{A}$ is only a *representation* of the operator $\hat{A}$ in the chosen basis sets $\{\hat{\mathrm{e}}_i\}$ (for $\mathcal{V}$) and $\{\hat{\mathbf{f}}_j\}$ (for $\mathcal{V}'$).

Q: If we had used different basis sets then the terms in $\mathbf{A}$ would change, but the dimensions of $\mathbf{A}$ would always be $M \times N$. Explain why.

Outline of the Talk

1 Goals of MT3

2 Vector Spaces

3 Norm of a vector

4 Gram–Schmidt Orthogonalization

5 Linear Operators

6 Matrices

7 Operations on matrices

8 Function Spaces & Fourier Transforms

Matrices (RHB §8.3) I

Properties of linear ops in index notation

$$(\hat{A}+\hat{B})\mathbf{a} = \hat{A}\mathbf{a} + \hat{B}\mathbf{a}$$

becomes

$$\sum_{j=1}^{N}(\mathbf{A}+\mathbf{B})_{ij}a_j = \sum_{j=1}^{N} A_{ij}a_j + \sum_{j=1}^{N} B_{ij}a_j$$

As this must hold $\forall \mathbf{a} \in \mathcal{V}$ we must have

$$(\mathbf{A}+\mathbf{B})_{ij} = A_{ij} + B_{ij}.$$

This defines matrix addition.

Matrices (RHB §8.3) II

$$(\hat{A}\hat{B})\mathbf{a} = \hat{A}(\hat{B}\mathbf{a})$$

becomes

$$\begin{aligned}\sum_{j}^{N}(\mathbf{AB})_{ij}a_j &= \sum_{k}^{N} A_{ik}(\mathbf{Ba})_k \\ &= \sum_{k}^{N} A_{ik}\sum_{j}^{N} B_{kj}a_j \\ &= \sum_{j}^{N}\left(\sum_{k}^{N} A_{ik}B_{kj}\right)a_j.\end{aligned}$$

Matrices (RHB §8.3) III

As this must hold $\forall \mathbf{a} \in \mathcal{V}$ we must have

$$(\mathbf{AB})_{ij} = \sum_{k}^{N} A_{ik} B_{kj}.$$

This defines matrix multplication.

Matrices (RHB §8.3) IV

Similarly, the simplified version of multplication by a scalar:

$$(\lambda \hat{A})\mathbf{a} = \lambda(\hat{A}\mathbf{a})$$

implies

$$(\lambda \mathbf{A})_{ij} = \lambda A_{ij},$$

which defines how matrices can be multplied by a scalar.

Examples are given in RHB §8.4.1 and RHB §8.4.2.

Matrices (RHB §8.3) V

$$(\mathbf{AB})_{ij} = \sum_{k}^{N} A_{ik} B_{kj}.$$

Let $(\mathbf{AB})_{ij} = P_{ij} = (\mathbf{P})_{ij}$, so

$$\mathbf{P} = \mathbf{AB}$$

Q: What are the dimensions of $\mathbf{P}$? See RHB §8.4.2.

Matrices (RHB §8.3) VI

Matrix multplication is associative

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}.$$

Q: Prove it.

Is matrix multplication commutative?

$$\mathbf{P} = \mathbf{AB} \stackrel{?}{=} \mathbf{BA} = \mathbf{Q}$$

Q: We can consider the commutation only if $\mathbf{A}$ is $M \times N$ and $\mathbf{B}$ is $N \times M$. Why? In this case, what are the dimensions of $\mathbf{P}$ and $\mathbf{Q}$?

Matrices (RHB §8.3) VII

Matrix multplication is not in general commutative

$$\mathbf{AB} \neq \mathbf{BA}$$

Q: When will matrix multplication be commutative? When the matrices are square? Diagonal? Any other case?

Matrix multplication is distributive under addition

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}.$$

Q: Prove this.

Matrices (RHB §8.3) VIII

Null matrix: $\mathbf{0}$

$$\mathbf{0A} = \mathbf{0} = \mathbf{A0}$$
$$\mathbf{0} + \mathbf{A} = \mathbf{A} = \mathbf{A} + \mathbf{0}$$

Identity matrix: $\mathbf{I}$

$$\mathbf{IA} = \mathbf{A} = \mathbf{AI}$$

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

Matrices (RHB §8.3) IX

Transpose: $\mathbf{A}^{\mathrm{T}}$

$$(\mathbf{A}^{\mathrm{T}})_{ij} = (\mathbf{A})_{ji}$$

If $\mathbf{A}$ is $M \times N$ then $\mathbf{A}^{\mathrm{T}}$ is $N \times M$.

Q: Prove that $(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}$.

Matrices (RHB §8.3) X

Complex conjugate: $\mathbf{A}^*$: $(\mathbf{A}^*)_{ij} = (A_{ij})^*$

Hermitian conjugate or adjoint: $\mathbf{A}^\dagger$

$$\mathbf{A}^\dagger = (\mathbf{A}^*)^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}})^*.$$

$$\begin{aligned}(\mathbf{A}^\dagger)_{ij} &= [(\mathbf{A}^*)^{\mathrm{T}}]_{ij} \\ &= (\mathbf{A}^*)_{ji} = (A_{ji})^*.\end{aligned}$$

Q: Prove that $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$.

Q: Show that if $\mathbf{A}$ is real then $\mathbf{A}^\dagger = \mathbf{A}^{\mathrm{T}}$.

Matrices (RHB §8.3) XI

Notation

In general:

$$\langle \mathbf{a}|\mathbf{b}\rangle = \sum_{i=1}^{N} a_i^* b_i$$

$$= \begin{bmatrix} a_1^* & a_2^* & \ldots & a_N^* \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{bmatrix}$$

$$= \mathbf{a}^\dagger \mathbf{b}$$

For real vectors this becomes:

$$\langle \mathbf{a}|\mathbf{b}\rangle = \mathbf{a}^{\mathrm{T}} \mathbf{b}$$

Matrices (RHB §8.3) XII

If $\mathbf{a}$ and $\mathbf{b}$ are operated on by $\hat{A}$ and $\hat{B}$:

$$\begin{aligned}\langle \hat{A}\mathbf{a} | \hat{B}\mathbf{b} \rangle &= (\mathbf{A}\mathbf{a})^{\dagger}(\mathbf{B}\mathbf{b}) \\ &= \mathbf{a}^{\dagger}\mathbf{A}^{\dagger}\mathbf{B}\mathbf{b}\end{aligned}$$

For real vectors this becomes:

$$\begin{aligned}\langle \hat{A}\mathbf{a} | \hat{B}\mathbf{b} \rangle &= (\mathbf{A}\mathbf{a})^{\mathrm{T}}(\mathbf{B}\mathbf{b}) \\ &= \mathbf{a}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{B}\mathbf{b}\end{aligned}$$

Matrices (RHB §8.3) XIII

Rotations

In 2D, on rotation by θ anti-clockwise, the basis vectors $\{\hat{\mathrm{e}}_1, \hat{\mathrm{e}}_2\}$ transform into $\{\hat{\mathrm{e}}'_1, \hat{\mathrm{e}}'_2\}$ which are given by:

$$\hat{\mathbf{e}}'_1 = \cos\theta\ \hat{\mathbf{e}}_1 + \sin\theta\ \hat{\mathbf{e}}_2$$
$$\hat{\mathbf{e}}'_2 = -\sin\theta\ \hat{\mathbf{e}}_1 + \cos\theta\ \hat{\mathbf{e}}_2$$

These basis sets are *both* orthonormal basis sets of the vector space. The rotation operator $\hat{R}$ is defined through its action on a basis vector:

$$\hat{\mathbf{e}}'_j = \hat{R}\hat{\mathbf{e}}_j = \sum_i R_{ij}\hat{\mathbf{e}}_i$$

This allows us to define the matrix $\mathbf{R}$:

$$\mathbf{R} = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Matrices (RHB §8.3) XIV

Orthogonal matrices

The rotation matrix is an example of an *orthogonal* matrix as

$$\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I} = \mathbf{R}\mathbf{R}^{\mathrm{T}}.$$

Q: Show this.

The inverse of an orthogonal matrix is particularly easy to compute as from the above definition it follows that

$$\mathbf{R}^{-1} = \mathbf{R}^{\mathrm{T}}.$$

Matrices (RHB §8.3) XV

As you have shown in the second exercise set, this kind of operator also preserves the lengths of vectors:

$$\langle \mathbf{a}|\mathbf{a}\rangle = \langle \hat{R}\mathbf{a}|\hat{R}\mathbf{a}\rangle$$

To show this you could first show that

$$\hat{R}\mathbf{a} = \sum_i \left(\sum_{i'} R_{ii'} a_{i'} \right) \hat{\mathbf{e}}_i$$

and then show that the R.H.S equals the L.H.S. But there is a faster way that uses a result we have demonstrated earlier:

$$\begin{aligned} \langle \hat{A}\mathbf{a}|\hat{B}\mathbf{b}\rangle &= (\mathbf{A}\mathbf{a})^{\mathrm{T}}(\mathbf{B}\mathbf{b}) \\ &= \mathbf{a}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{B}\mathbf{b} \end{aligned}$$

Matrices (RHB §8.3) XVI

We have $\hat{A} = \hat{B} = \hat{R}$ so we get

$$\begin{aligned} \langle \hat{R}\mathbf{a} | \hat{R}\mathbf{a} \rangle &= \mathbf{a}^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{R}\mathbf{a} \\ &= \mathbf{a}^{\mathrm{T}}\mathbf{I}\mathbf{a} \\ &= \mathbf{a}^{\mathrm{T}}\mathbf{a} = \langle \mathbf{a} | \mathbf{a} \rangle. \end{aligned}$$

Where we have used $\mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}$. This is much simpler a proof!

Matrices (RHB §8.3) XVII

Hermitian operators and matrices

A matrix $\mathbf{A}$ is said to be Hermitian iff:

$$\mathbf{A} = \mathbf{A}^{\dagger}.$$

Equivalently, an operator $\hat{A}$ is said to be Hermitian iff:

$$\langle \mathbf{a} | \hat{A}\mathbf{b} \rangle = \langle \hat{A}\mathbf{a} | \mathbf{b} \rangle.$$

The equivalence of these definitions can be seen as follows:

$$\begin{aligned} \langle \mathbf{a} | \hat{A}\mathbf{b} \rangle &= \mathbf{a}^{\dagger}\mathbf{A}\mathbf{b} \\ &= \mathbf{a}^{\dagger}\mathbf{A}^{\dagger}\mathbf{b} \quad \text{because } \mathbf{A} = \mathbf{A}^{\dagger} \\ &= (\mathbf{A}\mathbf{a})^{\dagger}\mathbf{b} = \langle \hat{A}\mathbf{a} | \mathbf{b} \rangle. \end{aligned}$$

Matrices (RHB §8.3) XVIII

Unitary matrices

This is a special kind of Hermitian matrix for which

$$\mathbf{U}^{-1} = \mathbf{U}^{\dagger}.$$

We get the orthogonal matrices as a special cases of unitary matrices when the elements of $\mathbf{U}$ are real.

These matrices also preserve the norms of vectors.

Outline of the Talk

1. Goals of MT3
2. Vector Spaces
3. Norm of a vector
4. Gram–Schmidt Orthogonalization
5. Linear Operators
6. Matrices
7. **Operations on matrices**
8. Function Spaces & Fourier Transforms

Operations on matrices I

Now we define some operations on matrices. These will be:

- The trace of a matrix: $\mathrm{Tr}\left(\mathbf{A}\right)$
- The determinant of a matrix: $|\mathbf{A}|$
- The inverse of a matrix: $\mathbf{A}^{-1}$

All these operations can only be defined for square matrices.

Operations on matrices II

The trace of a matrix

$$\mathrm{Tr}\left(\mathbf{A}\right) = \sum_{i=1}^{N} A_{ii} = A_{11} + A_{22} + \cdots + A_{NN}.$$

Notice that the trace can only be defined for a square matrix.

Q: Show that the trace is a linear operation. That is

$$\mathrm{Tr}\left(\mathbf{A} + \mathbf{B}\right) = \mathrm{Tr}\left(\mathbf{A}\right) + \mathrm{Tr}\left(\mathbf{B}\right).$$

Also show that

$$\mathrm{Tr}\left(\mathbf{AB}\right) = \mathrm{Tr}\left(\mathbf{BA}\right).$$

Operations on matrices III

The determinant of a matrix

$$\det(\mathbf{A}) \equiv |\mathbf{A}| = \begin{vmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1N} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & A_{M3} & \dots & A_{MN} \end{vmatrix}$$

The determinant is, if you will, the magnitude of a matrix. For a 2×2 matrix it is

$$\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} = A_{11}A_{22} - A_{12}A_{21}.$$

Operations on matrices IV

For larger matrices we use *recursion* to define $|\mathbf{A}|$:

- **Minor**: The minor M_{ij} of the element A_{ij} of $\mathbf{A}$ of dimension $N \times N$ is the determinant of the $(N-1) \times (N-1)$ matrix formed by removing the i^{th} row and j^{th} column of $\mathbf{A}$.
- **Cofactor**: $C_{ij} = (-1)^{i+j} M_{ij}$.
- **Determinant**: Choose any row i, or column j:

$$
\begin{aligned}
|\mathbf{A}| &= \sum_{k=1}^{N} A_{ik} C_{ik} \\
&= \sum_{k=1}^{N} A_{kj} C_{kj}
\end{aligned}
$$

This is the Laplace expansion for the determinant.

Operations on matrices V

Properties of the determinant
(Mostly without proof)

- $$\left|\mathbf{A}^{\mathrm{T}}\right| = |\mathbf{A}|$$

  This means any theorem established for the rows also applies to the columns.

- $$|\mathbf{A}^*| = |\mathbf{A}|^*, \quad \text{and}$$
  $$\left|\mathbf{A}^{\dagger}\right| = |\mathbf{A}|^*$$

- If any two rows (or columns) are interchanged

  $$|\mathbf{A}| = -\,|\mathbf{A}_{\mathbf{i}\leftrightarrow\mathbf{j}}|$$

Operations on matrices VI

- If $\mathbf{A}' = \lambda \mathbf{A}$, then

$$\left|\mathbf{A}'\right| = \lambda^N \left|\mathbf{A}\right|.$$

- If any two rows or columns are equal then $|\mathbf{A}| = 0$.
- If a multiple of any row (or column) is added to another row (or column) then the determinant of the resulting matrix is unchanged. That is, if $(\mathbf{A}')_{ij} = A_{ij} + \lambda A_{kj}$ then $|\mathbf{A}'| = |\mathbf{A}|$.
-

$$|\mathbf{AB}| = |\mathbf{A}|\,|\mathbf{B}| = |\mathbf{BA}|.$$

Operations on matrices VII

We will use determinants in this course so check to see if you have understood how to manipulate them by showing that

$$|\mathbf{A}| = \begin{vmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & -2 & 1 \\ 3 & -3 & 4 & -2 \\ -2 & 1 & -2 & -1 \end{vmatrix} = 0$$

The solution to this problem is given in RHB §8.9.1.

Operations on matrices VIII

The inverse of a matrix

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}.$$

To construct $\mathbf{A}^{-1}$:

- Construct the matrix $\mathbf{C}$ where $C_{ik} = \text{cofactor}(A_{ik})$.
- Now define the elements of $\mathbf{A}^{-1}$:

$$(\mathbf{A}^{-1})_{ik} = \frac{(\mathbf{C}^{\mathrm{T}})_{ik}}{|\mathbf{A}|} = \frac{(\mathbf{C})_{ki}}{|\mathbf{A}|}$$

- **Important**: The inverse is not defined if $|\mathbf{A}| = 0$.

Operations on matrices IX

To prove that this definition does lead to an inverse we need the following result:

$$\sum_k C_{ki} A_{kj} = |\mathbf{A}| \, \delta_{ij}.$$

The proof of this is quite simple and can be found in RHB §8.10. Now we will use this to show that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$:

$$\begin{aligned} (\mathbf{A}^{-1}\mathbf{A})_{ij} &= \sum_k (\mathbf{A}^{-1})_{ik} A_{kj} \\ &= \sum_k \frac{(\mathbf{C})_{ki}}{|\mathbf{A}|} A_{kj} \\ &= \frac{|\mathbf{A}|}{|\mathbf{A}|} \delta_{ij} = \delta_{ij}. \end{aligned}$$

This proves the result.

Operations on matrices X

Properties of the inverse

- $$(\mathbf{A}^{-\mathbf{1}})^{-1} = \mathbf{A}$$
- $$(\mathbf{A}^{\mathrm{T}})^{-1} = (\mathbf{A}^{-1})^{\mathrm{T}}$$
- $$(\mathbf{A}^{\dagger})^{-1} = (\mathbf{A}^{-1})^{\dagger}$$
- $$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

Q: Show that $\left|\mathbf{A}^{-\mathbf{1}}\right| = |\mathbf{A}|^{-1}$.

Operations on matrices XI

Eigenvectors and eigenvalues

Consider an operator $\hat{A}$ for which, for all $\mathbf{x} \in V$, $\hat{A}\mathbf{x} \in V$, then it is possible that for some $\mathbf{x}$,

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

If $\mathbf{x} \neq \mathbf{0}$ then $\mathbf{x}$ is called an *eigenvector* or *eigenfunction* of $\hat{A}$, and λ is the corresponding *eigenvalue*.

eigen means *proper* or *characteristic*.

Operations on matrices XII

In matrix form, the eigenvalue equation is

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

Since $\mathbf{x}' = \mu\mathbf{x}$ will also be an eigenfunction with the same eigenvalue, we use only normalized eigenfunctions, i.e.,

$$\langle\mathbf{x}|\mathbf{x}\rangle = \mathbf{x}^{\mathrm{T}}\mathbf{x} = 1.$$

Notation

The eigenfunctions of the square matrix $\mathbf{A}$ will be denoted by $\mathbf{x}^i$, and eigenvalues by λ_i.

Operations on matrices XIII

Practical methods for eigenvalues and eigenvectors

Write the eigenvalue equation as

$$\begin{aligned} \mathbf{0} &= \mathbf{A}\mathbf{x} - \lambda\mathbf{I} \\ &= (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} \\ &= \mathbf{B}\mathbf{x} \end{aligned}$$

Where $\mathbf{B} = \mathbf{A} - \lambda\mathbf{I}$. Now if $\mathbf{B}^{-1}$ exists, then we can multiply with this inverse on both sides to show that:

$$\mathbf{B}^{-1}\mathbf{0} = \mathbf{0} = \mathbf{B}^{-1}\mathbf{B}\mathbf{x} = \mathbf{x}.$$

This solution, $\mathbf{x} = \mathbf{0}$ is known as the *trivial* solution.

Operations on matrices XIV

On the other hand, if $\mathbf{B}^{-1}$ *does not exist* then we will find the more interesting solutions. For the inverse not to exist we must have

$$0 = |\mathbf{B}| = |\mathbf{A} - \lambda\mathbf{I}| \, .$$

This is known as the *characteristic equation* and it results in a polynomial of order N (the dimension of this space) in λ. This can be solved to yield N roots which will be the required eigenvalues.

See RHB §8.14 for more details and the following examples that we will also solve in class.

Operations on matrices XV

Some examples from RHB §8.14:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 3 \\ 1 & 1 & -3 \\ 3 & -3 & -3 \end{bmatrix}$$

Set up the characteristic equation:

$$\begin{aligned} 0 &= |\mathbf{A} - \lambda \mathbf{I}| \\ &= \begin{vmatrix} 1-\lambda & 1 & 3 \\ 1 & 1-\lambda & -3 \\ 3 & -3 & -3-\lambda \end{vmatrix} \end{aligned}$$

Operations on matrices XVI

This leads to the polynomial equation for λ:

$$0 = (\lambda - 2)(\lambda - 3)(\lambda + 6).$$

Therefore the eigenvalues are

$$\begin{aligned} \lambda_1 &= 2 \\ \lambda_1 &= 3 \\ \lambda_1 &= -6 \end{aligned}$$

To find the eigenvectors we need to solve for each i:

$$\mathbf{A}\mathbf{x}^i = \lambda_i \mathbf{x}^i.$$

Operations on matrices XVII

For the first eigenvalue this is

$$\mathbf{A}\mathbf{x}^1 = 2\mathbf{x}^1$$

If $\mathbf{x}^1 = (a \;\; b \;\; c)^{\mathrm{T}}$ then we get

$$\begin{bmatrix} 1 & 1 & 3 \\ 1 & 1 & -3 \\ 3 & -3 & -3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 2 \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

This must be solved for $a, b,$ and c to get $a = b$ and $c = 0$, so $\mathbf{x}^1 = (a \;\; a \;\; 0)^{\mathrm{T}}$, and this must be normalized to get $\mathbf{x}^1 = \frac{1}{\sqrt{2}}(1 \;\; 1 \;\; 0)^{\mathrm{T}}$.

Q: Find the two other eigenfunctions of $\mathbf{A}$ and show that the three eigenfunctions are mutually orthogonal.

Operations on matrices XVIII

In this example, also from RHB §8.14.1 we see how to tackle the case when the eigenvalues are *degenerate*:

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

The characteristic equation for this matrix leads to the polynomial:

$$0 = (4-\lambda)(\lambda+2)^2.$$

Therefore the eigenvalues are

$$\lambda_1 = 4 \quad \lambda_1 = -2 \quad \lambda_1 = -2$$

Q: How do we find the eigenfunction of this matrix?

Operations on matrices XIX

Properties of the eigenvectors

The eigenvalues of an Hermitian matrix are real.

Outline of proof
A matrix is Hermitian iff $\mathbf{A} = \mathbf{A}^\dagger$. Consider the eigenvalue equation for $\mathbf{A}$:

$$\mathbf{A}\mathbf{x}^i = \lambda_i \mathbf{x}^i.$$

Take the adjoint of this equation to get

$$(\mathbf{x}^i)^\dagger \mathbf{A}^\dagger = \lambda_i^* (\mathbf{x}^i)^\dagger, \quad \text{and therefore}$$
$$(\mathbf{x}^i)^\dagger \mathbf{A} = \lambda_i^* (\mathbf{x}^i)^\dagger$$

Now take the inner product with $\mathbf{x}^i$ of both equations and subtract one from the other to show that $\lambda_i = \lambda_i^*$.

Operations on matrices XX

Eigenvectors corresponding to different eigenvalues of an Hermitian matrix are orthogonal.

Outline of proof

Start with two eigenvalue equations

$$\mathbf{A}\mathbf{x}^i = \lambda_i \mathbf{x}^i \quad \text{and}$$
$$\mathbf{A}\mathbf{x}^j = \lambda_j \mathbf{x}^j$$

Take the adjoint of one of these and then the inner product with the other eigenfunction.

Full proof in RHB §8.13.2.

Operations on matrices XXI

Some more important results that you should prove:

Q: A matrix $\mathbf{A}$ is anti-Hermitian if $\mathbf{A}^\dagger = -\mathbf{A}$. Prove that the eigenvalues of an anti-Hermitian matrix are purely imaginary.

Q: Prove that the eigenvalues of a unitary matrix have unit modulus.

Q: Given the eigenvalues and eigenfunctions of matrix $\mathbf{A}$, find the corresponding eigenvalues and eigenfunctions of $\mathbf{A}^{-1}$.

All of these have solutions in RHB.

Operations on matrices XXII

Change of basis : similarity transformation

We followed RHB §8.15 and RHB §8.16. Please see the examples in these sections.

Outline of the Talk

1. Goals of MT3
2. Vector Spaces
3. Norm of a vector
4. Gram–Schmidt Orthogonalization
5. Linear Operators
6. Matrices
7. Operations on matrices
8. Function Spaces & Fourier Transforms

Function Spaces & Fourier Transforms I

We mainly followed the 2016 lecture notes on these topics. But we also used RHB §17.1 and RHB §17.2 for orthogonal polynomials and the adjoint and Hermitian conjugate of operators.