

About AICE1004

Andersen Ang

ECS, Uni. Southampton, UK
andersen.ang@soton.ac.uk
Homepage angms.science

Version: September 28, 2024

First draft: July 6, 2023

Why math

Instructors

Syllabus

Time, Assessment and Contact

12-week tentative plan

Example exam question

Computer

Readings

After AICE1004

Before AICE1004

Hardest module this sem

- ▶ Math is hard
- ▶ Math is abstract
- ▶ Math needs memory
- ▶ Math needs understanding
- ▶ Math needs imagination
- ▶ Math needs practise

the most important module

ex falso quodlibet
reductio ad absurdum

$$((P \Rightarrow \perp) \wedge P) \Rightarrow \neg P$$



$$\forall X [X \neq \emptyset \Rightarrow \exists f : X \rightarrow \bigcup X \quad \forall A \in X (f(A) \in A)]$$

$$\infty 2^{\aleph_\alpha} = \aleph_{\alpha+1}$$

$$\frac{df}{dx}$$



$$\lim_{x \rightarrow x_0} f(x) \triangleq \forall \epsilon > 0, \exists \delta > 0$$

$$|x - x_0| < \delta$$

$$\Rightarrow |f(x) - f(x_0)| < \epsilon$$

$$f : [0, 1] \rightarrow \mathbb{R}, f \in C^0$$



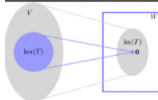
$$\forall \epsilon > 0, \exists p(x) \text{ s.t. } |f(x) - p(x)| < \epsilon$$



$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n = b_m \end{cases}$$



$$A^{-1} = \frac{1}{\det(A)} \sum_{s=0}^{n-1} A^s \sum_{k_1, \dots, k_{n-1}} \prod_{l=1}^{n-1} \frac{(-1)^{k_l+1}}{l^{k_l} k_l!} \text{tr}(A^l)^{k_l}$$



$$T: V \rightarrow W \quad \dim(\text{Im}T) + \dim(\text{null}T) = \dim(\text{dom}T)$$

$$\text{Im}T \oplus \text{Ker}T = V \quad \mathbf{e}_i \otimes \mathbf{e}_j$$

$$AB = 0 \iff \text{Im}B \subset \text{Ker}A$$



AI image generation

Input: "Students studying mathematics"

AI generate:



Why math is important

you need math to understand why AI works

$$p(\mathbf{x}_t | \mathbf{x}_{t-1}) = \prod_{i=1}^D p(x_{t,i} | \mathbf{x}_{t-1}) \quad p(\mathbf{x}_t | \mathbf{x}_{t-1}) = N(\mathbf{x}_t | \mu_t(\mathbf{x}_{t-1}), \Sigma_t(\mathbf{x}_{t-1})) \quad (1)$$

$$q(\mathbf{x}_{t-1} | \mathbf{x}_t) = \prod_{i=1}^D q(x_{t-1,i} | \mathbf{x}_t) \quad q(\mathbf{x}_{t-1} | \mathbf{x}_t) = N(\mathbf{x}_{t-1} | \sqrt{1 - \beta_t} \mathbf{x}_t, \beta_t \Sigma_t) \quad (2)$$

$$\mathbb{E}[-\log q(\mathbf{x}_t | \mathbf{x}_{t-1})] \leq \mathbb{E} \left[-\log \frac{p(\mathbf{x}_t | \mathbf{x}_{t-1})}{q(\mathbf{x}_t | \mathbf{x}_{t-1})} \right] = \mathbb{E} \left[-\log p(\mathbf{x}_t) - \sum_{i=1}^D \log \frac{p(x_{t,i} | \mathbf{x}_{t-1})}{q(x_{t,i} | \mathbf{x}_{t-1})} \right] = L \quad (3)$$

$$q(\mathbf{x}_t | \mathbf{x}_{t-1}) = N(\mathbf{x}_t | \sqrt{\beta_t} \mathbf{x}_{t-1}, (1 - \beta_t) \Sigma_t) \quad (4)$$

$$\mathbb{E} \left[\frac{p(x_{t,i} | \mathbf{x}_{t-1}) / q(x_{t,i} | \mathbf{x}_{t-1})}{\sum_{j=1}^D p(x_{t,j} | \mathbf{x}_{t-1}) / q(x_{t,j} | \mathbf{x}_{t-1})} \right] = \mathbb{E} \left[\frac{p(x_{t,i} | \mathbf{x}_{t-1})}{\sum_{j=1}^D p(x_{t,j} | \mathbf{x}_{t-1})} \right] \quad (5)$$

$$\mathbb{E} \left[\frac{p(x_{t,i} | \mathbf{x}_{t-1})}{\sum_{j=1}^D p(x_{t,j} | \mathbf{x}_{t-1})} \right] = \frac{p(x_{t,i} | \mathbf{x}_{t-1})}{\sum_{j=1}^D p(x_{t,j} | \mathbf{x}_{t-1})} \quad (6)$$

$$\mathbb{E} \left[\frac{p(x_{t,i} | \mathbf{x}_{t-1})}{\sum_{j=1}^D p(x_{t,j} | \mathbf{x}_{t-1})} \right] = \frac{p(x_{t,i} | \mathbf{x}_{t-1})}{\sum_{j=1}^D p(x_{t,j} | \mathbf{x}_{t-1})} \quad (7)$$

$$\mathbb{E} \left[\frac{p(x_{t,i} | \mathbf{x}_{t-1})}{\sum_{j=1}^D p(x_{t,j} | \mathbf{x}_{t-1})} \right] = \frac{p(x_{t,i} | \mathbf{x}_{t-1})}{\sum_{j=1}^D p(x_{t,j} | \mathbf{x}_{t-1})} \quad (8)$$

$$L_{t-1} - C = \mathbb{E} \left[\frac{1}{\beta_t} \left[\mu_t(\mathbf{x}_{t-1}) - \frac{1}{\sqrt{1 - \beta_t}} \mu_t(\mathbf{x}_t) - \mu_t(\mathbf{x}_t | \mathbf{x}_{t-1}) \right]^2 \right] \quad (9)$$

$$= \mathbb{E} \left[\frac{1}{\beta_t} \left[\frac{1}{\sqrt{1 - \beta_t}} \left(\mu_t(\mathbf{x}_{t-1}) - \frac{\beta_t}{\sqrt{1 - \beta_t}} \mu_t(\mathbf{x}_t) \right) - \mu_t(\mathbf{x}_t | \mathbf{x}_{t-1}) \right]^2 \right] \quad (10)$$

$$\mu_t(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mu_t \left(\mathbf{x}_{t-1} - \frac{1}{\sqrt{1 - \beta_t}} (\mathbf{x}_t - \sqrt{1 - \beta_t} \mathbf{x}_{t-1}) \right) = \frac{1}{\sqrt{1 - \beta_t}} \left(\mathbf{x}_{t-1} - \frac{\beta_t}{\sqrt{1 - \beta_t}} \mu_t(\mathbf{x}_t) \right) \quad (11)$$

$$\mathbb{E} \left[\frac{1}{\beta_t} \left[\frac{\beta_t}{\sqrt{1 - \beta_t}} \left(\mu_t(\mathbf{x}_{t-1}) - \mu_t(\mathbf{x}_t) \right) \right]^2 \right] \quad (12)$$

I intentionally blur the image to reduce fear

Your instructors

- ▶ Andersen (Man Shun) Ang (Module leader)

- ▶ Website: angms.science

- ▶ Math geek
Optimization Theorist
Your final boss this semester

$$\left\{ v \in \mathbb{R}^n \mid \begin{array}{l} \exists x_k \rightarrow x, f(x_k) \rightarrow f(x) \\ v_k \in \partial f(x_k), \lim_{k \rightarrow \infty} v_k = v \end{array} \right\}$$

- ▶ Office: Building 32 Room 3057

- ▶ Zhanxing Zhu

- ▶ Website: zhanxingzhu.github.io

- ▶ Badminton zealot
Deep Learning Theorist
DL uses lots of linear algebra \implies teach linear algebra

$$\begin{aligned} \mathcal{L}(N, D, T; f(x; \theta)) &= E + \frac{A}{N^\alpha} + \frac{B}{D^\beta} + \frac{C}{T^\gamma} \\ \text{s.t. } f(x; \theta) &= f_1 \circ f_2 \circ \dots \circ f_L(x) \end{aligned}$$

- ▶ Office: Building 32 Room 3XXX



Syllabus

Part 1. math of “formality” Logic & set

Proposition, truth, Id, \top , \perp , \neg , \wedge , \vee , $\bar{\wedge}$, $\underline{\vee}$, \forall , \exists , \implies , \iff , \vdash , \models , proofs

$\{\}$, \in , \notin , \subset , \cap , \cup , \setminus , Δ , A^c , \emptyset , 2^S , $|S|$, \aleph , ∞ , $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}, \mathbb{C}$, \times , R , $[\]$, $/ \sim$

dom, codom, $\text{Im}g$, $1 \rightarrow 1$, onto, 1-1

LEAN-4

Part 2. math of “change” Calculus

Number, elementary functions, inequality

calculus: differentiation, integration

analysis (calculus at proof level):
sequence, definition of limit, continuity,
L'Hospital rule, Taylor series, Riemann sum

Wolfram Alpha (Lab 1 M1)

Part 3. math of “structure” Linear algebra

Matrix-vector algebra, $A\mathbf{x} = \mathbf{b}$, Gaussian elimination

inverse, eigendecomposition, characteristic equation, determinant

Abstract linear algebra: combination, span, column space, independence, orthogonality, rank, dimension, kernel, quotient

NumPy (Lab 2 M2)

AICE1004 Mathematics for Artificial Intelligence and Computer Engineering (I)

Time commitment: 150 hours

Lecture	32h
Independent study	108h
Laboratories	10h

Assessment (School rule: $\geq 40\%$ to pass)

Computing Laboratories (lab)	5%
Coursework (mid-term test)	35%
Exam	60%

Online platform: Moodle

<https://moodle.ecs.soton.ac.uk/1>

Lecture notes

Exercises

Announcement

Lab

Ask question here so that everybody can see it

Lecture

Monday 11:00-11:55, 27/2001

Thursday 09:00-10:55, 35/1005

Friday 10:00-11:55, 27/2001

Map

<https://maps.southampton.ac.uk/>

Useful websites for new students

- ▶ SUSSED <https://sotonac.sharepoint.com/>
- ▶ StudentHub <https://www.southampton.ac.uk/studentservices/index.page>
 - ▶ information
 - ▶ fees, finance, accommodation
 - ▶ wellbeing, disability, careers
 - ▶ course administration, change program
- ▶ School rule
<https://www.southampton.ac.uk/about/governance/regulations-policies/taught-students/general/results-ug-integrated-masters>
9.1 Mark Scheme: Pass := ≥ 40
- ▶ <https://data.soton.ac.uk/>
 - ▶ Map, Building
 - ▶ Which week it is , Term date, University Calendar
- ▶ <https://timetable.soton.ac.uk>
- ▶ <https://moodle.ecs.soton.ac.uk/login>

12-week tentative plan

w	Topic, by	Content	Assessment
1	Logic, A	Proposition, truth, Id, \top , \perp , \neg , \wedge , \vee , $\bar{_}$, $\underline{_}$, \forall , \exists , \implies , \iff , \vdash , \models , proof	
2	Set, A	$\{\}$, \in , \notin , \subset , \cap , \cup , \emptyset , $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$, \times , R , $1 \rightarrow 1$, onto, $1-1$, 2^A , $ S $, \aleph	lab1 (2.5%)
3	Calculus, A	Number system, elementary functions,	mid-term (35%)
4	Calculus, A	Limit, continuity, derivative	
5	Calculus, A	Differentiation, Taylor series	
6	Calculus, A	antiderivative, integration, Riemann sum, Fundamental theorem of calculus	
7	LA, Z	Vector and matrix algebra, their geometric meaning	
8	LA, Z	Linear system, Gaussian Elimination, their geometric meaning	
9	LA, Z	Inverse, det, characteristic equation, Eigenvalue, eigenvector	
10	LA, Z	vector space: comb, span, Col, basis, dim, rank, ker, quotient	
11	Revision, A	revision of week 1 - week 6	lab2(2.5%)
12	Revision, Z	revision of week 7 - week 10	

Exam (60%)

	October 2024				November 2024				December 2024				January 2025				
Mon	7	14	21	28	4	11	18	25	2	9	16	23	6	13	20		
Tue	1	8	15	22	29	5	12	19	26	3	10	17	24	7	14	21	
Wed	2	9	16	23	30	6	13	20	27	4	11	18	25	1	8	15	22
Thu	3	10	17	24	31	7	14	21	28	5	12	19	26	2	9	16	23
Fri	4	11	18	25	1	8	15	22	29	6	13	20	27	3	10	17	24
Sat	5	12	19	26	2	9	16	23	30	7	14	21	28	4	11	18	25
Sun	6	13	20	27	3	10	17	24	1	8	15	22	29	5	12	19	26
Semester Wks	1	2	3	4	5	6	7	8	9	10	11			12	Ex	Ex	

Good news

No homework
(tons of exercises in MadBook)

You probably learned Calculus &
Linear Algebra

So you are familiar with it

Bad news

Mid-term in Week 6 & Final exam
(Mock midterm/exam will be provided)

We will look at them at abstract level &
deeper level

We do proof

Example exam question

Given a matrix $\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$.

The eigenvalues of \mathbf{A} are (there are more than one option that is correct, you get one point by picking all the correct options, you get zero point otherwise):

- (a) 2
- (b) -2
- (c) 5
- (d) -5

Solution: (a) and (d)

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= 0 \\ \iff \begin{vmatrix} 1 - \lambda & 2 \\ 3 & -4 - \lambda \end{vmatrix} &= 0 \\ \iff (1 - \lambda)(-4 - \lambda) - 2 \cdot 3 &= 0 \\ \iff (\lambda - 1)(\lambda + 4) - 6 &= 0 \\ \iff \lambda^2 + 3\lambda - 10 &= 0 \\ \iff (\lambda - 2)(\lambda + 5) &= 0 \\ \iff \lambda \in \{+2, -5\} \end{aligned}$$

Marking:

- Pick (a) and (d) 1 point
- Otherwise 0 point

Assessment (School rule: > 40% to pass)

- ▶ L_1 : result of lab1 2.5%
- ▶ L_2 : result of lab2 2.5%
- ▶ T : mid-term test 35%
- ▶ E : exam 60%
- ▶ G : grade
- ▶ Example: appear in all the labs, test and exam $G = 0.025L_1 + 0.025L_2 + 0.35T + 0.6E$.
- ▶ Example: missed lab1 $G = 0.025L_2 + 0.35T + 0.625E$
- ▶ Example: missed test $G = 0.025L_1 + 0.025L_2 + 0.95E$
- ▶ Example: missed lab1 and test $G = 0.025L_2 + 0.975E$
- ▶ Example: missed lab1, lab2 and test $G = E$
- ▶ Example: missed exam $G = 0.025L_1 + 0.025L_2 + 0.35T$
- ▶ Fun fact: you can actually use one equation to express all the situations above

Computer

~~Lab M0 Logic: LEAN 4.0~~

Lab M1 Calculus: Wolfram Alpha (Mathematica)

Lab M2 Linear Algebra: Python (NumPy) / MATLAB

Readings: Lecture notes and MadBookPro are sufficient

- ▶ Make your own notes!
- ▶ Want more?
 - ▶ Discrete mathematics and its applications by Kenneth H. Rosen QA39 ROS
Logic and Proofs, Sets, Functions, Matrices
 - ▶ Calculus : a complete course (10th ed) by Robert Adams & Christopher Essex QA303 ADA
Chapter 1,2,5,6,10
 - ▶ Calculus (3rd ed) by Michael Spivak QA303 SPI
First 400 pages
 - ▶ Schaum's outlines calculus QA303 AYR, library has ebook
Chapter 1-35 or the whole book
 - ▶ Schaum's outlines, advanced calculus QA303 SPI, library has ebook
Chapter 3,4,5,7
 - ▶ Introduction to Applied Linear Algebra - Vectors, Matrices, and Least Squares by Stephen Boyd and Lieven Vandenberghe
Free pdf <https://web.stanford.edu/~boyd/vmls/vmls.pdf>
Slide by Boyd & Vandenberghe <https://web.stanford.edu/~boyd/vmls/vmls-slides.pdf>
 - ▶ Introduction to linear algebra by Gilbert Strang QA184.2 STR
Solving linear equations, Vector spaces and subspaces, Orthogonality, Determinants, Eigenvalues and eigenvectors
 - ▶ Schaum's outlines, Linear algebra QA251 LIP, library has ebook
Chapter 1,2,3,4,8

AICE1004 (Academic Year 2024/25)

Reading for AICE1004

View Online



7 items

Discrete mathematics and its applications - Kenneth H. Rosen, 2019

[Book](#) | Additional

Schaum's Outline of Discrete Mathematics (3rd Edition) - Seymour Lipschutz, Marc Lipson, 2007

[Book](#) | Additional

Calculus : a complete course - Robert A. Adams, Christopher Essex, 2022

[Book](#) | Additional

Calculus - Frank Ayres, Elliott Mendelson, ©2013

[Book](#) | Additional

Advanced calculus - Robert C. Wrede, Murray R. Spiegel, ©2010

[Book](#) | Additional

Introduction to linear algebra - Gilbert Strang, 2016

[Book](#) | Additional

Linear algebra - Seymour Lipschutz, Marc Lipson, 2013

[Book](#) | Additional

YouTube

- ▶ Essence of calculus by 3Blue1Brown

<https://www.youtube.com/playlist?list=PLZHQObOWTQDMsr9K-rj53DwVRMY03t5Yr>

- ▶ Essence of linear algebra by 3Blue1Brown

https://www.youtube.com/playlist?list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab

- ▶ Video 1 - 7 in Vectors for multi-variable calculus by Michael Penn

<https://www.youtube.com/playlist?list=PL22w63XsKjqyC3bWVd5EjGEVN9pqq55yF>

- ▶ Techniques of Integration by Michael Penn

https://www.youtube.com/playlist?list=PL22w63XsKjqy_Hgl0PAhSNsTCx84UpARt

The remaining slides are used to warn you that math is hard

HARDEST module this sem

- ▶ Math is hard
- ▶ Math is abstract
- ▶ Math needs memory
- ▶ Math needs understanding
- ▶ Math needs imagination
- ▶ Math needs practise

the most important module in AI

In lecturer

Your eye/ear: oh I understand!

Your brain/hand: no you don't!

Watch somebody doing math \neq learning

you do math \neq learning

$(\text{you do math}) \wedge ((\text{get it correct}) \vee (\text{get it wrong but understand why wrong})) = \text{learning}$

Warning: Math needs memory

Math is like $A \rightarrow B \rightarrow C \rightarrow D \rightarrow \dots$

If you forgot A , you can't proceed to B, C, \dots

Example.

Find the eigenvector of $\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$

you first need to know what are *vectors*,
eigenvalues, so on

Warning: Math needs imagination

Math is abstract, pure memory is useless

you don't know what is it if you can't image a picture

Example.

$\det \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$

= the area of parallelogram of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 2 \\ -4 \end{bmatrix}$

Warning: Math needs logical thinking

Math is structural: knowing a bunch of unstructured information is useless

If you can't form a "logic chain" in your head, you will fail

Example. Scaling a row by a constant multiplies the determinant by that constant.

If we scale the first row of $\begin{bmatrix} -4 & 3 & 3 \\ 8 & 7 & 3 \\ 4 & 3 & 3 \end{bmatrix}$, then $\det \begin{bmatrix} -8 & 6 & 6 \\ 8 & 7 & 3 \\ 4 & 3 & 3 \end{bmatrix} = 2 \det \begin{bmatrix} -4 & 3 & 3 \\ 8 & 7 & 3 \\ 4 & 3 & 3 \end{bmatrix}$.

$$\det \begin{bmatrix} -8 & 6 & 6 \\ 8 & 7 & 3 \\ 4 & 3 & 3 \end{bmatrix} = \det \left(\begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -4 & 3 & 3 \\ 8 & 7 & 3 \\ 4 & 3 & 3 \end{bmatrix} \right)$$

$$= \det \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \det \begin{bmatrix} -4 & 3 & 3 \\ 8 & 7 & 3 \\ 4 & 3 & 3 \end{bmatrix} \quad \text{based on } \det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$$

$$= 2 \det \begin{bmatrix} -4 & 3 & 3 \\ 8 & 7 & 3 \\ 4 & 3 & 3 \end{bmatrix} \quad \text{det of diagonal matrix is the product of diagonal}$$

High school / A-level / IB / GCSE / blablabla math are nothing

Bad math in high school:

No worry, work hard will be ok

Good math in high school:

Nothing special, you are only δ -better

{bad high-school math} = 50

{good high-school math} = 85

{math} = 4000000000000000000000000000¹

{good high-school math} is only 0.000000000000000000000008% better

Me? May be 9000 \ll 4000000000000000000000000000

¹this number is the number of vertices of a “math knowledge graph”

In math you don't understand things, you just get used to them — von Neumann

See a new concept, **you don't feel “what's going on” until you do a tons of examples**

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \mathbf{B} = \begin{bmatrix} e & f \\ g & h \end{bmatrix} \implies \mathbf{A} \oplus \mathbf{B} = \begin{bmatrix} a & b & & \\ c & d & & \\ & & e & f \\ & & g & h \end{bmatrix}$$

$$\mathbf{A} = [a \quad b], \mathbf{B} = \begin{bmatrix} e \\ f \end{bmatrix} \implies \mathbf{A} \oplus \mathbf{B} = \begin{bmatrix} a & b & \\ & & e \\ & & f \end{bmatrix}$$

$$\mathbf{A} = [a \quad b], \mathbf{B} = [e \quad f] \implies \mathbf{A} \oplus \mathbf{B} = [a \quad b \quad e \quad f]$$

$$\mathbf{A} = \begin{bmatrix} a \\ b \end{bmatrix}, \mathbf{B} = \begin{bmatrix} e \\ f \end{bmatrix} \implies \mathbf{A} \oplus \mathbf{B} = \begin{bmatrix} a \\ b \\ e \\ f \end{bmatrix}$$

von Neumann is know as the “father of computer”, but he is also a math god: C*-algebra, von Neumann mutual information, von Neumann decomposition, von Neumann spectral theorem

Math is crazy

- ▶ Cantor's diagonal prove of ∞
- ▶ Godel's completeness Theorem
- ▶ Godel-Maltsev Compactness Theorem of 1st-order logic
- ▶ Zermelo-Fraenkel set theory and Axiom of Choice
- ▶ Zorn's Lemma of partially ordered set
- ▶ Axioms of integer and Axioms of real number
- ▶ Bolzano-Weierstrass Theorem
- ▶ Robinson's Transfer Principle and Hyper-reals
- ▶ Weierstrass Approximation Theorem of continuous function
- ▶ Fermat's rule, Rolle's Theorem, Lagrange Theorem, Cauchy Theorem
- ▶ Fundamental Theorem of calculus
- ▶ Rank-null Theorem
- ▶ Fundamental Theorem of Linear Algebra
- ▶ Fundamental Homomorphism Theorem of vector space

Math is crazy (things we may talk about in the course)

- ▶ Not everything can be a set, for example $R := \{x : x \in x\}$
- ▶ There are more real numbers in $[0, 1]$ than all integers
- ▶ There are things being “too large” to be a set, for example $C = \{\aleph_0, \aleph_1, \aleph_2, \dots\}$
- ▶ The sum of all real number in the interval $[0, 1]$ is impossible to compute
- ▶ Some function is impossible to differentiate
- ▶ $\frac{d^{0.5}}{d^{0.5}x}x = \text{????}$
- ▶ $\int \frac{1}{x} dx = \ln|x| + C$ is wrong.
- ▶ Some function is impossible to integrate, for example $\int_0^1 \sin\left(\frac{1}{x}\right) dx$
- ▶ All vector space has a basis

Example: proving converge of sequence using logic

A sequence $\{s_n\} \in \mathbb{R}$ converges to $L \in \mathbb{R}$ if and only if

$$(\forall \epsilon > 0)(\exists K \in \mathbb{N})(\forall n \geq K \implies |s_n - L| < \epsilon)$$

Now prove $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$ using logic

Prove $0.99999999... = 1$

Example: calculus

You may think calculus is easy: $\int x^5 dx = \frac{1}{6}x^6 + C$

How about: let a sequence $f_n(x) = \cos(nx)$

$f_1(x) = \cos x$, $f_2(x) = \cos(2x)$, $f_3(x) = \cos(3x)$, ...

Now,

$$\int \lim_{n \rightarrow \infty} f_n(x) dx \stackrel{?}{=} \lim_{n \rightarrow \infty} \int f_n(x) dx$$

- ▶ Does left hand side exists?
- ▶ Does right hand side exists?
- ▶ Does the equality holds?
- ▶ Why is it true?
- ▶ Why is it not true?
- ▶ What if $\int f^{-1}(x) dx$?

Example: abstract linear algebra (You will learn this)

Invertible matrix theorem If $A \in \mathbb{R}^{n \times n}$, then following are all equivalent (either all true or all false for A)

- \iff 1 A is non-singular $\iff A^{-1}$ exists
- \iff 2a A has left inverse under matrix multiplication \iff there exists B such that $BA = I$
- \iff 2b A has right inverse under matrix multiplication \iff there exists C such that $AC = I$
- \iff 2c A has inverse under matrix multiplication \iff there exists D such that $DA = AD = I$
- \iff 3a the reduced row-echelon form of A is I_n
- \iff 3b the reduced column-echelon form of A is I_n
- \iff 3c A has n pivots
- \iff 3d A can be expressed as a product of elementary matrices $\iff A = E_1 E_2 \cdots$
- \iff 4a $A : x \mapsto Ax$ is surjective \iff there exists at least one x that solves $Ax = b$
- \iff 4b $A : x \mapsto Ax$ is injective \iff there exists at most one x that solves $Ax = b$
- \iff 4c $A : x \mapsto Ax$ is bijective \iff there exists exactly one x that solves $Ax = b$
- \iff 4d $Ax = b$ is consistent for all $b \in \mathbb{R}^n$
- \iff 5a columns of A are linearly independent
- \iff 5b rows of A are linearly independent
- \iff 5c columns of A span \mathbb{R}^n
- \iff 5d rows of A span \mathbb{R}^n
- \iff 5e columns of A form a basis for \mathbb{R}^n
- \iff 5f rows of A form a basis for \mathbb{R}^n
- \iff 6a A has rank n
- \iff 6b A has nullity 0
- \iff 6c $Ax = 0$ has only the trivial solution $x = 0$ $\iff \ker A = \{0_n\}$
- \iff 6d the orthogonal complement of the null space of A is \mathbb{R}^n
- \iff 6e the orthogonal complement of the row space of A is 0_n
- \iff 6f the range of A is \mathbb{R}^n
- \iff 7a $\det(A) \neq 0$
- \iff 7b the eigenvalues of A contains no zero $\iff 0$ is not an eigenvalue of A
- \iff 8a A^T is invertible
- \iff 8b $A^T A$ is invertible

If you want even more exercises

- ▶ (3193) Problems in Mathematical Analysis
Boris Demidovich
- ▶ 3000 Solved Problems in Calculus
Elliot Mendelson
- ▶ 3000 Solved Problems in Linear Algebra
Lipschutz Seymour
- ▶ **How to study mathematics**
Pattern Recognition
Math is unlike other subject, all problems are unique, problem will not repeat, this is why memorization is useless in math

1466. $\int \sin\left(\frac{\pi}{4} - x\right)\sin\left(\frac{\pi}{4} + x\right) dx.$	1484. $\int \sinh x \cosh x dx.$
1467. $\int \tan^3\left(\frac{x}{2} + \frac{\pi}{4}\right) dx.$	1485. $\int \frac{\sinh \sqrt{1-x}}{\sqrt{1-x}} dx.$
1468. $\int \frac{dx}{2 \sin x + 3 \cos x - 5}.$	1486. $\int \frac{\sinh x \cosh x}{\sinh^2 x + \cosh^2 x} dx.$
1469. $\int \frac{dx}{2 + 3 \cos^2 x}.$	1487. $\int \frac{x}{\sinh^2 x} dx.$
1470. $\int \frac{dx}{\cos^2 x + 2 \sin x \cos x + 2 \sin^2 x}.$	1488. $\int \frac{dx}{e^{2x} - 2e^x}.$
1471. $\int \frac{dx}{\sin x \sin 2x}.$	1489. $\int \frac{e^x}{e^{2x} - 6e^x + 13} dx.$
1472. $\int \frac{dx}{(2 + \cos x)(3 + \cos x)}.$	1490. $\int \frac{e^{2x}}{(e^x + 1)^4} dx.$
1473. $\int \frac{\sec^2 x}{\sqrt{\tan^2 x + 4 \tan x + 1}} dx.$	1491. $\int \frac{2^x}{1 - 4^x} dx.$
1474. $\int \frac{\cos ax}{\sqrt{a^2 + \sin^2 ax}} dx.$	1492. $\int (x^2 - 1) 10^{-2x} dx.$
1475. $\int \frac{x dx}{\cos^2 3x}.$	1493. $\int \sqrt{e^x + 1} dx.$
1476. $\int x \sin^2 x dx.$	1494. $\int \frac{\arctan x}{x^2} dx.$
1477. $\int x^2 e^{x'} dx.$	1495. $\int x^3 \arcsin \frac{1}{x} dx.$
1478. $\int x e^{2x} dx.$	1496. $\int \cos(\ln x) dx.$
1479. $\int x^2 \ln \sqrt{1-x} dx.$	1497. $\int (x^3 - 3x) \sin 5x dx.$
1480. $\int \frac{x \arctan x}{\sqrt{1+x^2}} dx.$	1498. $\int x \arctan(2x+3) dx.$
1481. $\int \sin^2 \frac{x}{2} \cos \frac{3x}{2} dx.$	1499. $\int \arcsin \sqrt{x} dx.$
1482. $\int \frac{dx}{(\sin x + \cos x)^2}.$	1500. $\int x dx.$
1483. $\int \frac{dx}{(\tan x + 1) \sin^2 x}.$	

What comes after AICE1004?

- ▶ Vector calculus (AICE1008 Math II)
⊂ Linear algebra ⊕ Calculus
- ▶ Scientific Computing (AICE2001)
≈ Linear Algebra II
- ▶ Optimization (COMP6260)
queen of applied math
- ▶ Deep Learning
≈ Optimization ⊕ Statistics
queen of machine learning(?)

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{bmatrix}, \oint, \iint, (\nabla A) \cdot \mathbf{F}$$

$$G = (Df)^\top Df = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_2} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_2} \end{bmatrix}$$

$$du \otimes dv, \quad du \wedge dv$$

How to study math

- ▶ Bring pen, papers & computer,
 - ▶ Do math in all lectures
 - ▶ Drop notes in lectures
 - ▶ make your OWN notes
- ▶ Watch YouTube math videos
- ▶ Develop your “math brain”
Make sure you
 - ▶ fully understand basic concepts without needing to think
 - ▶ get VERY comfortable with the terminology
- ▶ study to FULL comprehension & be able to APPLY takes PATIENCE
Lectures are only 32 hours, you bear the responsibility of that 108 hours²

²<https://www.southampton.ac.uk/courses/modules/aice1004>

Things not-a-must-to-do but good for you

Learn LaTeX

Learn Beamer

Learn Markdown

My teaching style: no slide.

What's next: revision of high school math