

Searching for Universal Truths

Algebra

Sunghee Yun
sunghee.yun@gmail.com

Navigating Mathematical and Statistical Territories

- Notations & definitions & conventions
 - notations - 2
 - some definitions - 6
 - some conventions - 7
- Algebra - 8
 - inequalities - 9
 - number theory - 34
- Proof & references & indices
 - references - 42
 - index - 44

Notations

- sets of numbers
 - \mathbf{N} - set of natural numbers
 - \mathbf{Z} - set of integers
 - \mathbf{Z}_+ - set of nonnegative integers
 - \mathbf{Q} - set of rational numbers
 - \mathbf{R} - set of real numbers
 - \mathbf{R}_+ - set of nonnegative real numbers
 - \mathbf{R}_{++} - set of positive real numbers
 - \mathbf{C} - set of complex numbers
- sequences $\langle x_i \rangle$ and the like
 - finite $\langle x_i \rangle_{i=1}^n$, infinite $\langle x_i \rangle_{i=1}^\infty$ - use $\langle x_i \rangle$ whenever unambiguously understood
 - similarly for other operations, *e.g.*, $\sum x_i$, $\prod x_i$, $\cup A_i$, $\cap A_i$, $\times A_i$
 - similarly for integrals, *e.g.*, $\int f$ for $\int_{-\infty}^\infty f$
- sets
 - \tilde{A} - complement of A

- $A \sim B$ - $A \cap \tilde{B}$
- $A \Delta B$ - $(A \cap \tilde{B}) \cup (\tilde{A} \cap B)$
- $\mathcal{P}(A)$ - set of all subsets of A
- sets in metric vector spaces
 - \overline{A} - closure of set A
 - A° - interior of set A
 - **relint** A - relative interior of set A
 - **bd** A - boundary of set A
- set algebra
 - $\sigma(\mathcal{A})$ - σ -algebra generated by \mathcal{A} , *i.e.*, smallest σ -algebra containing \mathcal{A}
- norms in \mathbf{R}^n
 - $\|x\|_p$ ($p \geq 1$) - p -norm of $x \in \mathbf{R}^n$, *i.e.*, $(|x_1|^p + \cdots + |x_n|^p)^{1/p}$
 - *e.g.*, $\|x\|_2$ - Euclidean norm
- matrices and vectors
 - a_i - i -th entry of vector a
 - A_{ij} - entry of matrix A at position (i, j) , *i.e.*, entry in i -th row and j -th column
 - $\text{Tr}(A)$ - trace of $A \in \mathbf{R}^{n \times n}$, *i.e.*, $A_{1,1} + \cdots + A_{n,n}$

- symmetric, positive definite, and positive semi-definite matrices
 - $\mathbf{S}^n \subset \mathbf{R}^{n \times n}$ - set of symmetric matrices
 - $\mathbf{S}_+^n \subset \mathbf{S}^n$ - set of positive semi-definite matrices; $A \succeq 0 \Leftrightarrow A \in \mathbf{S}_+^n$
 - $\mathbf{S}_{++}^n \subset \mathbf{S}^n$ - set of positive definite matrices; $A \succ 0 \Leftrightarrow A \in \mathbf{S}_{++}^n$
- sometimes, use Python script-like notations (with serious abuse of mathematical notations)
 - use $f : \mathbf{R} \rightarrow \mathbf{R}$ as if it were $f : \mathbf{R}^n \rightarrow \mathbf{R}^n$, *e.g.*,

$$\exp(x) = (\exp(x_1), \dots, \exp(x_n)) \quad \text{for } x \in \mathbf{R}^n$$

and

$$\log(x) = (\log(x_1), \dots, \log(x_n)) \quad \text{for } x \in \mathbf{R}_{++}^n$$

which corresponds to Python code `numpy.exp(x)` or `numpy.log(x)` where `x` is instance of `numpy.ndarray`, *i.e.*, numpy array

- use $\sum x$ to mean $\mathbf{1}^T x$ for $x \in \mathbf{R}^n$, *i.e.*

$$\sum x = x_1 + \dots + x_n$$

which corresponds to Python code `x.sum()` where `x` is numpy array

- use x/y for $x, y \in \mathbf{R}^n$ to mean

$$\begin{bmatrix} x_1/y_1 & \cdots & x_n/y_n \end{bmatrix}^T$$

which corresponds to Python code `x / y` where `x` and `y` are 1-d numpy arrays

- use X/Y for $X, Y \in \mathbf{R}^{m \times n}$ to mean

$$\begin{bmatrix} X_{1,1}/Y_{1,1} & X_{1,2}/Y_{1,2} & \cdots & X_{1,n}/Y_{1,n} \\ X_{2,1}/Y_{2,1} & X_{2,2}/Y_{2,2} & \cdots & X_{2,n}/Y_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m,1}/Y_{m,1} & X_{m,2}/Y_{m,2} & \cdots & X_{m,n}/Y_{m,n} \end{bmatrix}$$

which corresponds to Python code `X / Y` where `X` and `Y` are 2-d numpy arrays

Some definitions

Definition 1. [infinitely often - i.o.] *statement P_n , said to happen infinitely often or i.o. if*

$$(\forall N \in \mathbf{N}) (\exists n > N) (P_n)$$

Definition 2. [almost everywhere - a.e.] *statement $P(x)$, said to happen almost everywhere or a.e. or almost surely or a.s. (depending on context) associated with measure space (X, \mathcal{B}, μ) if*

$$\mu\{x | P(x)\} = 1$$

or equivalently

$$\mu\{x | \sim P(x)\} = 0$$

Some conventions

- (for some subjects) use following conventions

- $0 \cdot \infty = \infty \cdot 0 = 0$

- $(\forall x \in \mathbf{R}_{++})(x \cdot \infty = \infty \cdot x = \infty)$

- $\infty \cdot \infty = \infty$

Algebra

Inequalities

Jensen's inequality

- strictly convex function: for any $x \neq y$ and $0 < \alpha < 1$ (Definition ??)

$$\alpha f(x) + (1 - \alpha)f(y) > f(\alpha x + (1 - \alpha)y)$$

- convex function: for any x, y and $0 < \alpha < 1$ (Definition ??)

$$\alpha f(x) + (1 - \alpha)f(y) \geq f(\alpha x + (1 - \alpha)y)$$

Inequality 1. [Jensen's inequality - for finite sequences] *for convex function f and distinct x_i and $0 < \alpha_i < 1$ with $\alpha_1 + \cdots = \alpha_n = 1$*

$$\alpha_1 f(x_1) + \cdots + \alpha_n f(x_n) \geq f(\alpha_1 x_1 + \cdots + \alpha_n x_n)$$

- *if f is strictly convex, equality holds if and only if $x_1 = \cdots = x_n$*

Jensen's inequality - for random variables

- discrete random variable interpretation of Jensen's inequality in summation form - assume $\mathbf{Prob}(X = x_i) = \alpha_i$, then

$$\mathbf{E} f(X) = \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n) \geq f(\alpha_1 x_1 + \cdots + \alpha_n x_n) = f(\mathbf{E} X)$$

- true for any random variables X

Inequality 2. [Jensen's inequality - for random variables] for random vector X
(page ?? for definition)

$$\mathbf{E} f(X) \geq f(\mathbf{E} X)$$

if probability density function (PDF) p_X given,

$$\int f(x) p_X(x) dx \geq f\left(\int x p_X(x) dx\right)$$

Proof for $n = 3$

- for any x, y, z and $\alpha, \beta, \gamma > 0$ with $\alpha + \beta + \gamma = 1$

$$\begin{aligned}\alpha f(x) + \beta f(y) + \gamma f(z) &= (\alpha + \beta) \left(\frac{\alpha}{\alpha + \beta} f(x) + \frac{\beta}{\alpha + \beta} f(y) \right) + \gamma f(z) \\ &\geq (\alpha + \beta) f \left(\frac{\alpha}{\alpha + \beta} x + \frac{\beta}{\alpha + \beta} y \right) + \gamma f(z) \\ &\geq f \left((\alpha + \beta) \left(\frac{\alpha}{\alpha + \beta} x + \frac{\beta}{\alpha + \beta} y \right) + \gamma z \right) \\ &= f(\alpha x + \beta y + \gamma z)\end{aligned}$$

Proof for all n

- use mathematical induction
 - assume that Jensen's inequality holds for $1 \leq n \leq m$
 - for distinct x_i and $\alpha_i > 0$ ($1 \leq i \leq m+1$) with $\alpha_1 + \cdots + \alpha_{m+1} = 1$

$$\begin{aligned}
 \sum_{i=1}^{m+1} \alpha_i f(x_i) &= \left(\sum_{j=1}^m \alpha_j \right) \sum_{i=1}^m \left(\frac{\alpha_i}{\sum_{j=1}^m \alpha_j} f(x_i) \right) + \alpha_{m+1} f(x_{m+1}) \\
 &\geq \left(\sum_{j=1}^m \alpha_j \right) f \left(\sum_{i=1}^m \left(\frac{\alpha_i}{\sum_{j=1}^m \alpha_j} x_i \right) \right) + \alpha_{m+1} f(x_{m+1}) \\
 &= \left(\sum_{j=1}^m \alpha_j \right) f \left(\frac{1}{\sum_{j=1}^m \alpha_j} \sum_{i=1}^m \alpha_i x_i \right) + \alpha_{m+1} f(x_{m+1}) \\
 &\geq f \left(\sum_{i=1}^m \alpha_i x_i + \alpha_{m+1} x_{m+1} \right) = f \left(\sum_{i=1}^{m+1} \alpha_i x_i \right)
 \end{aligned}$$

1st and 2nd order conditions for convexity

- 1st order condition (assuming differentiable $f : \mathbf{R} \rightarrow \mathbf{R}$) - f is strictly convex *if and only if* for any $x \neq y$

$$f(y) > f(x) + f'(x)(y - x)$$

- 2nd order condition (assuming twice-differentiable $f : \mathbf{R} \rightarrow \mathbf{R}$)
 - if $f''(x) > 0$, f is strictly convex
 - f is convex *if and only if* for any x

$$f''(x) \geq 0$$

Jensen's inequality examples

- $f(x) = x^2$ is strictly convex

$$\frac{a^2 + b^2}{2} \geq \left(\frac{a + b}{2}\right)^2$$

- $f(x) = x^4$ is strictly convex

$$\frac{a^4 + b^4}{2} \geq \left(\frac{a + b}{2}\right)^4$$

- $f(x) = \exp(x)$ is strictly convex

$$\frac{\exp(a) + \exp(b)}{2} \geq \exp\left(\frac{a + b}{2}\right)$$

- equality holds *if and only if* $a = b$ for all inequalities

1st and 2nd order conditions for convexity - vector version

- 1st order condition (assuming differentiable $f : \mathbf{R}^n \rightarrow \mathbf{R}$) - f is strict convex *if and only if* for any x, y

$$f(y) > f(x) + \nabla f(x)^T (y - x)$$

where $\nabla f(x) \in \mathbf{R}^n$ with $\nabla f(x)_i = \partial f(x) / \partial x_i$

- 2nd order condition (assuming twice-differentiable $f : \mathbf{R}^n \rightarrow \mathbf{R}$)
 - if $\nabla^2 f(x) \succ 0$, f is strictly convex
 - f is convex *if and only if* for any x

$$\nabla^2 f(x) \succeq 0$$

where $\nabla^2 f(x) \in \mathbf{R}^{n \times n}$ is Hessian matrix of f evaluated at x , *i.e.*, $\nabla^2 f(x)_{i,j} = \partial^2 f(x) / \partial x_i \partial x_j$

Jensen's inequality examples - vector version

- assume $f : \mathbf{R}^n \rightarrow \mathbf{R}$
- $f(x) = \|x\|_2 = \sqrt{\sum x_i^2}$ is strictly convex

$$(\|a\|_2 + 2\|b\|_2)/3 \geq \|(a + 2b)/3\|_2$$

– equality holds *if and only if* $a = b \in \mathbf{R}^n$

- $f(x) = \|x\|_p = (\sum |x_i|^p)^{1/p}$ ($p > 1$) is strictly convex

$$\frac{1}{k} \left(\sum_{i=1}^k \|x^{(i)}\|_p \right) \geq \left\| \frac{1}{k} \sum_{i=1}^k x^{(i)} \right\|_p$$

– equality holds *if and only if* $x^{(1)} = \dots = x^{(k)} \in \mathbf{R}^n$

AM \geq GM

- for all $a, b > 0$

$$\frac{a+b}{2} \geq \sqrt{ab}$$

– equality holds if and only if $a = b$

- below most general form holds

Inequality 3. [AM-GM inequality] for any n $a_i > 0$ and $\alpha_i > 0$ with $\alpha_1 + \cdots + \alpha_n = 1$

$$\alpha_1 a_1 + \cdots + \alpha_n a_n \geq a_1^{\alpha_1} \cdots a_n^{\alpha_n}$$

where equality holds if and only if $a_1 = \cdots = a_n$

- let's prove these incrementally (for rational α_i)

Proof of $AM \geq GM$ - simplest case

- use fact that $x^2 \geq 0$ for any $x \in \mathbf{R}$

- for any $a, b > 0$

$$\begin{aligned} & (\sqrt{a} - \sqrt{b})^2 \geq 0 \\ \Leftrightarrow & a^2 - 2\sqrt{ab} + b^2 \geq 0 \\ \Leftrightarrow & a + b \geq 2\sqrt{ab} \\ \Leftrightarrow & \frac{a+b}{2} \geq \sqrt{ab} \end{aligned}$$

- equality holds if and only if $a = b$

Proof of $AM \geq GM$ - when $n = 4$ and $n = 8$

- for any $a, b, c, d > 0$

$$\frac{a + b + c + d}{4} \geq \frac{2\sqrt{ab} + 2\sqrt{cd}}{4} = \frac{\sqrt{ab} + \sqrt{cd}}{2} \geq \sqrt{\sqrt{ab}\sqrt{cd}} = \sqrt[4]{abcd}$$

- equality holds if and only if $a = b$ and $c = d$ and $ab = cd$ if and only if $a = b = c = d$

- likewise, for $a_1, \dots, a_8 > 0$

$$\begin{aligned} \frac{a_1 + \dots + a_8}{8} &\geq \frac{\sqrt{a_1 a_2} + \sqrt{a_3 a_4} + \sqrt{a_5 a_6} + \sqrt{a_7 a_8}}{4} \\ &\geq \sqrt[4]{\sqrt{a_1 a_2} \sqrt{a_3 a_4} \sqrt{a_5 a_6} \sqrt{a_7 a_8}} \\ &= \sqrt[8]{a_1 \dots a_8} \end{aligned}$$

- equality holds if and only if $a_1 = \dots = a_8$

Proof of $AM \geq GM$ - when $n = 2^m$

- generalized to cases $n = 2^m$

$$\left(\sum_{a=1}^{2^m} a_i \right) / 2^m \geq \left(\prod_{a=1}^{2^m} a_i \right)^{1/2^m}$$

– equality holds if and only if $a_1 = \cdots = a_{2^m}$

- can be proved by *mathematical induction*

Proof of $AM \geq GM$ - when $n = 3$

- proof for $n = 3$

$$\begin{aligned}\frac{a+b+c}{3} &= \frac{a+b+c+(a+b+c)/3}{4} \geq \sqrt[4]{abc(a+b+c)/3} \\ \Rightarrow \left(\frac{a+b+c}{3}\right)^4 &\geq abc(a+b+c)/3 \\ \Leftrightarrow \left(\frac{a+b+c}{3}\right)^3 &\geq abc \\ \Leftrightarrow \frac{a+b+c}{3} &\geq \sqrt[3]{abc}\end{aligned}$$

– equality holds if and only if $a = b = c = (a+b+c)/3$ if and only if $a = b = c$

Proof of $AM \geq GM$ - for all integers

- for any integer $n \neq 2^m$
- for m such that $2^m > n$

$$\begin{aligned}
 \frac{a_1 + \cdots + a_n}{n} &= \frac{a_1 + \cdots + a_n + (2^m - n)(a_1 + \cdots + a_n)/n}{2^m} \\
 &\geq \sqrt[2^m]{a_1 \cdots a_n \cdot ((a_1 + \cdots + a_n)/n)^{2^m - n}} \\
 \Leftrightarrow \left(\frac{a_1 + \cdots + a_n}{n} \right)^{2^m} &\geq a_1 \cdots a_n \cdot \left(\frac{a_1 + \cdots + a_n}{n} \right)^{2^m - n} \\
 \Leftrightarrow \left(\frac{a_1 + \cdots + a_n}{n} \right)^n &\geq a_1 \cdots a_n \\
 \Leftrightarrow \frac{a_1 + \cdots + a_n}{n} &\geq \sqrt[n]{a_1 \cdots a_n}
 \end{aligned}$$

– equality holds if and only if $a_1 = \cdots = a_n$

Proof of $AM \geq GM$ - rational α_i

- given n positive rational α_i , we can find n natural numbers q_i such that

$$\alpha_i = \frac{q_i}{N}$$

where $q_1 + \cdots + q_n = N$

- for any n positive $a_i \in \mathbf{R}$ and positive n $\alpha_i \in \mathbf{Q}$ with $\alpha_1 + \cdots + \alpha_n = 1$

$$\alpha_1 a_1 + \cdots + \alpha_n a_n = \frac{q_1 a_1 + \cdots + q_n a_n}{N} \geq \sqrt[N]{a_1^{q_1} \cdots a_n^{q_n}} = a_1^{\alpha_1} \cdots a_n^{\alpha_n}$$

– equality holds if and only if $a_1 = \cdots = a_n$

Proof of $AM \geq GM$ - real α_i

- exist n rational sequences $\{\beta_{i,1}, \beta_{i,2}, \dots\}$ ($1 \leq i \leq n$) such that

$$\beta_{1,j} + \dots + \beta_{n,j} = 1 \quad \forall j \geq 1$$

$$\lim_{j \rightarrow \infty} \beta_{i,j} = \alpha_i \quad \forall 1 \leq i \leq n$$

- for all j

$$\beta_{1,j}a_1 + \dots + \beta_{n,j}a_n \geq a_1^{\beta_{1,j}} \dots a_n^{\beta_{n,j}}$$

hence

$$\lim_{j \rightarrow \infty} (\beta_{1,j}a_1 + \dots + \beta_{n,j}a_n) \geq \lim_{j \rightarrow \infty} a_1^{\beta_{1,j}} \dots a_n^{\beta_{n,j}}$$

$$\Leftrightarrow \alpha_1a_1 + \dots + \alpha_na_n \geq a_1^{\alpha_1} \dots a_n^{\alpha_n}$$

- *cannot* prove equality condition from above proof method

Proof of $AM \geq GM$ using Jensen's inequality

- $(-\log)$ is strictly convex function because

$$\frac{d^2}{dx^2} (-\log(x)) = \frac{d}{dx} \left(-\frac{1}{x} \right) = \frac{1}{x^2} > 0$$

- Jensen's inequality implies for $a_i > 0$, $\alpha_i > 0$ with $\sum \alpha_i = 1$

$$-\log \left(\prod a_i^{\alpha_i} \right) = -\sum \log(a_i^{\alpha_i}) = \sum \alpha_i (-\log(a_i)) \geq -\log \left(\sum \alpha_i a_i \right)$$

- $(-\log)$ strictly monotonically decreases, hence $\prod a_i^{\alpha_i} \leq \sum \alpha_i a_i$, having just proved

$$\alpha_1 a_1 + \cdots + \alpha_n a_n \geq a_1^{\alpha_1} \cdots a_n^{\alpha_n}$$

where equality if and only if a_i are equal (by Jensen's inequality's equality condition)

Cauchy-Schwarz inequality

Inequality 4. [Cauchy-Schwarz inequality] for any $a_i, b_i \in \mathbf{R}$

$$(a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2) \geq (a_1b_1 + \cdots + a_nb_n)^2$$

- middle school proof

$$\begin{aligned} & \sum (ta_i + b_i)^2 \geq 0 \quad \forall t \in \mathbf{R} \\ \Leftrightarrow & t^2 \sum a_i^2 + 2t \sum a_i b_i + \sum b_i^2 \geq 0 \quad \forall t \in \mathbf{R} \\ \Leftrightarrow & \Delta = \left(\sum a_i b_i \right)^2 - \sum a_i^2 \sum b_i^2 \leq 0 \end{aligned}$$

- equality holds if and only if $\exists t \in \mathbf{R}, ta_i + b_i = 0$ for all $1 \leq i \leq n$

Cauchy-Schwarz inequality - another proof

- $x \geq 0$ for any $x \in \mathbf{R}$, hence

$$\begin{aligned}
 & \sum_i \sum_j (a_i b_j - a_j b_i)^2 \geq 0 \\
 \Leftrightarrow & \sum_i \sum_j (a_i^2 b_j^2 - 2a_i a_j b_i b_j + a_j^2 b_i^2) \geq 0 \\
 \Leftrightarrow & \sum_i \sum_j a_i^2 b_j^2 + \sum_i \sum_j a_j^2 b_i^2 - 2 \sum_i \sum_j a_i a_j b_i b_j \geq 0 \\
 \Leftrightarrow & 2 \sum_i a_i^2 \sum_j b_j^2 - 2 \sum_i a_i b_i \sum_j a_j b_j \geq 0 \\
 \Leftrightarrow & \sum_i a_i^2 \sum_j b_j^2 - \left(\sum_i a_i b_i \right)^2 \geq 0
 \end{aligned}$$

- equality holds if and only if $a_i b_j = a_j b_i$ for all $1 \leq i, j \leq n$

Cauchy-Schwarz inequality - still another proof

- for any $x, y \in \mathbf{R}$ and $\alpha, \beta > 0$ with $\alpha + \beta = 1$

$$\begin{aligned}
 (\alpha x - \beta y)^2 &= \alpha^2 x^2 + \beta^2 y^2 - 2\alpha\beta xy \\
 &= \alpha(1 - \beta)x^2 + (1 - \alpha)\beta y^2 - 2\alpha\beta xy \geq 0 \\
 \Leftrightarrow \alpha x^2 + \beta y^2 &\geq \alpha\beta x^2 + \alpha\beta y^2 + 2\alpha\beta xy = \alpha\beta(x + y)^2 \\
 \Leftrightarrow x^2/\alpha + y^2/\beta &\geq (x + y)^2
 \end{aligned}$$

- plug in $x = a_i, y = b_i, \alpha = A/(A + B), \beta = B/(A + B)$ where $A = \sqrt{\sum a_i^2}, B = \sqrt{\sum b_i^2}$

$$\begin{aligned}
 \sum (a_i^2/\alpha + b_i^2/\beta) &\geq \sum (a_i + b_i)^2 \Leftrightarrow (A + B)^2 \geq A^2 + B^2 + 2 \sum a_i b_i \\
 \Leftrightarrow AB &\geq \sum a_i b_i \Leftrightarrow A^2 B^2 \geq \left(\sum a_i b_i \right)^2 \Leftrightarrow \sum a_i^2 \sum b_i^2 \geq \left(\sum a_i b_i \right)^2
 \end{aligned}$$

Cauchy-Schwarz inequality - proof using determinant

- almost the same proof as first one - but using 2-by-2 matrix determinant

$$\begin{aligned}
 & \sum (xa_i + yb_i)^2 \geq 0 \quad \forall x, y \in \mathbf{R} \\
 \Leftrightarrow & x^2 \sum a_i^2 + 2xy \sum a_i b_i + y^2 \sum b_i^2 \geq 0 \quad \forall x, y \in \mathbf{R} \\
 \Leftrightarrow & \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \sum a_i^2 & \sum a_i b_i \\ \sum a_i b_i & \sum b_i^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \geq 0 \quad \forall x, y \in \mathbf{R} \\
 \Leftrightarrow & \left| \begin{array}{cc} \sum a_i^2 & \sum a_i b_i \\ \sum a_i b_i & \sum b_i^2 \end{array} \right| \geq 0 \Leftrightarrow \sum a_i^2 \sum b_i^2 - \left(\sum a_i b_i \right)^2 \geq 0
 \end{aligned}$$

– equality holds *if and only if*

$$(\exists x, y \in \mathbf{R} \text{ with } xy \neq 0) (xa_i + yb_i = 0 \quad \forall 1 \leq i \leq n)$$

- allows *beautiful generalization* of Cauchy-Schwarz inequality

Cauchy-Schwarz inequality - generalization

- want to say something like $\sum_{i=1}^n (xa_i + yb_i + zc_i + wd_i + \dots)^2$
- run out of alphabets . . . - use double subscripts

$$\sum_{i=1}^n (x_1 A_{1,i} + x_2 A_{2,i} + \dots + x_m A_{m,i})^2 \geq 0 \quad \forall x_i \in \mathbf{R}$$

$$\Leftrightarrow \sum_{i=1}^n (x^T a_i)^2 = \sum_{i=1}^n x^T a_i a_i^T x = x^T \left(\sum_{i=1}^n a_i a_i^T \right) x \geq 0 \quad \forall x \in \mathbf{R}^m$$

$$\Leftrightarrow \begin{vmatrix} \sum_{i=1}^n A_{1,i}^2 & \sum_{i=1}^n A_{1,i} A_{2,i} & \cdots & \sum_{i=1}^n A_{1,i} A_{m,i} \\ \sum_{i=1}^n A_{1,i} A_{2,i} & \sum_{i=1}^n A_{2,i}^2 & \cdots & \sum_{i=1}^n A_{2,i} A_{m,i} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n A_{1,i} A_{m,i} & \sum_{i=1}^n A_{2,i} A_{m,i} & \cdots & \sum_{i=1}^n A_{m,i}^2 \end{vmatrix} \geq 0$$

where $a_i = \begin{bmatrix} A_{1,i} & \cdots & A_{m,i} \end{bmatrix}^T \in \mathbf{R}^m$

– equality holds *if and only if* $\exists x \neq 0 \in \mathbf{R}^m$, $x^T a_i = 0$ for all $1 \leq i \leq n$

Cauchy-Schwarz inequality - three series of variables

- let $m = 3$

$$\begin{aligned} & \begin{bmatrix} \sum a_i^2 & \sum a_i b_i & \sum a_i c_i \\ \sum a_i b_i & \sum b_i^2 & \sum b_i c_i \\ \sum a_i c_i & \sum b_i c_i & \sum c_i^2 \end{bmatrix} \succeq 0 \\ \Rightarrow & \sum a_i^2 \sum b_i^2 \sum c_i^2 + 2 \sum a_i b_i \sum b_i c_i \sum c_i a_i \\ & \geq \sum a_i^2 \left(\sum b_i c_i \right)^2 + \sum b_i^2 \left(\sum a_i c_i \right)^2 + \sum c_i^2 \left(\sum a_i b_i \right)^2 \end{aligned}$$

– equality holds if and only if $\exists x, y, z \in \mathbf{R}$, $xa_i + yb_i + zc_i = 0$ for all $1 \leq i \leq n$

- questions for you
 - what does this mean?
 - any real-world applications?

Cauchy-Schwarz inequality - extensions

Inequality 5. [Cauchy-Schwarz inequality - for complex numbers] *for $a_i, b_i \in \mathbb{C}$*

$$\sum |a_i|^2 \sum |b_i|^2 \geq \left| \sum a_i b_i \right|^2$$

Inequality 6. [Cauchy-Schwarz inequality - for infinite sequences] *for two complex infinite sequences $\langle a_i \rangle_{i=1}^{\infty}$ and $\langle b_i \rangle_{i=1}^{\infty}$*

$$\sum_{i=1}^{\infty} |a_i|^2 \sum_{i=1}^{\infty} |b_i|^2 \geq \left| \sum_{i=1}^{\infty} a_i b_i \right|^2$$

Inequality 7. [Cauchy-Schwarz inequality - for complex functions] *for two complex functions $f, g : [0, 1] \rightarrow \mathbb{C}$*

$$\int |f|^2 \int |g|^2 \geq \left| \int f g \right|^2$$

- note that *all these can be further generalized as in page 31*

Number Theory - Queen of Mathematics

Integers

- integers (**Z**) - $\dots - 2, -1, 0, 1, 2, \dots$
 - first defined by Bertrand Russell
 - algebraic structure - commutative ring
 - addition, multiplication defined, but division *not* defined
 - addition, multiplication are associative
 - multiplication distributive over addition
 - addition, multiplication are commutative
- natural numbers (**N**)
 - $1, 2, \dots$

Division and prime numbers

- divisors for $n \in \mathbf{N}$

$$\{d \in \mathbf{N} \mid d \text{ divides } n\}$$

- prime numbers
 - p is primes if 1 and p are only divisors

Fundamental theorem of arithmetic

Theorem 1. [fundamental theorem of arithmetic] *integer $n \geq 2$ can be factored uniquely into products of primes, i.e., exist distinct primes, p_1, \dots, p_k , and $e_1, \dots, e_k \in \mathbf{N}$ such that*

$$n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$$

- hence, integers are *factorial ring* (Definition ??)

Elementary quantities

- greatest common divisor (gcd) (of a and b)

$$\gcd(a, b) = \max\{d \mid d \text{ divides both } a \text{ and } b\}$$

– for definition of gcd for general entire rings, refer to Definition ??

- least common multiple (lcm) (of a and b)

$$\text{lcm}(a, b) = \min\{m \mid \text{both } a \text{ and } b \text{ divides } m\}$$

- a and b coprime, relatively prime, mutually prime $\Leftrightarrow \gcd(a, b) = 1$

Are there infinite number of prime numbers?

- yes!
- proof
 - assume there only exist finite number of prime numbers, *e.g.*, $p_1 < p_2 < \cdots < p_n$
 - but then, $p_1 \cdot p_2 \cdots p_n + 1$ is prime, but which is greater than p_n , hence contradiction

Integers modulo n

Definition 3. [modulo] *when n divides $a - b$, a , said to be equivalent to b modulo n , denoted by*

$$a \equiv b \pmod{n}$$

read as “ a congruent to b mod n ”

- $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$ imply
 - $a + c \equiv b + d \pmod{n}$
 - $ac \equiv bd \pmod{n}$

Definition 4. [congruence class] *classes determined by modulo relation, called congruence or residue class under modulo*

Definition 5. [integers modulo n] *set of equivalence classes under modulo, denoted by $\mathbb{Z}/n\mathbb{Z}$, called integers modulo n or integers mod n*

Euler's theorem

Definition 6. [Euler's totient function] for $n \in \mathbf{N}$,

$$\varphi(n) = (p_1 - 1)p_1^{e_1-1} \cdots (p_k - 1)p_k^{e_k-1} = n \prod_{\text{prime } p \text{ dividing } n} (1 - 1/p)$$

called Euler's totient function, also called Euler φ -function

- e.g., $\varphi(12) = \varphi(2^2 \cdot 3^1) = 1 \cdot 2^1 \cdot 2 \cdot 3^0 = 4$, $\varphi(10) = \varphi(2^1 \cdot 5^1) = 1 \cdot 2^0 \cdot 4 \cdot 5^0 = 4$

Theorem 2. [Euler's theorem - number theory] for coprime n and a

$$a^{\varphi(n)} \equiv 1 \pmod{n}$$

- e.g., $5^4 \equiv 1 \pmod{12}$ whereas $4^4 \equiv 4 \not\equiv 1 \pmod{12}$
- Euler's theorem underlies RSA cryptosystem, which is pervasively used in internet communication

References

References

- [HLP52] G. Hardy, J.E. Littlewood, and G. Polya. *Inequalities*. Cambridge Mathematical Library, 2nd edition, 1952.

Index

a.e.

almost everywhere, [6](#)

a.s.

almost surely, [6](#)

almost everywhere, [6](#)

almost everywhere - a.e., [6](#)

almost surely, [6](#)

AM-GM inequality, [18](#)

boundary

set, [3](#)

Cauchy, Augustin-Louis

Cauchy-Schwarz inequality, [27](#)

extension, [33](#)

for complex functions, [33](#)

for complex numbers, [33](#)

for infinite sequences, [33](#)

generalization, [31](#)

Cauchy-Schwarz inequality, [27](#)

extension, [33](#)

for complex functions, [33](#)

for complex numbers, [33](#)

for infinite sequences, [33](#)

generalization, [31](#)

Cauchy-Schwarz inequality - for complex functions,
[33](#)

Cauchy-Schwarz inequality - for complex numbers,
[33](#)

Cauchy-Schwarz inequality - for infinite sequences,
[33](#)

closure

set, [3](#)

complement

set, [2](#)

complex number, [2](#)

congruence class, [40](#)

integers, [40](#)

convex functions, [10](#)

first order condition, [14](#)

vector functions, [16](#)

second order condition, [14](#)

vector functions, [16](#)

strictly, [10](#)

definitions

almost everywhere - a.e., [6](#)

congruence class, [40](#)

Euler's totient function, [41](#)

infinitely often - i.o., [6](#)

integers modulo n , [40](#)

modulo, [40](#)

difference

set, [3](#)

Euler φ -function, [41](#)

Euler phi-function, [41](#)

Euler's theorem, [41](#)

Euler's theorem - number theory, [41](#)

Euler's totient function, [41](#)

Euler, Leonhard

φ -function, [41](#)

Euler's theorem, [41](#)

Euler's totient function, [41](#)

phi-function, [41](#)

finite sequence, [2](#)

fundamental theorem

of arithmetic, [37](#)

fundamental theorem of arithmetic, [37](#)

greatest common divisor, [38](#)

integers, [38](#)

i.o.

infinitely often, [6](#)

inequalities

AM-GM inequality, [18](#)

Cauchy-Schwarz inequality, [27](#)

Cauchy-Schwarz inequality - for complex functions, [33](#)

Cauchy-Schwarz inequality - for complex numbers, [33](#)

Cauchy-Schwarz inequality - for infinite sequences, [33](#)

Jensen's inequality, [10](#)

Jensen's inequality - for finite sequences, [10](#)

Jensen's inequality - for random variables, [11](#)

infinite sequence, [2](#)

infinitely often, [6](#)

infinitely often - i.o., [6](#)

integer, [2](#)

integers

congruence class, [40](#)

integers mod n , [40](#)

integers modulo n , [40](#)

residue class under modulo, [40](#)

integers mod n , [40](#)

integers modulo n , [40](#)

integers modulo n , [40](#)

interior

set, [3](#)

Jensen's inequality, [10](#)

for finite sequences, [10](#)

for random variables, [11](#)

Jensen's inequality - for finite sequences, [10](#)

Jensen's inequality - for random variables, [11](#)

Jensen, Johan Ludwig William Valdemar

inequality, [10](#)

Jensen's inequality

for finite sequences, [10](#)

for random variables, [11](#)

least common multiple, [38](#)

integers, [38](#)

matrix

positive definite, [4](#)

positive semi-definite, [4](#)

symmetric, [4](#)

trace, [3](#)

modulo, [40](#)

natural number, [2](#)

norm

vector, [3](#)

number

complex number, [2](#)

integer, [2](#)

natural number, [2](#)

rational number, [2](#)

real number, [2](#)

positive definite matrix, [4](#)

positive semi-definite matrix, [4](#)

rational number, [2](#)

real number, [2](#)

relative interior

set, [3](#)

residue class under modulo, [40](#)

integers, [40](#)

Schwarz, Hermann

Cauchy-Schwarz inequality, [27](#)

extension, [33](#)

for complex functions, [33](#)

for complex numbers, [33](#)

for infinite sequences, [33](#)

generalization, [31](#)

sequence, [2](#)

finite sequence, [2](#)

infinite sequence, [2](#)

set

boundary, [3](#)

closure, [3](#)

complement, [2](#)

difference, [3](#)

Sunghee Yun

July 14, 2025

interior, [3](#)

relative interior, [3](#)

smallest σ -algebra containing subsets, [3](#)

symmetric matrix, [4](#)

theorems

Euler's theorem - number theory, [41](#)

fundamental theorem of arithmetic, [37](#)

trace

matrix, [3](#)

vector

norm, [3](#)

ZZ-todo

0 - apply new comma conventions, [0](#)

1 - convert bullet points to proper theorem, definition, lemma, corollary, proposition, etc., [0](#)

CANCELED - < 2024 0421 - python script extracting important list, [0](#)

DONE - 2024 0324 - change tocpageref and funpageref to hyperlink, [0](#)

DONE - 2024 0324 - python script extracting figure list → using “list of figures” functionality on doc, [0](#)

DONE - 2024 0324 - python script extracting theorem-like list → using “list of theorem” functionality on doc, [0](#)

DONE - 2024 0324 - python script for converting slides to doc, [0](#)

DONE - 2025 0414 - 1 - change mathematicians' names, [0](#)