Convex Functions on $\mathbb{R}$ and Inequalities

S. Kumaresan
School of Math. and Stat.
University of Hyderabad
Hyderabad 500046
kumaresa@gmail.com

The aim of this article is to introduce the reader to some basic results on convex functions and their applications to inequalities, the bread and butter of analysis. Unfortunately, convex functions are not taught nowadays in standard curriculum. It is a well-known fact that the study of convex functions unifies the treatment of almost all classical inequalities. Convex functions also arise when we deal with the existence of points of global extrema of real valued functions. Knowledge of convex functions is useful in sketching the graphs of functions from $\mathbb{R}$ to $\mathbb{R}$.

1 Basic Results

Let I, J be intervals in $\mathbb{R}$. A function $f : I \to \mathbb{R}$ is said to be *convex* if for all $t \in [0, 1]$ and for all $x, y \in I$, the following inequality holds:

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y). \tag{1}$$

Recall that the line joining $(y, f(y))$ and $(x, f(x))$ in $\mathbb{R}^2$ is given by

$$t(x, f(x)) + (1-t)(y, f(y)) = (tx + (1-t)y, tf(x) + (1-t)f(y)), \qquad 0 \le t \le 1.$$

Then the y-coordinates of points on this line are given by $tf(x) + (1-t)f(y)$. Thus the geometric meaning of the definition is that the line segment joining $(x, f(x))$ and $(y, f(y))$ is never below the graph of the function. See Fig. 1.

If for $0 < t < 1$, strict inequality holds in (1), then the function is said to be *strictly convex.*

The expression $tx+(1-t)y$ is called a convex combination of x and y. More generally, if x_j, $1 \le j \le n$ are points of the interval and $t_j \in [0, 1]$ such that $\sum_{j=1}^n t_j = 1$, then $t_1x_1 + \cdots + t_nx_n$ is called a convex linear combination of x_j's. A convex combination is a weighted average. The convexity inequality (1) can be interpreted as saying that every weighted average of two function values is greater than the function value at the corresponding weighted average of the arguments.

Ex. 1. Show that the functions $f(x) = ax + b$, $g(x) = x^2$ and $h(x) = |x|$ are convex. (Later, we shall see an easy method of proof for the first two functions.)

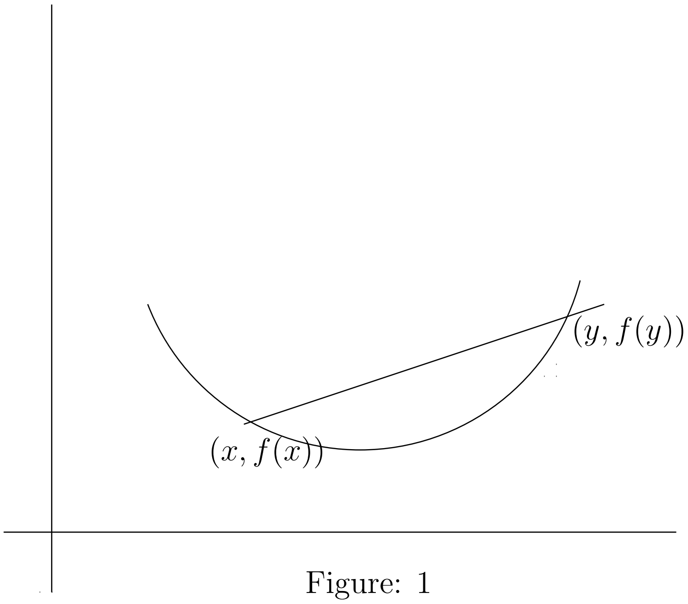


Figure: 1

Ex. 2. A function $f: I \to \mathbb{R}$ is convex iff $f: J \to \mathbb{R}$ is convex for every subinterval $J \subset I$.

Theorem 3 (Jensen's Theorem)**.** *Let $f: J \to \mathbb{R}$ be convex. Let $z = \sum_{j=1}^n t_j x_j$ be a convex combination of points of J. Then*

$$f(t_1x_1 + \cdots + t_nx_n) \leq t_1 f(x_1) + \cdots + t_n f(x_n), \qquad \text{for } t_j \in [0,1] \text{ with } \sum_j t_j = 1. \tag{2}$$

Proof. We prove the result by induction. When $n = 2$, it is the definition of convexity. Assume the result for all convex linear combinations with $n-1$ terms. Let $z = \sum_j t_j x_j$ be given. Let $s_j = \frac{t_j}{1-t_1}$. Then $z = t_1x_1 + (1-t_1)(\sum_{j=2}^n s_j x_j)$. Using the inequality for $n = 2$ and again for $n-1$, we get

$$\begin{aligned} f(t_1x_1 + \cdots + t_nx_n) &\leq t_1 f(x_1) + (1-t_1) f(\sum_{j=2}^n s_j x_j) \\ &\leq t_1 f(x_1) + (1-t_1) \sum_{j=2}^n s_j f(x_j) \\ &= t_1 f(x_1) + t_2 f(x_2) + \cdots + t_n f(x_n). \end{aligned}$$

□

Let $m_{x,y} := \frac{f(y)-f(x)}{y-x}$ for any $x, y \in I$. Note that this is the slope of the line joining $(x, f(x))$ and $(y, f(y))$.

Lemma 4. *Let $f: I \to \mathbb{R}$ be given. Let $x, y \in I$. Fix $t \in [0,1]$ and $z := (1-t)x + ty$. Let $m_{u,v} := \frac{f(v)-f(u)}{v-u}$ for any $u, v \in I$. Then the following are equivalent:*
(a) $f(z) \leq (1-t)f(x) + tf(y)$.
(b) $m_{x,z} \leq m_{x,y}$.
(c) $m_{x,y} \leq m_{z,y}$.

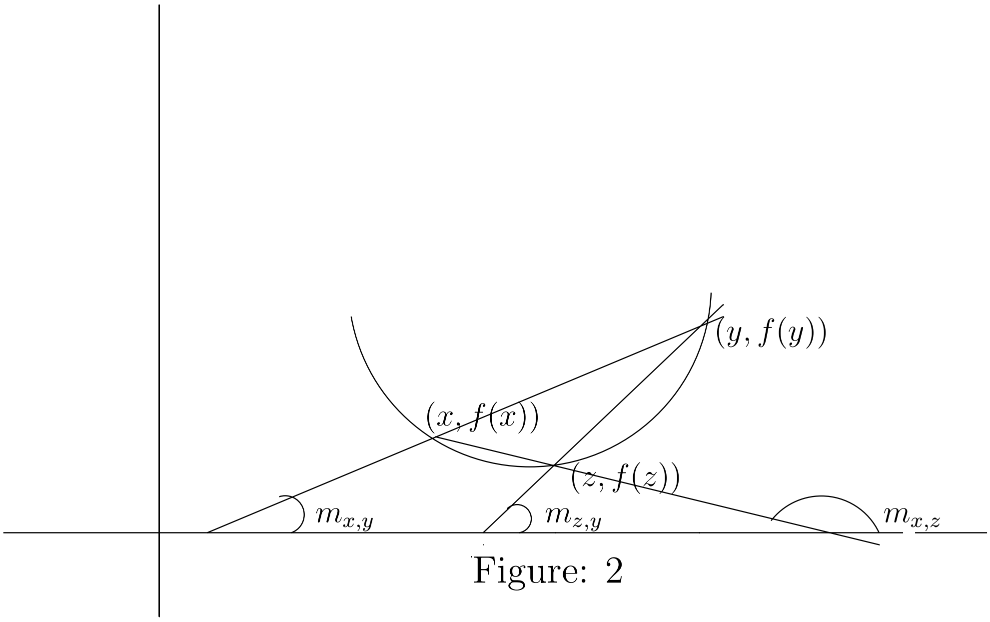


Figure: 2

Proof. See Fig. 2.

(a) $\iff$ (b): Note that (a) can rewritten as $f(z) - f(x) \le t[f(y) - f(x)]$. Since $t = (z-x)/(y-x)$, this can be cast as

$$f(z) - f(x) \le \frac{z-x}{y-x}[f(y) - f(x)]. \tag{3}$$

Since the (b) is obviously true if t is either 0 or 1, we may assume $0 < t < 1$ so that $z - x > 0$. We can divide both sides of the inequality (3) by $z - x$ to obtain $m_{x,z} \le m_{x,y}$. Clearly, these steps are reversible.

(a) $\iff$ (c): We write (a) as $f(z) - f(y) \le (1-t)[f(x) - f(y)]$. Proceeding as above we get the result. ☐

Theorem 5. *A differentiable function f on an interval I is convex iff f' is non-decreasing on I.*

Proof. Assume first that f is differentiable and convex on I. Let $x, y \in I$ with $x < y$. By the above lemma, for any $z \in [x, y]$, we have

$$m_{x,z} \le m_{x,y} \le m_{z,y}.$$

Letting z decrease to x in the first inequality yields $f'(x) \le m_{x,y}$. Letting z increase to y in the second inequality yields $m_{x,y} \le f'(y)$. Hence $f'(x) \le f'(y)$. That is, f' is non-decreasing.

Let us assume that f is differentiable and f' is non-decreasing. If f' were not convex, then by the lemma, there exist points $x < z < y$ in I such that $m_{x,z} > m_{x,y}$ and $m_{x,y} > m_{z,y}$. (Is this clear why such points exist? Think carefully.) By the mean value theorem, there exist points u and v with $x < u < z$ and $z < v < y$ such that $f'(u) = m_{x,z}$ and $f'(v) = m_{z,y}$.

But then $f'(u) > f'(v)$ and hence f' is not non-decreasing. This contradicts our hypothesis. Hence we conclude that f is convex. □

Corollary 6. *Let $f\colon I \to \mathbb{R}$ be twice differentiable. Then f is convex iff f'' is nonnegative on I.*

If $f'' > 0$, then f is strictly convex.

Proof. If $f'' > 0$ and if it were not strictly convex, then there exist points $x, y \in I$ and a $t \in (0,1)$ such that if $z := (1-t)x + ty$ then $f(z) = (1-t)f(x) + tf(y)$. That is, the points $(x, f(x))$, $(y, f(y))$ and $(z, f(z))$ are collinear. By Rolle's theorem, there exist points u and v with $u < v$ and such that $f'(u) = f'(v)$. But then there exists $w \in (u, v)$ such that $f''(w) = 0$. □

Since Corollary 6 is the most useful one, we give a more direct proof of it in the following

Theorem 7. *Let J be an open interval in $\mathbb{R}$. Assume that $f\colon J \to \mathbb{R}$ is twice differentiable. Then f is convex iff $f''(x) \geq$ for all $x \in J$.*

Proof. We prove the necessity part. Let $a \in J$. Let h is such that $a - h, a + h \in J$. Then $a = [(a+h) + (a-h)]/2$. Since f is convex on J, we have

$$f(a) = f(\frac{a+h}{2} + \frac{a-h}{2}) \leq \frac{1}{2}f(a+h) + \frac{1}{2}f(a-h).$$

It follows that

$$f(a+h) - 2f(a) + f(a-h) \geq 0.$$

Since $h^2 \geq 0$, it follows from (6) in Ex. 8 below that $f''(a) \geq 0$.

We now prove the sufficiency part using Taylor's theorem. Let $x, y \in J$ and $0 < t < 1$. Let $z = (1-t)x + ty$. We apply Taylor's theorem to f at z to find a point c_1 between x and z such that

$$f(x) = f(z) + f'(z)(x-z) + \frac{1}{2}f''(c_1)(x-z)^2. \tag{4}$$

Similarly there exists c_2 between y and z such that

$$f(y) = f(z) + f'(z)(y-z) + \frac{1}{2}f''(c_1)(y-z)^2. \tag{5}$$

Using (4) and (5), we obtain

$$\begin{aligned}
(1-t)f(x) + tf(y) &= f(z) + f'(z)[(1-t)x + ty - z] \\
&\quad + \frac{1}{2}(1-t)f''(c_1)(x-z)^2 \\
&\quad + \frac{1}{2}tf''(c_2)(y-z)^2 \\
&= f(z) + \frac{1}{2}(1-t)f''(c_1)(x-z)^2 + \frac{1}{2}tf''(c_2)(y-z)^2 \\
&\geq f(z) = f((1-t)x + ty).
\end{aligned}$$

Hence f is convex. □

Ex. 8. Let $f\colon J \to \mathbb{R}$ be such that $f''(a)$ exists. Then

$$f''(a) := \lim_{h\to 0} \frac{f(x+h) - 2f(x) + f(x-h)}{h^2}. \tag{6}$$

Give an example to show that the converse if not true.

Corollary 6 (that is, Theorem 7) gives a lot of examples of convex functions.

Ex. 9. Prove the following:
(i) e^x is (strictly) convex on $\mathbb{R}$.
(ii) x^α is convex on $(0, \infty)$ for $\alpha \geq 1$.
(iii) $-x^\alpha$ is strictly convex on $(0, \infty)$ for $0 < \alpha < 1$.
(iv)) $x \log x$ is strictly convex on $(0, ?)$.
(v) $f(x) = x^4$ is strictly convex but $f''(0) = 0$.
(vi) $f(x) = x + (1/x)$ is convex on $(0, \infty)$.
(vii) $f(x) = 1/x$ is convex on $(0, \infty)$.

Ex. 10. Let f, g be convex functions and $\alpha \in \mathbb{R}$. Which of the functions $f + g$, αf and $f \cdot g$ are convex?

Ex. 11. Let $f\colon [a, b] \to \mathbb{R}$ be convex. Prove that the maximum of f on $[a, b]$ is either $f(a)$ or $f(b)$.

Ex. 12. Let f be twice differentiable with $f'' \geq 0$ on an interval $[a, b]$. What is the maximum possible value of $f((a + b)/2)$? For what functions, this bound is attained?

Ex. 13. Which polynomials of odd degree are convex on $\mathbb{R}$?

Ex. 14. Characterize the fourth degree polynomials that are convex on $\mathbb{R}$ by giving a necessary and sufficient condition on the coefficients.

Ex. 15. Let $f\colon [a, b] \to \mathbb{R}$ be continuous and increasing. Then $g(x) := \int_a^x f(t)\,dt$ is convex on $[a, b]$.

We now make a geometric observation which paves the way for defining convex functions on more general spaces, say, for example, on Riemannian manifolds.

Proposition 16. *Let f be convex and differentiable on (a, b). If $y \in (a, b)$, then $f(x) - f(y) \geq f'(y)(x - y)$, for all $x \in (a, b)$.*

Remark 17. The geometric interpretation of this is that the graph of the function lies above any tangent drawn to the graph. See Fig. 3. This suggests the following definition.

Let $\varphi\colon (a, b) \to \mathbb{R}$ be convex and $\alpha \in (a, b)$. A *supporting line* is a line $y = M(x - \alpha) + \varphi(\alpha)$ passing through the point $(\alpha, \varphi(\alpha))$ and such that it always lies below the graph of φ, i.e., is such that

$$\varphi(x) \geq M(x - \alpha) + \varphi(\alpha), \qquad \text{for all } x \in (a, b).$$

Hence, Proposition 16 tells us that if φ is differentiable, then the tangent line to the graph at $(\alpha, \varphi(\alpha))$ is a supporting line at that point.

In fact, existence of supporting lines at each point of the graph is the characterizing property of convex functions.

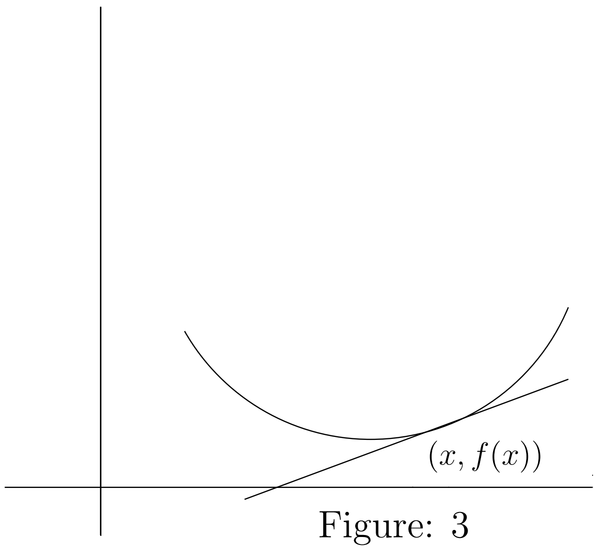


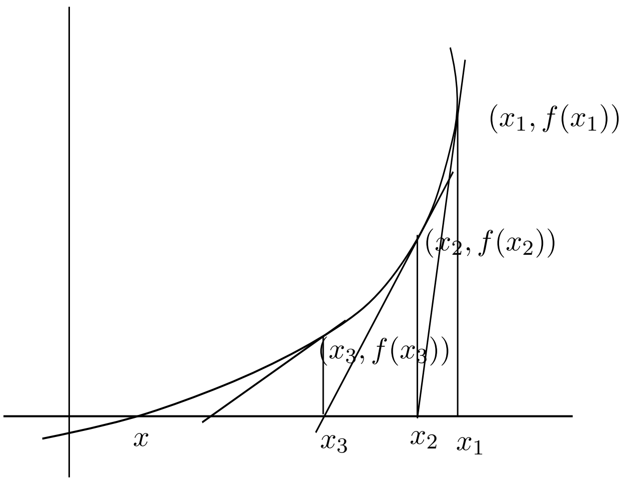
Figure: 3

Proof. By mean value theorem, there exists z between x and y such that $\frac{f(x)-f(y)}{x-y} = f'(z)$. Since f is convex, f' increases. Hence $f'(z) \geq f'(y)$ if $x > y$ or $f'(z) \leq f'(y)$ if $x < y$. ☐

Ex. 18. Use Prop. 16 to justify the **Newton-Raphson method:** Let $\phi \colon \mathbb{R} \to \mathbb{R}$ be a strictly increasing, convex, differentiable function on $\mathbb{R}$. Assume that $\phi(x) = 0$. If $x_1 > x$, and x_n is defined inductively as

$$x_n = x_{n-1} - \frac{\phi(x_{n-1})}{\phi'(x_{n-1})},$$

then show that $x_n \to x$ as $n \to \infty$. See Fig. 4.



Ex. 19. Assume f is differentiable and convex on $\mathbb{R}$. If f is bounded then f is a constant.

Lemma 20. *Let $\psi \colon (a, b) \to \mathbb{R}$ be convex. Let $t \in (a, b)$. Then there exists a supporting line to the graph of ψ at the point $(t, \psi(t))$.*

Remark 21. In fact, Lemma 23 tells us the right and left hand derivatives exist at each point of the domain of a convex function. In light of it, we see that any line passing through the point of the graph and whose slope lies between the one-sided derivatives at the point is a supporting line at that point. The following proof is a direct attack on the problem.

Proof. Let

$$M := \sup_{s\in(a,t)} \left\{ \frac{\psi(t)-\psi(s)}{t-s} \right\}. \tag{7}$$

Let $u \in (t, b)$. From Lemma 4, we have

$$M \leq \frac{\psi(u)-\psi(t)}{u-t}. \tag{8}$$

Then (8) implies

$$\psi(u) \geq M(u-t) + \psi(t), \qquad t \leq u \leq b, \tag{9}$$

and (7) implies

$$\psi(s) \geq M(s-t) + \psi(t), \qquad \leq s \leq t. \tag{10}$$

Then (9) and (10) imply that

$$\psi(x) \geq \psi(t) + M(x-t) \text{ for any } x \in (a,b) \tag{11}$$

□

Proposition 22. *If $f : (a,b) \to \mathbb{R}$ is convex and $c \in (a,b)$ is a local minimum, then c is a minimum for f on (a,b). That is, local minima of convex functions are global minima.*

Proof. Recall that c is a local minimum for f if and only if there exists $\varepsilon > 0$ such that $f(x) \geq f(c)$ for all $x \in (c-\varepsilon, c+\varepsilon) \subseteq (a,b)$.

If c were not a point of global minimum, there exists a $d \in (a,b)$ ($d \leq c$ or $d \geq c$ with $f(d) < f(c)$. Consider the curve $c(t) = (1-t)c + td$. Then $c(0) = c$, $c(1) = d$ and

$$\begin{aligned} f(c(t)) &\leq (1-t)f(c) + tf(d) \\ &< (1-t)f(c) + tf(c), \qquad t \neq 0 \\ &= f(c) \qquad \text{for all } t \in (0, 1\ ] \end{aligned} \tag{12}$$

But for t sufficiently small, $c(t) \in (c-\varepsilon, c+\varepsilon)$ so that

$$f(c(t)) \geq f(c) \quad \text{for } 0 < t < \varepsilon.$$

which contradicts (12). □

Lemma 23. *Let f be convex on an open interval I. Then the limits*

$$\lim_{h\to 0^-} \frac{f(x+h)-f(x)}{h} \text{ and } \lim_{h\to 0^+} \frac{f(x+h)-f(x)}{h}$$

both exist for every $x \in I$. Thus f has both right and left side derivatives on I.

Proof. Let $h_1, h_2 \in \mathbb{R}$ be such that $0 < h_1 < h_2$. Using Lemma 4, we get

$$\frac{f(x+h_1)-f(x)}{h_1} \le \frac{f(x+h_2)-f(x)}{h_2}$$

Thus if we define the function

$$F(h) := \frac{f(x+h)-f(x)}{h}$$

then $F(h_1) \le F(h_2)$ if $0 < h_1 < h_2$, i.e., F *increases* on $(0, \delta)$ for all δ sufficiently small. Hence $\lim_{h\to 0^+} F(h)$ exists. (Why?) □

Example 24. What is the relevance of the function $f(x) := |x|$ to the above lemma?

Theorem 25. *Let f be convex on the open interval I. Then f is continuous on I.*

Proof. By the above lemma, we have

$$\begin{aligned} \lim_{h\to 0^-} [f(x+h) - f(x)] &= \lim_{h\to 0^-} \frac{f(x+h)-f(x)}{h} \{\lim_{h\to 0^-} h\} \\ &= 0. \end{aligned}$$

Similarly $\lim_{h\to 0^+}[f(x+h) - f(x)] = 0$. □

Ex. 26. What can you say when the interval I in the theorem is not open?

Proposition 27. *A continuous function $f : I \to \mathbb{R}$ is convex iff it is midpoint convex, i.e. iff*

$$f(\frac{x+y}{2}) \le \frac{1}{2}[f(x) + f(y)] \text{ for all } x, y \in I.$$

Proof. By the inequality 2 we have

$$f(\frac{x_1 + \cdots + x_n}{n}) \le \frac{1}{n}[f(x_1) + \cdots + f(x_n)], \quad x_i \in I. \tag{13}$$

Let r be rational, $0 < r < 1$. Write $r = \frac{m}{n}$. If $s = 1 - r$, then $s = \frac{n-m}{n}$. We apply (13) with $x_1 = x_2 = \cdots = x_m = x$, $x_{m+1} = \cdots = x_n = y$. Then

$$\begin{aligned} f(rx+sy) &= f(\frac{m}{n}x + \frac{n-m}{n}y) \\ &= f(\frac{mx + (n-m)y}{n}) \\ &= f(\frac{\overbrace{x + \cdots + x}^{m \text{ times}} + \overbrace{y + \cdots + y}^{n-m \text{ times}}}{n}) \\ &\le \frac{1}{n}[\overbrace{f(x) + \cdots + f(x)}^{m \text{ times}} + \overbrace{f(y) + \cdots + f(y)}^{n-m \text{ times}}] \\ &= \frac{1}{n}[mf(x) + (n-m)f(y)] \\ &= rf(x) + sf(y). \end{aligned} \tag{14}$$

If α is irrational, consider a sequence $\{\alpha_n\}$ of rational numbers such that $\alpha_n \to \alpha$. Then $\beta_n = 1 - \alpha_n \to 1 - \alpha = \beta$ as $n \to \infty$. We have

$$f(\alpha_n x + \beta_n y) \leq \alpha_n f(x) + \beta_n f(y), \quad n \in \mathbb{N}. \tag{15}$$

Since f is continuous, taking limits on both sides of (15), we see that

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y).$$

That is, f is convex. □

Project. A function $f\colon I \to \mathbb{R}$ is said to be *concave* if the inequality (1) is reversed. Find examples of concave functions. Formulate results (analogous to those for convex functions) for concave functions. Prove them.

2 Applications: Inequalities

The most important use of the theory of convex functions is the derivation of inequalities. We start with a most important inequality, second only to the triangle inequality.

Ex. 28 (Arithmetic Mean and Geometric Mean Inequality)**.** Let $\alpha_j \geq 0$ for $1 \leq j \leq n$. Then the arithmetic mean of these numbers is $\frac{\alpha_1 + \cdots + \alpha_n}{n}$, and their geometric mean is the positive n-th root of their product, viz. $(\alpha_1 \cdots \alpha_n)^{1/n}$. Then we have

$$(\alpha_1 \cdots \alpha_n)^{1/n} \leq \frac{1}{n}(\alpha_1 + \cdots + \alpha_n). \tag{16}$$

Equality holds iff α_j are equal. *Hint:* Write $\alpha_j = e^{x_j}$ and apply the inequality 2 in Jensen's theorem to the strictly convex function e^x.

Ex. 29. Generalized AM-GM Inequality. Let $x_i \in \mathbb{R}$, $x_i \geq 0$. Let $\alpha_i \in \mathbb{R}$, $\alpha_i > 0$ such that $\sum \alpha_i = 1$. Then

$$x_1^{\alpha_1} \cdots x_n^{\alpha_n} \leq \alpha_1 x_1 + \cdots + \alpha_n x_n. \tag{17}$$

Equality holds iff all x_j are equal. *Hint:* Let $y_i := \log x_i$ and $f(t) = e^t$.

Ex. 30. Deduce the standard AM-GM Inequality as a special case of the generalized AM-GM inequality (29).

Ex. 31. Young's Inequality. Yet another special case of the generalized AM-GM inequality:

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}, \text{ for positive } x, y \text{ and } p > 1. \tag{18}$$

Equality holds iff $x^p = y^q$.

Let $\alpha_1 = \frac{1}{p}$ and $\alpha_2 = \frac{1}{q}$, where $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, i.e. $\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-q}{p}$. Take $x_1 = x^p$ and $x_2 = y^q$ in the inequality 29.

You may also directly derive this inequality as follows: Write $x = e^{a/p}$ and $y = e^{b/p}$ and use the convexity of e^t.

Ex. 32. This is a special case of the inequality (18). Put $a = x^p$, $b = y^q$ to get

$$a^{\frac{1}{p}} \cdot b^{\frac{1}{q}} \leq \frac{a}{p} + \frac{b}{q}. \tag{19}$$

Equality holds iff $a = b$.

Ex. 33. Hölder's Inequality. Let $p > 1$. Let q be defined by the equation: $\frac{1}{p} + \frac{1}{q} = 1$. We prove first the discrete version of Hölder's inequality:

$$\sum_{j=1}^{n} |z_j||w_j| \leq (\sum_{j=1}^{n} |z_j|^p)^{1/p} (\sum_{j=1}^{n} |w_j|^q)^{1/q}, \qquad z, w \in \mathbb{C}^n. \tag{20}$$

Equality holds iff there exists a nonnegative constant λ such that $|z_j|^p = \lambda |w_j|^q$ for $1 \leq j \leq n$. *Hint:* Let $\|z\|_p := (\sum_{j=1}^n |z_j|^p)^{1/p}$ and $\|w\|_q := (\sum_{j=1}^n |w_j|^q)^{1/q}$. We apply the inequality (18) to $x = \frac{|z_j|}{\|z\|_p}$ and $y = \frac{|w_j|}{\|w\|_q}$ and sum over the index j. We get the result upon simplification.

We also have a continuous version of Hölder's inequality: Let $f, g \in C[0,1]$. Let $\|f\|_p := [\int_0^1 |f|^p]^{\frac{1}{p}}$ etc. Then we have, for $f, g \in C[0,1]$,

$$\int_0^1 |fg| \leq \|f\|_p \|g\|_q. \tag{21}$$

Equality holds iff there exists α, β not both zero in $\mathbb{R}$ such that $\alpha f^p + \beta g^q \equiv 0$ on $[0,1]$. *Hint:* Take $a = \frac{|f|}{\|f\|_p}$, $b = \frac{|g|}{\|g\|_q}$. Apply the inequality 18 to this and integrate over the interval to get Holder's inequality.

The case $p = 2$ is known as the (Cauchy-) Schwarz inequality.

Ex. 34 (Minkowski's Inequality)**.** Let the notation be as in the Hölder's inequality. We have

$$\|z + w\|_p \leq \|z\|_p + \|w\|_p, \qquad z, w \in \mathbb{C}^n, \quad p > 1. \tag{22}$$

Equality holds iff there exists a nonnegative λ such that $w_j = \lambda z_j$ for $1 \leq j \leq n$. *Hint:* Let $\alpha := \|z\|_p$ and $\beta := \|w\|_p$. Let a_j and b_j be defined by $|z_j| = \alpha a_j$ and $|w_j| = \beta b_j$. Then $\|a\|_p = 1 = \|b\|_p$. If we take $t = \alpha/(\alpha + \beta)$, then $1 - t = \beta/(\alpha + \beta)$. We have

$$\begin{aligned} |z_j + w_j|^p &\leq (|z_j| + |w_j|)^p \\ &= [\alpha a_j + \beta b_j]^p \\ &\leq (\alpha + \beta)^p [t a_j + (1-t) b_j]^p \\ &\leq (\alpha + \beta)^p t a_j^p + (1-t) b_j^p, \end{aligned}$$

by the convexity of the function $\varphi(t) = t^p$ on $[0, \infty)$ for $p > 1$. Summing both the extreme sides of the above inequality will yield the result $\|z + w\|_p^p \leq (\|z\|_p + \|w\|_p)^p$.

We can derive Minkowski's inequality from Hólder's as follows:

$$\begin{aligned} |z_j + w_j|^p &= |z_j + w_j|^{p-1} |z_j + w_j| \\ &\leq |z_j + w_j|^{p-1} |z_j| + |z_j + w_j|^{p-1} |w_j| \end{aligned}$$

Summing over j, we get

$$\sum_j |z_j + w_j|^p \leq \sum_j |z_j + w_j|^{p-1}|z_j| + \sum_j |z_j + w_j|^{p-1}|w_j|.$$

Apply Hölder's inequality to the two sums on right side with exponents q and p respectively and simplify the result.

Ex. 35. Formulate and prove a continuous version of Minkowski's inequality.

Theorem 36 (Jensen's Inequality)**.** *Let $f\colon [0,1] \to (a,b)$ be a continuous function. Assume that $\psi\colon (a,b) \to \mathbb{R}$ is continuous and convex. Then*

$$\psi(\int_0^1 f(t)dt) \leq \int_0^1 (\psi \circ f)(t)dt.$$

Proof. Since $\psi \circ f \in \mathcal{C}[0,1]$, the right hand side of the inequality makes sense. Let $\alpha := \int_0^1 f(r)dr$. Then $\alpha \in (a,b)$. By Lemma 20, there exists a constant M such that

$$\psi(x) \geq \psi(\alpha) + M(x - \alpha) \text{ for any } x \in (a,b) \tag{23}$$

Now we take $x = f(r)$ for $r \in [0,1]$ and get from (23)

$$\psi \circ f(r) \geq \psi(\alpha) + M(f(r) - \alpha). \tag{24}$$

We integrate (24) and simplify the result to get Jensen's inequality. □

Ex. 37. One can view the inequality of Jensen's theorem (Eq. 2) as the discrete version of Jensen's inequality: *Hint:* We use the notation of Eq. 2. Let $\alpha := \sum_j t_j x_j$. Consider a supporting function at $(\alpha, f(\alpha))$ and proceed as in the continuous version.

Ex. 38 (Hadamard's Inequality)**.** Let $f\colon (0,1) \to \mathbb{R}$ be continuous and convex. Then

$$f(1/2) \leq \int_0^1 f(t)\, dt.$$

There is a physical interpretation of this inequality: Let $f(t)$ be the velocity of a particle moving along a straight line with an increasing acceleration from time $t = 0$ to time $t = 1$. Then its velocity at $t = 1/2$ cannot exceed the average velocity of the entire trip.

Acknowledgements: I thank Mr. Ajit Kumar for drawing the pictures and for a critical reading of the earlier versions.

This article is an expanded version of the lectures given at S.P. College, Pune on February 9, 1999.

References

1. G.H. Hardy, J.E. Littlewood and G. Polya, *Inequalities*, Cambridge University Press, 1952.
2. E.F. Beckenbach and R. Bellman. *Inequalities*, Springer Verlag, 1965
3. W. Fleming, *Functions of several variables*, Springer-Verlag, UTM series.

The first one is a classic though it uses a bit old-fashioned notation. The second one is more modern and contains newer results. The last is ideal if you want to learn about convex functions of several variables.