Deriving Some Inequalities from a Property of Convex Functions

Stefan M. Stefanov
Department of Informatics, Neofit Rilski South-Western University
Blagoevgrad, BULGARIA; e-mail: stefm@swu.bg

MATHEMATICS EDUCATION

Abstract

This paper is devoted to deriving some inequalities, among them the inequality of the arithmetic and geometric means, triangle inequality, Minkowski's inequality, etc., from Jensen's inequality applied for some convex functions. Additional properties of convex functions are also considered.

Key words and phrases. Inequalities, convex functions, Jensen's inequality.
AMS Subject Classification. 52A40, 52A41.

1 Introduction. Preliminaries

Definition 1 A numerical function $f : X \to \mathbb{R}$ defined on a convex nonempty set $X \subset \mathbb{R}^n$ is said to be *convex* if for every $\mathbf{x}_1, \mathbf{x}_2 \in X, \quad t \in [0,1]$ we have that

$$f(t\mathbf{x}_1 + (1-t)\mathbf{x}_2) \le tf(\mathbf{x}_1) + (1-t)f(\mathbf{x}_2), \tag{1.1}$$

and it is said to be *strictly convex* if for every $\mathbf{x}_1, \mathbf{x}_2 \in X, \quad \mathbf{x}_1 \ne \mathbf{x}_2, \quad 0 < t < 1$ we have that

$$f(t\mathbf{x}_1 + (1-t)\mathbf{x}_2) < tf(\mathbf{x}_1) + (1-t)f(\mathbf{x}_2). \tag{1.2}$$

Obviously, each strictly convex function is convex.

If we take " $\ge$ " instead " $\le$ " in (1.1) and " $>$ " instead " $<$ " in (1.2), we obtain definitions of *concave* and *strictly concave* function, respectively.

Theorem 1 *Let f be a numerical differentiable function on an open convex set $X \subset \mathbb{R}^n$. A necessary and sufficient condition that f be convex is that for each $\mathbf{x}_1, \mathbf{x}_2 \in X$,*

$$\langle f'(\mathbf{x}_2) - f'(\mathbf{x}_1), \mathbf{x}_2 - \mathbf{x}_1 \rangle \ge 0.$$

Theorem 2 *Let f be a numerical twice-differentiable function on an open convex set $X \subset \mathbb{R}^n$. f is convex on X if and only if $f''(\mathbf{x})$ is positive semidefinite on X, that is, for each $\mathbf{x} \in X$,*

$$\langle \mathbf{y}, f''(\mathbf{x})\mathbf{y} \rangle \geq 0$$

for all $\mathbf{y} \in \mathbb{R}^n$.

If f is a function in one variable then the following (more specific) propositions hold.

Proposition 1 ([2, pp. 133-135], [8, p. 45]) *A real-valued function $f : (a, b) \to \mathbb{R}$ is convex on (a, b) if and only if the "slope" function (the difference quotient)*

$$\Delta_{x_0}(x) \stackrel{\text{def}}{=} \frac{f(x) - f(x_0)}{x - x_0}, \ x \in (a, b)\backslash\{x_0\}$$

is monotone nondecreasing in $(a, b)\backslash\{x_0\}$ for an arbitrary $x_0 \in (a, b)$ and for all choices of x not equal to x_0.

Similarly, strict convexity is characterized by strict increasing "slope" function $\Delta_{x_0}(x)$ of $x \in (a, b)\backslash\{x_0\}$.

Proposition 2 *Let $f : (a, b) \to \mathbb{R}$ and $f'(x)$ exists for any $x \in (a, b)$. f is convex (strictly convex) in (a, b) if and only if f' is monotone nondecreasing (monotone increasing) in (a, b).*

Proposition 3 *Let $f : (a, b) \to \mathbb{R}$ and $f''(x)$ exists for any $x \in (a, b)$. Necessary and sufficient condition for $f(x)$ to be convex (strictly convex) in (a, b) is $f''(x) \geq 0 (f''(x) > 0)$ for any x in (a, b).*

Theorem 3 (*Jensen's inequality* 1906, Johann Ludwig Jensen 1859–1925, [9])
Let f be a real-valued function defined on a convex subset D of $\mathbb{R}^n$, and let $\lambda_1 x_1 + \ldots + \lambda_n x_n$ be a convex combination of points $x_1, \ldots, x_n$ in D, that is, $\lambda_i \geq 0, i = 1, \ldots, n$ and $\lambda_1 + \ldots + \lambda_n = 1$. Then f is convex if and only if

$$f(\lambda_1 x_1 + \ldots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n).$$

Proof. Without loss of generality let $\lambda_i > 0$, otherwise the number of terms in the sum is less than m.

Necessity. (By induction) Let f be convex. If $m = 2$ then f is convex by definition.

Let the inequality be satisfied for $m = k - 1$. Verify it for $m = k$. Consider the point

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i \quad \text{where} \quad \mathbf{x}_i \in X, \ \lambda_i > 0, i = 1, \ldots, k, \ \sum_{i=1}^{k} \lambda_i = 1.$$

Then

$$f\left(\lambda_1 \mathbf{x}_1 + \sum_{i=2}^{k} \lambda_i \mathbf{x}_i\right) = f\left(\lambda_1 \mathbf{x}_1 + \lambda \sum_{i=2}^{k} \frac{\lambda_i}{\lambda} \mathbf{x}_i\right)$$

where $\lambda = \sum_{i=2}^{k} \lambda_i$. Since $\frac{\lambda_i}{\lambda} > 0$, $\sum_{i=2}^{k} \frac{\lambda_i}{\lambda} = \frac{1}{\lambda}\sum_{i=2}^{k} \lambda_i = 1$, $\mathbf{x}_i \in X$ and X is a convex set, then

$$\mathbf{y} \stackrel{\text{def}}{=} \sum_{i=2}^{k} \frac{\lambda_i}{\lambda}\mathbf{x}_i \in X.$$

Because f is convex and $\mathbf{x}_1, \mathbf{y} \in X$, $\lambda_1 + \lambda = 1$, $\lambda_1, \lambda \in [0,1]$, then

$$f(\mathbf{x}) = f(\lambda_1\mathbf{x}_1 + \lambda\mathbf{y}) \le \lambda_1 f(\mathbf{x}_1) + \lambda f(\mathbf{y}).$$

On the other hand,

$$f(\mathbf{y}) \equiv f\left(\sum_{i=2}^{k} \frac{\lambda_i}{\lambda}\mathbf{x}_i\right) \le \sum_{i=2}^{k} \frac{\lambda_i}{\lambda} f(\mathbf{x}_i)$$

according to the inductive hypothesis. From both inequalities it follows that

$$f\left(\sum_{i=1}^{k} \lambda_i\mathbf{x}_i\right) \equiv f\left(\lambda_1\mathbf{x}_1 + \lambda\sum_{i=2}^{k} \frac{\lambda_i}{\lambda}\mathbf{x}_i\right) \equiv f(\lambda_1\mathbf{x}_1 + \lambda\mathbf{y}) \le$$

$$\le \lambda_1 f(\mathbf{x}_1) + \lambda\sum_{i=2}^{k} \frac{\lambda_i}{\lambda} f(\mathbf{x}_i) = \lambda_1 f(\mathbf{x}_1) + \sum_{i=2}^{k} \lambda_i f(\mathbf{x}_i) = \sum_{i=1}^{k} \lambda_i f(\mathbf{x}_i).$$

Alternative proof of "necessity" part of Theorem 3.

As above, without loss of generality let $\lambda_i > 0, i = 1, \ldots, m$. If $\mathbf{x}_i \notin \mathit{dom}\, f, i = 1, \ldots, m$ then $f(\mathbf{x}_i) = +\infty$, $\lambda_i f(\mathbf{x}_i) = +\infty$ and Jensen's inequality is trivially satisfied, where *dom* f is the effective domain of f.

Otherwise, let $\mathbf{x}_i \in \mathit{dom}\, f$, $i = 1, \ldots, m$. Since the epigraph *epi* f is a convex set (f is convex) then $(\mathbf{x}_i, f(\mathbf{x}_i)) \in \mathit{epi}\, f$, $i = 1, \ldots, m$ imply

$$(\lambda_1\mathbf{x}_1 + \ldots + \lambda_m\mathbf{x}_m,\ \lambda_1 f(\mathbf{x}_1) + \ldots + \lambda_m f(\mathbf{x}_m)) \in \mathit{epi} f$$

according to a property of the epigraph *epi* f. Therefore

$$f(\lambda_1\mathbf{x}_1 + \ldots + \lambda_m\mathbf{x}_m) \le \lambda_1 f(\mathbf{x}_1) + \ldots + \lambda_m f(\mathbf{x}_m)$$

by definition of *epi* f.

Sufficiency. Let Jensen's inequality be satisfied for f with arbitrary m. Then this inequality with $m = 2$ implies convexity of f. ■

Definition 2 ([6], p. 205) Let δ be a positive integer. A real-valued function f defined on a convex cone C with vertex at θ is said to be *homogeneous of degree* δ if $f(\lambda x) = \lambda^\delta f(x)$ for any $x \in C$ and any $\lambda \ge 0$.

Theorem 4 ([6], p. 205) *Suppose that f is a convex function defined on a convex cone C with vertex at θ and that $f(x) > 0$ for all $x \neq \theta$ in C. If f is homogeneous of degree δ then the function g defined on C by $g(x) \equiv [f(x)]^{\frac{1}{\delta}}$ is convex on C.*

Examples.

1. Affine (linear) function $l(\mathbf{x}) = \langle \mathbf{c}, \mathbf{x} \rangle + a$, $\mathbf{c}, \mathbf{x} \in \mathbb{R}^n, a \in \mathbb{R}$, where $\langle \mathbf{x}, \mathbf{y} \rangle$ denotes the scalar (inner) product of $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, is both convex and concave.
2. Quadratic function $f(x) = x^2$, $x \in \mathbb{R}$ is strictly convex according to Proposition 3.
3. $f(x) = |x|$, $x \in \mathbb{R}$ is a convex function.
4. $f(x) = x^p$, $p > 1, x \in (0, \infty)$ is strictly convex according to Proposition 3.
5. $f(x) = \ln x$, $x \in (0, \infty)$ is strictly concave on $(0, \infty)$ according to Proposition 3.
6. $f(x) = \ln \sin x$ is strictly concave on $(0, \pi)$ according to Proposition 3.
7. $f(x) = \ln \sin x - \ln x$ is strictly concave for $x \in (0, \pi)$ according to Proposition 3.

2 Applications of Jensen's Inequality

1. Weighted (Generalized) Arithmetic-Geometric Mean Inequality

Prove the inequality

$$x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha^n} \leq \alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n, \tag{2.1}$$

where $x_i \geq 0,\ \alpha_i \geq 0, i = 1, \ldots, n,\ \sum\limits_{i=1}^{n} \alpha_i = 1$.

Using Jensen's inequality with the (strictly) convex function $f(x) = -\ln x$, $x \in (0, \infty)$ (Example 5) we obtain

$$-\ln \left(\sum_{i=1}^{n} \alpha_i x_i \right) \leq \sum_{i=1}^{n} \alpha_i (-\ln x_i),\ \alpha_i \geq 0, i = 1, \ldots, n,\ \sum_{i=1}^{n} \alpha_i = 1,$$

$$\sum_{i=1}^{n} \alpha_i \ln x_i \leq \ln \left(\sum_{i=1}^{n} \alpha_i x_i \right),$$

$$\sum_{i=1}^{n} \ln x_i^{\alpha_i} \leq \ln \left(\sum_{i=1}^{n} \alpha_i x_i \right),$$

$$\ln \left(\prod_{i=1}^{n} x_i^{\alpha_i} \right) \leq \ln \left(\sum_{i=1}^{n} \alpha_i x_i \right).$$

Since $\ln x$ is an increasing function with $x > 0$ (and $e > 1$) then

$$\prod_{i=1}^{n} x_i^{\alpha_i} \leq \sum_{i=1}^{n} \alpha_i x_i$$

for $x_i \geq 0, \alpha_i \geq 0, i = 1, \ldots, n,\ \sum\limits_{i=1}^{n} \alpha_i = 1$.

Equality holds if and only if $x_1 = x_2 = \ldots = x_n$.

When $\alpha_i = \frac{1}{n}, i = 1, \ldots, n$ (these α_i's satisfy the requirements) we obtain the traditional arithmetic-geometric mean inequality

$$\left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^{n} x_i, \quad x_i \geq 0, i = 1, \ldots, n.$$

2. (Weighted) Triangle Inequality
Prove the inequality

$$\left|\sum_{i=1}^{n} \alpha_i x_i\right| \le \sum_{i=1}^{n} \alpha_i |x_i|, \quad \alpha_i \ge 0, i = i, \ldots, n, \ \sum_{i=1}^{n} \alpha_i = 1, \tag{2.2}$$

Application of Jensen's inequality to the convex function $f(x) = |x|, x \in \mathbb{R}$ (Example 3) leads to (2.2), the weighted triangle inequality with weights $\alpha_i, i = 1, \ldots, n$.

When $\alpha_i = \frac{1}{n}, i = 1, \ldots, n$, we have

$$\left|\sum_{i=1}^{n} x_i\right| \le \sum_{i=1}^{n} |x_i|,$$

the triangle inequality.

When $n = 2$ we have the traditional triangle inequality for two elements.

3. Prove that the following inequality holds

$$\sqrt[n]{\sin x_1 \sin x_2 \ldots \sin x_n} \le \sin \frac{x_1 + x_2 + \ldots + x_n}{n}, \tag{2.3}$$

where n is a positive integer, and $x_1, x_2, \ldots, x_n$ are in $(0, \frac{\pi}{2})$.

Proof. Application of Jensen's inequality to the strictly convex function $f(x) = -\ln \sin x$ over the interval $(0, \frac{\pi}{2})$ (Example 6) leads to the inequality (2.3).

Remark 1. *Condition $x_i \in (0, \frac{\pi}{2}), i = 1, \ldots, n$ guarantees that $\sin x_i \ge 0, i = 1, \ldots, n$ and nonnegativity of the radicand in the inequality (2.3).*

4. Prove the inequality

$$\frac{\frac{1}{n}\sum_{i=1}^{n} x_i}{\sin\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)} \le \sqrt[n]{\frac{x_1 x_2 \ldots x_n}{\sin x_1 \sin x_2 \ldots \sin x_n}}, \tag{2.4}$$

where n is a positive integer, and $x_1, x_2, \ldots, x_n$ are in $(0, \frac{\pi}{2})$.

By using Jensen's inequality for the strictly convex function $f(x) = \ln x - \ln \sin x$, $x \in (0, \pi)$ (Example 7) we get

$$\ln\left(\sum_{i=1}^{n} \alpha_i x_i\right) - \ln \sin\left(\sum_{i=1}^{n} \alpha_i x_i\right) \le \sum_{i=1}^{n} \alpha_i (\ln x_i - \ln \sin x_i),$$

$$\ln \frac{\sum_{i=1}^{n} \alpha_i x_i}{\sin\left(\sum_{i=1}^{n} \alpha_i x_i\right)} \le \sum_{i=1}^{n} \alpha_i \ln \frac{x_i}{\sin x_i},$$

$$\ln \frac{\sum_{i=1}^{n} \alpha_i x_i}{\sin\left(\sum_{i=1}^{n} \alpha_i x_i\right)} \le \sum_{i=1}^{n} \ln\left(\frac{x_i}{\sin x_i}\right)^{\alpha_i},$$

$$\ln \frac{\sum_{i=1}^{n} \alpha_i x_i}{\sin\left(\sum_{i=1}^{n} \alpha_i x_i\right)} \le \ln \prod_{i=1}^{n} \left(\frac{x_i}{\sin x_i}\right)^{\alpha_i}.$$

Since $\ln x$ is an increasing function of x with $x > 0$ and $e > 1$ then

$$\frac{\sum_{i=1}^{n} \alpha_i x_i}{\sin\left(\sum_{i=1}^{n} \alpha_i x_i\right)} \le \prod_{i=1}^{n} \left(\frac{x_i}{\sin x_i}\right)^{\alpha_i}.$$

In the special case when $\alpha_i = \frac{1}{n}, i = 1, \dots, n$, we obtain the inequality (2.4).

5. Minkowski's Inequality 1896 (Hermann Minkowski 1864–1909)
Prove the inequality

$$\left(\sum_{i=1}^{n} (\alpha_i + \beta_i)^p\right)^{\frac{1}{p}} \le \left(\sum_{i=1}^{n} \alpha_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} \beta_i^p\right)^{\frac{1}{p}}, \tag{2.5}$$

where $\alpha_1, \alpha_2, \dots, \alpha_n$ and $\beta_1, \beta_2, \dots, \beta_n$ are nonnegative real numbers, and p is a positive integer (that is, $p \ge 1$).

Denote $\mathbf{x} \equiv (\alpha_1, \dots, \alpha_n)$, $\mathbf{y} \equiv (\beta_1, \dots, \beta_n)$. Consider the function $f(\mathbf{x}) \equiv \sum_{i=1}^{n} \alpha_i^p$.

Our purpose is to prove that $f(\mathbf{x})$ is convex and homogeneous of degree p on the nonnegative orthant in $\mathbb{R}^n$. Indeed,

$$\begin{aligned}
f(t\mathbf{x} + (1-t)\mathbf{y}) &= \sum_{i=1}^{n} [t\alpha_i + (1-t)\beta_i]^p = \\
&= \sum_{i=1}^{n} \left\{\sum_{k=0}^{p} \binom{p}{k} (t\alpha_i)^k [(1-t)\beta_i]^{p-k}\right\} \equiv \\
&\equiv \sum_{i=1}^{n} \left\{(1-t)^p \beta_i^p + \sum_{k=1}^{p-1} \binom{p}{k} (t\alpha_i)^k [(1-t)\beta_i]^{p-k} + t^p \alpha_i^p\right\} < \\
&< \sum_{i=1}^{n} \{(1-t)^p \beta_i^p + t^p \alpha_i^p\} \le \\
&\le \sum_{i=1}^{n} \{(1-t)\beta_i^p + t\alpha_i^p\} \equiv (1-t)f(\mathbf{y}) + tf(\mathbf{x}),
\end{aligned}$$

where we have used that

$$\sum_{k=1}^{p-1} \binom{p}{k} (t\alpha_i)^p[(1-t)\beta_i]^{p-k} \geq 0$$

with nonnegative $\alpha_i, \beta_i, i = 1, \ldots, n,\ t \in [0,1], p \geq 1$; and $(1-t)^p \leq 1-t,\ t^p \leq t$ with $t \in [0,1], p \geq 1$. Therefore, $f(x)$ is convex by definition.

Moreover

$$f(\lambda \mathbf{x}) = \sum_{i=1}^{n} (\lambda \alpha_i)^p = \lambda^p \sum_{i=1}^{n} \alpha_i^p = \lambda^p f(\mathbf{x}),$$

p is a positive integer. Therefore f is homogeneous of degree p according to Definition 2. From Theorem 4 it follows that function $\left(\sum_{i=1}^{n} \alpha_i^p\right)^{\frac{1}{p}}$ is convex on $\mathbb{R}_+^n$. Hence

$$\left[\sum_{i=1}^{n} \left(\frac{1}{2}\alpha_i + \frac{1}{2}\beta_i\right)^p\right]^{\frac{1}{p}} \leq \frac{1}{2}\left(\sum_{i=1}^{n} \alpha_i^p\right)^{\frac{1}{p}} + \frac{1}{2}\left(\sum_{i=1}^{n} \beta_i^p\right)^{\frac{1}{p}}$$

according to Jensen's inequality.

Multiplying both sides by 2, we obtain Minkowski's inequality (2.5).

3 Conclusions

Other inequalities can also be obtained by using the approach discussed in this paper. For example, *Hölder's inequality* (1889, Ludwig Otto Hölder 1859–1937)

$$\sum_{i=1}^{n} a_i b_i \leq \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n} b_i^q\right)^{\frac{1}{q}}$$

where a_i and $b_i, i = 1, \ldots, n$ are nonnegative real numbers, $p > 1, q > 1, \frac{1}{p} + \frac{1}{q} = 1$ can be obtained by using Jensen's inequality with the convex function $f(x) = x^p, x \geq 0, p > 1$. In the special case when $p = q = 2$, Hölder's inequality is known as *Cauchy-Schwarz inequality* (1821, Baron Augustin Louis Cauchy 1789–1857, Karl Hermann Amandus Schwarz 1843–1921).

4 Bibliographical Notes

Various aspects of inequalities are discussed in [1], [3], [5], [10].

Some inequalities and related topics are considered in [2], [4, Appendix B], etc.

Convex functions and convex sets, as well as their properties are studied in [6], [7], [8], [9].

References

[1] Edwin Beckenbach, Richard Bellman, *An Introduction to Inequalities*, The Mathematical Association of America, Random House, New York – Toronto, 1961.

[2] A. Dorogovtzev, *Mathematical Analysis*, Visha Shkola Publishing House, Kiev, 1985 (in Russian).

[3] G.H. Hardy, J.E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, 1988.

[4] Roger A. Horn and Charles A. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1990.

[5] Nicholas D. Kazarinoff, *Geometric Inequalities*, The Mathematical Association of America, Random House, New York – Toronto, 1961.

[6] Steven R. Lay, *Convex Sets and Their Applications*, John Wiley & Sons, New York, 1982.

[7] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[8] R. Tyrrell Rockafellar, Roger J.-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.

[9] S.M. Stefanov, *Separable Programming: Theory and Methods*, Kluwer Academic Publishers, Dordrecht-Boston-London, 2001.

[10] G. Stephenson, *Inequalities and Optimal Problems in Mathematics and the Sciences*, Longman, London, 1971.