Thai Journal of Mathematics
Special Issue: The 17^{th} IMT-GT ICMSA 2021
Pages 98–105 (2022)

http://thaijmath.in.cmu.ac.th

ISSN 1686-0209

A proof of the Minkowski inequalities based on convex homogeneous functions

Seppo Karrila[1], Taewee Karrila[2] and Alex Karrila[3,*]

[1] *Faculty of Science and Industrial Technology, Prince of Songkla University, Surat Thani Campus, Surat Thani 84000, Thailand*
e-mail : `seppo.karrila@gmail.com`

[2] *Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, Pattani 94000, Thailand*
e-mail : `taewee.t@psu.ac.th`

[3] *Faculty of Science and Engineering, Åbo Akademi University, Henriksgatan 2, 20500 Åbo, Finland*
e-mail : `alex.karrila@abo.fi`

Abstract The triangle inequality for p-norms, also known as the Minkowski inequality, is often proven with algebra relying on the Hlder inequality. We give an appealing alternative proof relying on elementary convex analysis that we hope is pedagogically useful. The core lemma is the following. Let $K \subset \mathbb{R}^n$ be a convex cone and $g : K \to \mathbb{R}_{\geq 0}$ be a positively homogeneous function with $g(x) > 0$ for $x \neq 0$. Then, g is convex (resp. concave) if and only if the sublevel set $\{x \in K : g(x) \leq 1\}$ (resp. its complement) is convex. This yields a nice characterization of a norm via its unit ball. As roots and powers preserve the sublevel set at height 1, another immediate consequence is the following: if $f : K \to \mathbb{R}_{\geq 0}$ is a convex (resp. concave) positively homogeneous function of degree $p \geq 1$ (resp. $0 < p \leq 1$), with $f(x) > 0$ for $x \neq 0$, then $g(x) := [f(x)]^{1/p}$ is convex (resp. concave). This readily implies the Minkowski and reverse Minkowski inequalities; also some other applications are briefly exemplified.

MSC: 46N10; 26B25; 39B62; 97I99
Keywords: functional analysis; convexity of real functions of several variables; functional inequalities

Submission date: 15.03.2022 / Acceptance date: 31.03.2022

1. Introduction

The Minkowski inequality is the triangle inequality for p-norms, and most readers of mathematics have certainly seen it derived, perhaps through several different proofs. The 'classic' derivation (see, e.g., [1, 2]) establishes first Young's, then Hölder's, and finally Minkowski's inequalities. While this strategy does provide a short proof and introduce useful intermediate steps, it may to a newcomer appear as a "rabbit out of the hat" type approach, where the overall scheme is not transparent. The proof and the intermediate results are also rather specific for the p^{th} power expressions.

*Corresponding author.

Published by The Mathematical Association of Thailand.
Copyright © 2022 by TJM. All rights reserved.

We provide here a proof where the Minkowski inequality is deduced as a standard "verify convexity" task. We believe this can be pedagogically useful as the link from convexity to the triangle inequality is then immediate (see Equation (3.1) below). This approach is further supported by the increasing weight put on convex analysis in mathematics curricula, due to its enormous importance to especially numerical optimization and computing. The idea of deriving the Minkowski inequality via convex analysis was pursued also in [3], with a different verification of convexity than in this note.

The core of the convexity verification here is Lemma 3.1 which states that a positive, positively homogeneous function is convex (resp. concave) if and only if its sublevel set at height one (resp. its complement) is convex. This should be compared to the general function composition rules of convex analysis (see Section 2.2 below), which for instance guarantee that composing with an increasing *convex* function preserves convexity: by Lemma 3.1, if the composite function is positive homogeneous, then actually composing with *any* increasing function will do. Applying this principle for the p^{th} root function effortlessly yields the Minkowski inequality for $1 \leq p < \infty$, and the concave analogue implies the reverse Minkowski inequality for $0 < p \leq 1$; also other applications of the general convex analysis results are illustrated in Section 4.

The analysis of convex functions here can be linked to the gauge functions, also known as Minkowski functionals [1, 4], while for concave functions the results presented seem to lack such connection to textbook materials in convex analysis. The authors did eventually find a proof of the Minkowski inequality based on positively homogeneous functions in prior literature, namely in the book by Güler [5]. That version does not emphasize the role of sublevel sets, but instead presents clever algebraic manipulations.

2. Preliminaries

In this section, we provide the necessary and sufficient amount of basic convex analysis for the main results, to facilitate the discussion and to explicate the required background knowledge for pedagogical use.

Throughout this paper we treat for simplicity subsets of the Euclidean space $\mathbb{R}^n$ and real-valued functions on them; $\mathbb{R}^n$ could, however, be replaced with any real vector space without any modifications in the arguments.

2.1. Basic definitions

Definition 2.1. A set $K \subset \mathbb{R}^n$ is *convex* if, for any $x, y \in K$ and $\theta \in [0,1]$, also the point $\theta x + (1-\theta)y$ is in K.

Thus, the line segment connecting any two points of K must also be in K. Informally, the boundary of K "bulges outwards", not leaving outside any such point that could obstruct line of sight between some two points of K.

Definition 2.2. A real valued function f defined on a convex set K is *convex*, if, for any $x, y \in K$ and $\theta \in [0,1]$, we have

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y). \tag{2.1}$$

A redundant but useful term is concavity; if the reverse of the inequality above holds, we say that f is *concave*, i.e., f is convex if and only if its negative $-f$ is concave.

Pictorially, convexity means that along the straight line segment connecting any x and y, the function value is below its linear interpolation based on $f(x)$ and $f(y)$. The graph of the restriction function along the line segment thus has a "smile shape".

This study concerns functions defined on the following type of sets.

Definition 2.3. A subset K in Euclidean space $\mathbb{R}^n$ is a *convex cone* if it is closed with respect to summing and non-negative scaling of vectors, i.e., for all $x, y \in K$ and $t \geq 0$, we have $x + y \in K$ and $t \cdot x \in K$.

Note also that, for a convex cone K, the sets tK and K are equal whenever $t > 0$. Equivalent to the definition, a subset K of $\mathbb{R}^n$ is a convex cone if and only if it is convex and closed with respect to non-negative scaling. Thus, to apply convexity theory to the class of functions below, convex cones are the natural domain type to consider.

Definition 2.4. Let K be a subset in the Euclidean space that is closed with respect to non-negative scaling. A function $f : K \to \mathbb{R}_{\geq 0}$ is *positively homogeneous of degree* p, if, for any $x \in K$ and $t \geq 0$, we have

$$f(tx) = t^p f(x);$$

in the case $p = 1$, f is called *positively homogeneous*, for short.

For instance, any vector norm in $\mathbb{R}^n$ is a positively homogeneous function.

2.2. Some elementary results

We will use the following (very) elementary convexity theory results.

i) Let $K \subset \mathbb{R}^n$ be convex. If $f : K \to \mathbb{R}$ is a convex function, then all its sublevel sets

$$S_a^f = \{x \in K : f(x) \leq a\}, \qquad a \in \mathbb{R}$$

are (empty or) convex sets. Likewise, sublevel sets defined with a strict inequality are convex.

ii) An increasing union of convex sets is convex.

These results follow directly from the definitions. Note that the converse implication of (i) is *in general not true* (think of the function $\sqrt{|x|}$ on $\mathbb{R}$) but we will later in Lemma 3.1 show that the converse *is true for suitable positively homogeneous functions.*

iii) A set $K \subset \mathbb{R}^n$ is convex if and only if, for all $a, b \geq 0$, we have

$$aK + bK = (a + b)K.$$

Proof. The equation above is equivalent to $\frac{a}{a+b}K + \frac{b}{a+b}K = \theta K + (1 - \theta)K = K$ (where we excluded the trivial case $a + b = 0$ and set $\theta = \frac{a}{a+b} \in [0, 1]$). In $\theta K + (1 - \theta)K = K$, the inclusion "$\supset$" holds for any $K \subset \mathbb{R}^n$, while "$\subset$" is the very definition of convexity. ∎

Finally, we will use the following *convex transformations* that the reader may either prove as an exercise or find in a textbook such as [6].[1]

iv) If $f : K \to I$, with a real interval I, is convex and $h : I \to \mathbb{R}$ is convex and increasing, then $h \circ f : K \to \mathbb{R}$ is convex.

[1]This particular textbook has been made available for download as a PDF file by one of its authors (search for "cvxbook"), and online lectures by Prof. Boyd are also accessible on the internet.

v) If $f : K \to I$, with a real interval I, is concave and $h : I \to \mathbb{R}$ is concave and increasing, then $h \circ f : K \to \mathbb{R}$ is concave.

Redundant transform rules are obtained by changing h to $-h$ (which reverses convexity/concavity and increasingness/decreasingness).

3. Proofs of the main statements

We are now ready to prove the main results.

In the following lemma, let K be a convex cone and the function $g : K \to \mathbb{R}_{\geq 0}$ be positively homogeneous (of degree one) with $g(x) > 0$ for $x \neq 0$. Let S denote the sublevel set $S := S_1^g$ and $C := K \setminus S$ its complement. Observe that by the homogeneity, $S_a^g = aS$ for all $a \geq 0$.

Lemma 3.1. *In the setup given above, g is a convex (resp. concave) function if and only if S (resp. C) is a convex set.*

Proof. g is convex $\Rightarrow$ S is convex: this is an instance of convexity fact (i).

S is convex $\Rightarrow$ g is convex:

Denote $g(x) = a$ and $g(y) = b$, so $x \in aS$ and $y \in bS$. Hence, $(x + y) \in aS + bS = (a + b)S$ where we used convexity fact (iii). As $(a + b)S = S_{a+b}^g$, we thus have $g(x + y) \leq a + b = g(x) + g(y)$ for any $x, y \in K$. Applying this formula and homogeneity again, we get, for any $\theta \in [0, 1]$,

$$g(\theta x + (1-\theta)y) \leq g(\theta x) + g((1-\theta)y) = \theta g(x) + (1-\theta)g(y),$$

i.e., g is convex.

The case where g is concave and C is convex is almost identical, and we leave it to the reader. ∎

The previous lemma yields a nice characterization of a norm. As indicated in the beginning of this text, the analogous characterization remains valid for every real vector space.

Corollary 3.2. *Let g be a positively homogeneous function (of degree one) on $\mathbb{R}^n$, with $g(x) > 0$ for $x \neq 0$ and $g(-x) = g(x)$. Then, the following are equivalent: i) g is convex; ii) the sublevel set $S = S_1^g$ is convex; and iii) g is a norm on $\mathbb{R}^n$.*

Proof. Statements (i) and (ii) are equivalent by the previous lemma. We prove (i) $\Leftrightarrow$ (iii). With the assumptions given, g is a norm if and only if it satisfies the triangle inequality

$$g(x + y) \leq g(x) + g(y).$$

In the proof above, we saw that this implies convexity for the homogeneous function g. Conversely, if g is convex and positively homogeneous, then

$$g(x + y) = 2g\left(\frac{x + y}{2}\right) \leq 2 \cdot \frac{g(x) + g(y)}{2} = g(x) + g(y). \tag{3.1}$$

∎

Another corollary of Lemma 3.1 is the following.

Corollary 3.3. *Let $K \subset \mathbb{R}^n$ be a convex cone and $f : K \to \mathbb{R}_{\geq 0}$ be a positively homogeneous function of degree $p \geq 1$, with $f(x) > 0$ for $x \neq 0$. Set $g(x) := [f(x)]^{1/p}$. Then, g is convex if and only if f is convex.*

Let us make some remarks here. First, combining Corollary 3.3 with convexity fact (iv), if f as above is convex, then so is $[f(x)]^q$ for any $q \geq 1/p$. Note also that applying fact (iv) without reference to Corollary 3.3 would only prove this for $q \geq 1$, a significantly weaker result.

Second, as regards restricting to $p \geq 1$ and convex functions f, note that a (non-zero) positively homogeneous functions of degree $p < 1$ (resp. $p > 1$) cannot be convex (resp. concave), because along a ray from the origin such function is strictly concave (resp. convex). The complementary case when convex analysis can be applied is thus when $0 < p \leq 1$ and f is concave, and it is treated below.

Proof. Suppose first that f is a convex function. Then, its sublevel set $S_1^f := \{x \in K : f(x) \leq 1\}$ is a convex set by fact (i). The latter coincides with the sublevel set $S := \{x \in K : g(x) \leq 1\}$ of g. The convexity of g now follows by the previous Lemma 3.1.

Suppose then that g is a convex function to the interval $I = \mathbb{R}_{\geq 0}$. Define an increasing convex function $h : I \to \mathbb{R}$ by $h(x) = x^p$. Thus, $f = h \circ g$ is convex by the convex transformations introduced in the previous section. ∎

Example 3.4 (The Minkowski inequality)**.** Let $p \geq 1$. On Euclidean space $\mathbb{R}^n$ (also a convex cone), the function $f(x) := \sum_{i=1}^n |x_i|^p$ is a sum of convex terms so it is convex, and obviously positively homogeneous of degree p. By the previous corollary, also $g(x) := [f(x)]^{1/p}$ is convex, and by Corollary 3.2 (whose assumptions are clearly satisfied here), it is a norm and satisfies the triangle inequality, i.e., the Minkowski inequality.

We yet formulate what happens in the concave case.

Corollary 3.5. *Let $K \subset \mathbb{R}^n$ be a convex cone and $f : K \to \mathbb{R}_{\geq 0}$ be a positively homogeneous function of degree $0 < p \leq 1$, with $f(x) > 0$ for $x \neq 0$. Set $g(x) := [f(x)]^{1/p}$. Then, the following are equivalent: i) g is concave; ii) f is concave; iii) the set $K \setminus \{0\}$ is convex and the function $[1/g(x)]^q$ on it is convex iii)(a) for any $q > 0$ iii)(b) for some $q > 0$.*

Analogously to the convex case we note that, by fact (v), if f as above is concave, then $[f(x)]^q$ is actually concave whenever $q \leq 1/p$. Note also that the third characterization has no analogue in the convex case, and poses a restriction on the set K through the existence of a concave f.

Proof. The proof of (i) $\Leftrightarrow$ (ii) is a concave analogue of that of Corollary 3.3.

To prove (i) $\Rightarrow$ (iii)(a), the set $K \setminus \{0\}$ is the union of the increasing sets $K \setminus S^g_{1/k} = \frac{1}{k}(K \setminus S) = \frac{1}{k}C$, $k \in \mathbb{N}$. Here C is the (strict) sublevel set of the convex function $-g$, hence convex. The convexity of $K \setminus \{0\}$ then follows by fact (ii). Talking about the convexity of the function $[1/g(x)]^q$ now makes sense in the first place. Next, raising to a negative power $-q$ is convex and decreasing on $I = \mathbb{R}_{>0}$, so this converts concave g to a convex function, by fact (v).

The implication (iii)(a) $\Rightarrow$ (iii)(b) is clear. For (iii)(b) $\Rightarrow$ (i), note that C is the (strict) sublevel set of the convex function $[1/g(x)]^q$, hence convex, and then apply Lemma 3.1. ∎

Example 3.6 (Reverse Minkowski for $0 < p \leq 1$)**.** Let $K \subset \mathbb{R}^n$ be the (positive) orthant $K = \{x \in \mathbb{R}^n : x_i \geq 0 \text{ for all } 1 \leq i \leq n\}$. Define $f : K \to \mathbb{R}_{\geq 0}$ via

$$f(x) := \sum_{i=1}^n x_i^p,$$

where this time $p \in (0, 1]$. As a sum of concave terms — Note that here it is crucial that we only consider one orthant instead of whole $\mathbb{R}^n$! — f is a concave function. By the previous corollary, also $g(x) := [f(x)]^{1/p}$ is concave, and combining with its homogeneity,

$$g(x+y) = 2g(\frac{x+y}{2}) \geq 2 \cdot \frac{g(x)+g(y)}{2} = g(x) + g(y). \tag{3.2}$$

This is the reverse Minkowski inequality on an orthant.

4. Some additional applications

We proved the Minkowski inequalities as a consequence of general convexity results. Let us illustrate some further applications.

First, as mentioned in the very beginning, the same convexity results hold with essentially identical proofs in any real vector space. This allows one to deduce the Minkowski integral inequalities.

Example 4.1 (The Minkowski integral inequalities)**.** Let $0 < p < +\infty$, and let $V = V(p)$ be the real vector space of Lebesgue measurable functions ϕ on $\mathbb{R}^m$ such that $\int_{\mathbb{R}^m} |\phi(w)|^p dw < \infty$.[2] Define the homogeneous function $f : V \to \mathbb{R}$ of degree p by

$$f(\phi) = \int_{\mathbb{R}^m} |\phi(w)|^p dw.$$

For $1 \leq p < +\infty$, $f : V \to \mathbb{R}$ is convex essentially due to the convexity of the real function $|\cdot|^p$. If we identify functions that coincide almost everywhere (i.e., we actually study the vector space $V/\sim$, where $\sim$ is the corresponding equivalence relation), f also becomes strictly positive. Applying Corollary 3.3 and then Corollary 3.2 (more precisely, their analogues on the vector space $V/\sim$), $g(\phi) := [f(\phi)]^{1/p}$ is a norm on $V/\sim$ and satisfies the triangle inequality, i.e., the Minkowski integral inequality. (For Riemann integrable functions, the same inequality can be proven by applying Example 3.4 to the Riemann sums of the integral $f(\phi)$.)

For $0 < p \leq 1$, study the set $K \subset V/\sim$ of (equivalence classes of) functions $\phi \in V$ that are non-negative Lebesgue almost everywhere, $\phi(w) \geq 0$ for a.e. $w \in \mathbb{R}^m$. K is easily shown to be a convex cone, and f is concave on it essentially due to the concavity of the real function $|\cdot|^p$ on the non-negative half-line $\mathbb{R}_{\geq 0}$. (Note again that the non-negativity of ϕ is crucial here.) By Corollary 3.5 on the vector space $V/\sim$, $g(\phi) := [f(\phi)]^{1/p}$ is concave on K, and using the homogeneity as in (3.2), we readily obtain the reverse Minkowski inequality for integrals: for all $\phi_1, \phi_2 \in K$,

$$g(\phi_1 + \phi_2) \geq g(\phi_1) + g(\phi_2).$$

Along similar lines as above, one can study p-integrable real functions on a general measure space. Another interesting application is a determinant inequality for symmetric, positive definite matrices. Also this inequality is due to Minkowski.

Example 4.2. (Minkowski's determinant inequality) In what follows, identify $\mathbb{R}^{m\times m} \cong \mathbb{R}^{m^2}$ to apply our previous results for the matrix space $\mathbb{R}^{m\times m}$, and for $A \in \mathbb{R}^{m\times m}$ denote $\det A = |A|$ for short. Denote by $\mathrm{SPD}(m)$ the set of symmetric, positive definite matrices

[2] To conclude that this set of functions indeed is a vector space, it is clearly closed under scaling, and for summation, note that $|\phi_1(w) + \phi_2(w)|^p \leq (2\max\{|\phi_1(w)|, |\phi_2(w)|\})^p \leq 2^p(|\phi_1(w)|^p + |\phi_2(w)|^p)$, which implies $\int_{\mathbb{R}^m} |\phi_1(w) + \phi_2(w)|^p dw < \infty$.

in $\mathbb{R}^{m\times m}$, i.e., symmetric matrices $A \in \mathbb{R}^{m\times m}$ such that $x^T Ax > 0$ for all $x \in \mathbb{R}^{m\times 1}$, $x \neq 0$. Minkowski's determinant inequality now reads

$$|A+B|^{1/m} \geq |A|^{1/m} + |B|^{1/m} \qquad \text{for all } A, B \in \mathrm{SPD}(m). \tag{4.1}$$

We give a short proof of this result.

First, the definition directly implies that $\mathrm{SPD}(m)$ is a convex set in $\mathbb{R}^{m\times m}$, and it is fairly standard that $\log |A|$ is concave on $\mathrm{SPD}(m)$,[3] equivalently,

$$\begin{aligned} & |A| = e^{\phi(A)}, && \phi : \mathrm{SPD}(m) \to \mathbb{R} \text{ concave} \\ \Leftrightarrow \quad & 1/|A| = e^{\psi(A)}, && \psi = -\phi \text{ convex.} \end{aligned}$$

By convexity fact (iv), exponentials of convex functions are themselves convex functions. Combining this with convexity fact (i),

$$\{A \in \mathrm{SPD}(m) \ : \ 1/|A| < 1\} \qquad \text{is a convex set.}$$

Finally, let K be the convex cone $\mathrm{SPD}(m) \cup \{0\}$ in $\mathbb{R}^{m\times m}$. Note that $g : K \to \mathbb{R}$ defined by $g(A) = |A|^{1/m}$ is a positive homogeneous function, and by the above

$$K \setminus S_1^g = \{A \in K \ : \ |A|^{1/m} > 1\} \qquad \text{is a convex set.}$$

Applying Lemma 3.1, we get that g is concave, and combining with its homogeneity as in (3.2) yields (4.1).

A different proof of (4.1) can be found in, e.g., [7, Theorem 7.8.8]. We conclude this example by remarking two simple consequences. First, by continuity, the inequality (4.1) remains valid even if we only require A and B to be symmetric and positive *semi*-definite. Second, for symmetric matrices, the trace is the sum of eigenvalues and the determinant is their product, so using the arithmetic–geometric inequality, we get

$$|A|^{1/m} + |B|^{1/m} \leq |A+B|^{1/m} \leq \mathrm{trace}(A+B)/m \quad \text{for all } A, B \in \mathrm{SPD}(m).$$

[3] Proofs can be found in, e.g., [6, Section 3.1.5] or [7, Theorem 7.6.7], or given through matrix algebra as follows: It suffices here to test concavity with $\theta = 1/2$ in the reverse of inequality (2.1) (this property is called *midpoint convexity/concavity*, which coincides with convexity/concavity for continuous functions). Using matrix square roots, we observe that the midpoint concavity $|\frac{A+B}{2}| \geq |A|^{1/2}|B|^{1/2} = |A^{1/2}B^{1/2}|$ is equivalent to $|\frac{I+A^{-1/2}B^{1/2}B^{1/2}A^{-1/2}}{2}| \geq |A^{-1/2}B^{1/2}|$. The latter then follows from the real inequality $\frac{1+\lambda^2}{2} \geq \lambda$ for the singular values λ of $A^{-1/2}B^{1/2}$.

Acknowledgements

We would like to thank the referees for their comments and suggestions on the manuscript. A.K. wishes to thank Mikael Lindström for discussions and the Academy of Finland (grant #339515) for financial support.

References

[1] L. V. Kantorovich, G. Akilov, Functional Analysis, Pergamon Press; Oxford, UK, 1982.

[2] M. Schechter, Principles of Functional Analysis, American Mathematical Society, Providence, 2002.

[3] G. Brown, Convexity and Minkowski's inequality, Amer. Math. Monthly 112(8) (2005) 740–742.

[4] W. Rudin, Functional Analysis, Tata McGraw-Hill; New Delhi, India, 1978.

[5] O. Güler, Foundations of Optimization, Graduate Texts in Mathematics 258; Springer, New York, 2010.

[6] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press; Cambridge, UK, 2004.

[7] R. A. Horn, C. R. Johnson, Matrix Analysis, Cambridge University Press; Cambridge, UK, 1985.