Schedae Informaticae Vol. 21 (2012): 55–63
doi: 10.4467/20838476SI.12.004.0814

Twice Differentiable Characterizations of Convexity Notions for Functions on Full Dimensional Convex Sets

Oliver Stein
Institute of Operations Research, Karlsruhe Institute of Technology (KIT),
Kaiserstr. 12, 76131 Karlsruhe, Germany
e-mail: *stein@kit.edu*

Abstract. We derive C^2-characterizations for convex, strictly convex, as well as strongly convex functions on full dimensional convex sets. In the cases of convex and strongly convex functions this weakens the well-known openness assumption on the convex sets. We also show that, in a certain sense, the full dimensionality assumption cannot be weakened further. In the case of strictly convex functions we weaken the well-known sufficient C^2-condition for strict convexity to a characterization. Several examples illustrate the results.

Keywords: convexity, strict convexity, strong convexity, true convexity, differentiable characterization, full dimensional convex set.

1. Introduction

This article derives C^2-characterizations for three basic convexity notions of functions on full dimensional convex sets. Recall (e.g. from [1, 3]) that for a nonempty convex set $M \subset \mathbb{R}^n$ a function $f : M \to \mathbb{R}$ is called *convex* if

$$\text{for all } \; x, y \in M, \; \lambda \in (0,1) : \; f(\,(1-\lambda)x + \lambda y\,) \;\leq\; (1-\lambda)f(x) + \lambda f(y) \qquad (1)$$

holds, *strictly convex* if the inequality in (1) is strict for $x \neq y$, and *strongly* (or *uniformly*) *convex* if $f(x) - \frac{c}{2}x^\intercal x$ is convex for some $c > 0$.

In the framework of function minimization, these convexity notions have important and well-known implications. In fact, let $M \subset \mathbb{R}^n$ be nonempty and convex.

Then for a convex function f on M the set of minimal points is convex (but possibly empty), for a strictly convex function f on M the set of minimal points contains at most one element, and for a strongly convex function f on M the set of minimal points contains exactly one element, if M is also closed (but not necessarily bounded).

For functions f which are twice continuously differentiable on some open set containing M (briefly, $f \in C^2(M, \mathbb{R})$), also the following relations of these concepts to properties of the Hessians $D^2 f(x)$ of f on M are well-known.

Theorem 1 (e.g. [1]) *Let $M \subset \mathbb{R}^n$ be nonempty and convex, and let $f \in C^2(M, \mathbb{R})$. Then the following assertions hold.*

(i) If $D^2 f(x)$ is positive semi-definite (briefly: $D^2 f(x) \succeq 0$) for all $x \in M$, then f is convex on M. If M is open, also the converse direction holds.

(ii) If $D^2 f(x)$ is positive definite (briefly: $D^2 f(x) \succ 0$) for all $x \in M$, then f is strictly convex on M.

(iii) If there exists some $c > 0$ such that the smallest eigenvalue of $D^2 f(x)$ is bounded below by c (briefly: $\lambda_{\min}(D^2 f(x)) \geq c$) on M, then f is strongly convex on M. If M is open, also the converse direction holds.

As opposed to the C^2-conditions for convexity and strong convexity, the C^2-condition for strict convexity in Theorem 1(ii) is only sufficient but not necessary, even on open sets M. The latter is illustrated by the function $f(x) = x^4$ which is strictly convex on $\mathbb{R}$, but $f''(0)$ vanishes. This discrepancy between the two function classes is the motivation to introduce the following terminology (see also [5]).

Definition 1 *For a nonempty convex set $M \subset \mathbb{R}^n$ we call $f \in C^2(M, \mathbb{R})$* truly convex *on M if $D^2 f(x)$ is positive definite for all $x \in M$.*

For C^2-functions f on M we have the obvious implications

$$\text{strongly convex} \quad \Rightarrow \quad \text{truly convex} \quad \Rightarrow \quad \text{strictly convex} \quad \Rightarrow \quad \text{convex},$$

and on $\mathbb{R}$, for example, $f(x) = x^2$ is strongly convex, $f(x) = \exp(x)$ is truly, but not strongly convex, $f(x) = x^4$ is strictly, but not truly convex, and $f(x) = x$ is convex, but not strictly convex.

The contribution of this paper is twofold. In Section 2, for the C^2-characterizations of convexity and strong convexity, we weaken the openness assumption on convex sets to their full dimensionality. We also show that, in a certain sense, the full dimensionality assumption cannot be weakened further.

In the subsequent sections we contribute C^2-characterizations of strict convexity on full dimensional convex sets. Section 3 prepares these results by a characterization of strict convexity on line segments which does not need full dimensionality or smoothness assumptions. Section 4 deals with univariate functions, while Section 5 treats the multivariate case. We conclude this paper with some final remarks in Section 6.

2. Twice differentiable characterizations of convexity and strong convexity on full dimensional convex sets

The converse direction in the C^2-characterization of convexity from Theorem 1(i) says that, if $f \in C^2(M, \mathbb{R})$ is convex on some nonempty *open* convex set $M \subset \mathbb{R}^n$, then $D^2 f(x) \succeq 0$ holds for all $x \in M$. In fact, the well-known example $f(x) = x_1^2 - x_2^2$ on $M = \mathbb{R} \times \{0\}$ shows that f may be convex on a non-open set M, while $D^2 f(x)$ is indefinite, even for all $x \in M$.

However, for practical applications, in particular in optimization, the openness assumption does not directly cover many relevant situations in which M is a closed convex proper subset of $\mathbb{R}^n$. Fortunately, the openness assumption can be weakened to *full dimensionality* of M, as we will see in the subsequent results. Recall that a convex set $M \subset \mathbb{R}^n$ is called *full dimensional* if its affine hull $\mathrm{aff}(M)$ has dimension n, that $\mathrm{aff}(M)$ is the smallest affine set containing M, and that its dimension is the dimension of the parallel linear space. Note that, here and in the following, we shall often assume that f is twice continuously differentiable on some open set containing the topological closure $\mathrm{cl}(M)$ of M which, for a non-closed set M, is a slightly stronger assumption than $f \in C^2(M, \mathbb{R})$.

Lemma 1 *Let $f \in C^2(\mathrm{cl}(M), \mathbb{R})$ be convex on some full dimensional convex set $M \subset \mathbb{R}^n$. Then $D^2 f(x) \succeq 0$ holds for all $x \in \mathrm{cl}(M)$.*

Proof. By [3, Th. 6.3], $\mathrm{cl}(M)$ coincides with $\mathrm{cl}(\mathrm{int}(M))$ where $\mathrm{int}(M)$, the topological interior of M, is nonempty and convex in view of [3, Th. 6.2]. As f is convex on the nonempty open convex set $\mathrm{int}(M)$, Theorem 1(i) implies $D^2 f(x) \succeq 0$ for all $x \in \mathrm{int}(M)$. The continuity of the eigenvalues of the symmetric matrix $D^2 f(x)$ immediately implies $D^2 f(x) \succeq 0$ for all $x \in \mathrm{cl}(\mathrm{int}(M)) = \mathrm{cl}(M)$. •

Due to $M \subset \mathrm{cl}(M)$ for any set $M \subset \mathbb{R}^n$, the combination of Theorem 1(i) and Lemma 1 yields the following theorem.

Theorem 2 (C^2-characterization of convexity)
Let $M \subset \mathbb{R}^n$ be a full dimensional convex set. Then $f \in C^2(\mathrm{cl}(M), \mathbb{R})$ is convex if and only if $D^2 f(x) \succeq 0$ holds for all $x \in M$.

Note that the set $M = \mathbb{R} \times \{0\}$, on which the function $f(x) = x_1^2 - x_2^2$ with indefinite Hessian is convex, is not full dimensional.

In contrast to the openness assumption on M, its full dimensionality cannot be weakened further, when arbitrary convex C^2 functions shall possess positive semi-definite Hessians on M. This follows from the subsequent result, which generalizes the above example to arbitrary 'flat' convex sets.

Theorem 3 *Let $M \subset \mathbb{R}^n$ be nonempty and convex with $\dim(M) < n$. Then there exists a function $f \in C^2(M, \mathbb{R})$ which is convex on M but has an indefinite Hessian $D^2 f(x)$ for all $x \in M$.*

Proof. The assertion is trivial in the case when M is a singleton. Hence, in the following we will assume $1 \leq m < n$ for $m = \dim(M)$. Let $u \in M$, let the columns of the $(n, m)-$matrix V form an orthonormal basis of $\mathrm{aff}(M) - u$, and let the columns of the $(n, n-m)-$matrix W form an orthonormal basis of $(\mathrm{aff}(M) - u)^{\perp}$. Then the function

$$f(x) \;=\; \tfrac{1}{2}(x-u)^{\intercal}\,(VV^{\intercal} - WW^{\intercal})\,(x-u)$$

possesses the Hessian $D^2 f(x) = VV^{\intercal} - WW^{\intercal}$ which satisfies

$$W^{\intercal} D^2 f(x) W \;=\; W^{\intercal} V V^{\intercal} W - W^{\intercal} W W^{\intercal} W \;=\; -I,$$

so that $D^2 f(x)$ is negative definite on $\mathrm{span}(W)$. Analogously, it is not hard to see that $D^2 f(x)$ is positive definite on $\mathrm{span}(V) = \mathrm{span}(W)^{\perp}$, so that $D^2 f(x)$ is indefinite for each $x \in \mathbb{R}^n$.

On the other hand, each $x \in M$ is an element of $\mathrm{aff}(M)$ and can, thus, be written in the form $x = u + Vy$ with some $y \in \mathbb{R}^m$. In fact, the affine transformation $x = T(y, z) = u + Vy + Wz$ with $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^{n-m}$ satisfies $T(\mathbb{R}^m \times \{0_{n-m}\}) = \mathrm{aff}(M)$. For the transformed function $f \circ T$ and all $y \in \mathbb{R}^m$ we obtain

$$(f \circ T)(y, 0) \;=\; f(u + Vy) \;=\; \tfrac{1}{2} y^{\intercal}\,(V^{\intercal} V V^{\intercal} V - V^{\intercal} W W^{\intercal} V)\,y \;=\; \tfrac{1}{2} y^{\intercal} y,$$

so that $f \circ T$ is convex on $\mathbb{R}^m \times \{0_{n-m}\}$. Consequently, f is convex on $\mathrm{aff}(M)$ and, thus, on M. •

We close this section with a simple consequence of Theorems 1(iii) and 2 for the C^2-characterization of strongly convex functions.

Theorem 4 (C^2-characterization of strong convexity)
Let $M \subset \mathbb{R}^n$ be a full dimensional convex set. Then $f \in C^2(\mathrm{cl}(M), \mathbb{R})$ is strongly convex on M if and only if there exists some $c > 0$ such that $\lambda_{\min}(D^2 f(x)) \geq c$ holds for all $x \in M$.

Proof. In view of Theorem 1(iii) we only have to show that for a strongly convex function $f \in C^2(\mathrm{cl}(M), \mathbb{R})$ there exists some $c > 0$ such that $\lambda_{\min}(D^2 f(x)) \geq c$ holds for all $x \in M$. In fact, for a strongly convex function f on M there exists some $c > 0$ such that $f(x) - \frac{c}{2} x^{\intercal} x$ is convex on M. By Theorem 2, the Hessian $D^2(f(x) - \frac{c}{2} x^{\intercal} x) = D^2 f(x) - cI$ is positive semi-definite for all $x \in M$, which implies the assertion. •

3. A characterization of strict convexity on line segments

The following characterization of strict convexity does not depend on any smoothness or full dimensionality assumptions and seems obvious. As the proof is not straightforward, we include it for the sake of completeness. In the following, for a convex set $M \subset \mathbb{R}^n$ and $x, y \in M$

$$[x, y] \ := \ \{(1-\lambda)x + \lambda y| \ \lambda \in [0,1]\}$$

will denote a (closed) *line segment* in M. In the case $x \neq y$ we call it *line segment of positive length.*

Condition A. *f is convex on M, and f is not linear on any line segment of positive length in M.*

Lemma 2 *Let $M \subset \mathbb{R}^n$ be nonempty and convex. Then f is strictly convex on M if and only if Condition A holds.*

Proof. Clearly, f is not strictly convex on M if Condition A is violated. On the other hand, let f be convex, but not strictly convex on M. Then there are points $\bar{x}, \bar{y} \in M$, $\bar{x} \neq \bar{y}$, as well as some $\bar{\lambda} \in (0,1)$ with

$$f(\,(1-\bar{\lambda})\bar{x} + \bar{\lambda}\bar{y}\,) \ = \ (1-\bar{\lambda})f(\bar{x}) + \bar{\lambda} f(\bar{y}).$$

Assume that there exists some $\hat{\lambda} \in (0,1)$ with

$$f(\,(1-\hat{\lambda})\bar{x} + \hat{\lambda}\bar{y}\,) \ < \ (1-\hat{\lambda})f(\bar{x}) + \hat{\lambda} f(\bar{y}).$$

Without loss of generality, let $\hat{\lambda} < \bar{\lambda}$. Then

$$\mu \ = \ \frac{\bar{\lambda} - \hat{\lambda}}{1 - \hat{\lambda}}$$

is an element of $(0,1)$. With $\bar{z} = (1-\hat{\lambda})\bar{x} + \hat{\lambda}\bar{y}$ we obtain the contradiction

$$\begin{aligned}
(1-\bar{\lambda})f(\bar{x}) + \bar{\lambda} f(\bar{y}) \ &= \ f(\,(1-\bar{\lambda})\bar{x} + \bar{\lambda}\bar{y}\,) \\
&= \ f(\,(1-\mu)(1-\hat{\lambda})\bar{x} + ((1-\mu)\hat{\lambda} + \mu)\bar{y}\,) \\
&= \ f(\,(1-\mu)\bar{z} + \mu\bar{y}\,) \\
&\leq \ (1-\mu)f(\bar{z}) + \mu f(\bar{y}) \\
&< \ (1-\mu)\left[(1-\hat{\lambda})f(\bar{x}) + \hat{\lambda} f(\bar{y})\right] + \mu f(\bar{y}) \\
&= \ (1-\bar{\lambda})f(\bar{x}) + \bar{\lambda} f(\bar{y}).
\end{aligned}$$

Consequently we have

$$\text{for all } \ \lambda \in [0,1] : \ f(\,(1-\lambda)\bar{x} + \lambda\bar{y}\,) \ = \ (1-\lambda)f(\bar{x}) + \lambda f(\bar{y})$$

so that the graph of f on $[\bar{x}, \bar{y}]$ is the line segment of positive length

$$[(\bar{x}, f(\bar{x})), (\bar{y}, f(\bar{y}))].$$

This shows the assertion. •

4. Strict convexity in the univariate case

For $n = 1$ the following C^2-characterization of strict convexity is inspired by a C^1−characterization of strict monotonicity in [2, IV.24 Folgerung(b)].

Condition B. *The function f'' is nonnegative on M, and the set $\{x \in M|\ f''(x) = 0\}$ does not contain interior points.*

Theorem 5 (C^2-char. of strict convexity, univariate case)
Let the convex set $M \subset \mathbb{R}^1$ contain at least two distinct elements, and let $f \in C^2(\mathrm{cl}(M), \mathbb{R})$. Then f is strictly convex on M if and only if Condition B holds.

Proof. We use Lemma 2 and show the equivalence of Conditions A and B. By Theorem 2, the convexity of f on M is equivalent to the nonnegativity of f'' on M, as a one dimensional convex set with at least two distinct elements is full dimensional. Furthermore, the linearity of f on some line segment of positive length $[\bar{x}, \bar{y}] \subset M$ is equivalent to $[\bar{x}, \bar{y}] \subset \{x \in M|\ f''(x) = 0\}$. This shows the assertion. •

Example 1 *The function $f(x) = x^4$ is not truly convex on $\mathbb{R}$. However, it satisfies Condition B, so that its strict convexity follows from Theorem 5.*

Example 2 *With*

$$g(x) = \begin{cases} (x+1)^4, & x < -1 \\ 0, & x \in [-1, 1] \\ (x-1)^4, & x > 1 \end{cases}$$

and any $a, b \in \mathbb{R}$ the function $f(x) = g(x) + ax + b$ is C^2, convex, but not strictly convex on $\mathbb{R}$. Moreover, it violates Condition B.

5. Strict convexity in the multivariate case

For $n > 1$ the generalization of Condition B is not straightforward. In view of Condition A, at least it is obvious that the interior point property from Condition B refers to line segments of positive length in the multivariate case, as in the following condition.

Condition C. *The Hessian $D^2 f$ is positive semi-definite on M, and the set $\{x \in M|\ D^2 f(x) \text{ is singular}\}$ does not contain any line segment of positive length.*

Condition C is weaker than true convexity on M, since for any truly convex function f the set $\{x \in M|\ D^2 f(x) \text{ is singular}\}$ is empty. The following theorem shows, however, that Condition C is sufficiently strong to imply strict convexity of f on M.

Theorem 6 *Let $M \subset \mathbb{R}^n$ be nonempty and convex (but not necessarily full dimensional), let $f \in C^2(\mathrm{cl}(M), \mathbb{R})$, and let Condition C be satisfied. Then f is strictly convex on M.*

Proof. We use Lemma 2 and show that Condition C implies Condition A. In fact, under Condition C the convexity of f on M follows from Theorem 1(i). Assume that Condition A is violated. Then there exists a line segment $[\bar{x}, \bar{y}] \subset M$ of positive length such that the function

$$\varphi(\lambda) \;:=\; f(\bar{x} + \lambda(\bar{y} - \bar{x})) \tag{2}$$

is linear on $[0, 1]$. Hence, its second derivative

$$\varphi''(\lambda) \;=\; (\bar{y} - \bar{x})^{\intercal} D^2 f(\bar{x} + \lambda(\bar{y} - \bar{x}))(\bar{y} - \bar{x}) \tag{3}$$

vanishes on $[0, 1]$. As f is convex on M, we have $D^2 f(\bar{x} + \lambda(\bar{y} - \bar{x})) \succeq 0$ for all $\lambda \in [0, 1]$. If $D^2 f(\bar{x} + \bar{\lambda}(\bar{y} - \bar{x}))$ was nonsingular for some $\bar{\lambda} \in [0, 1]$, this would mean $D^2 f(\bar{x} + \bar{\lambda}(\bar{y} - \bar{x})) \succ 0$, so that $\bar{y} - \bar{x} \neq 0$ leads to the contradiction $\varphi''(\bar{\lambda}) > 0$. Consequently, $D^2 f$ is singular on $[\bar{x}, \bar{y}]$, and Condition C is violated. •

Example 3 *Consider $f(x) = x_1^4 + x_2^2$ on $M(t) = \{x \in \mathbb{R}^2 |\ (x_1 - t)^2 + x_2^2 \leq 1\}$ with $t \in \mathbb{R}$. For all $t < -1$ as well as all $t > 1$ the function f is truly and, thus, strictly convex on $M(t)$. For $t \in \{\pm 1\}$ the function f is not truly convex on $M(t)$, as $D^2 f(0)$ is singular. However, since $D^2 f(x) \succ 0$ holds for all $x \in M(t) \setminus \{0\}$, Condition C is satisfied and, by Theorem 6, f is strictly convex on $M(t)$.*

Unfortunately, Condition C is too strong to characterize strict convexity, as the next example illustrates.

Example 4 *For f and $M(t)$ from Example 3 and $t \in (-1, 1)$ Condition C is violated as the set*

$$\{x \in \mathbb{R}^2 |\ D^2 f(x) \text{ is singular}\} \;=\; \{x \in \mathbb{R}^2 |\ x_1 = 0\}$$

contains a line segment of positive length in $M(t)$. However, it is not hard to verify Condition A (see also Example 5 below), so that f is strictly convex on $M(t)$.

In the following Condition D, for $v \in \mathbb{R}^n \setminus \{0\}$ we denote by $A|_{\mathrm{span}(v)}$ the restriction of the symmetric $(n, n)-$matrix A to the one dimensional space $\mathrm{span}(v)$, that is, the scalar $v^{\intercal} A v$.

Condition D. *The Hessian $D^2 f$ is positive semi-definite on M, and there is no line segment $[\bar{x}, \bar{y}]$ of positive length in M such that $D^2 f(x)|_{\mathrm{span}(\bar{y} - \bar{x})}$ vanishes for all $x \in [\bar{x}, \bar{y}]$.*

Theorem 7 (C^2-char. of strict convexity, multivariate case)
Let $M \subset \mathbb{R}^n$ be a full dimensional convex set, and let $f \in C^2(\mathrm{cl}(M), \mathbb{R})$. Then f is strictly convex on M if and only if Condition D holds.

Table 1. C^2-characterizations of convexity notions on full dimensional convex sets

convexity notion	C^2-characterization
f convex on M	$D^2 f \succeq 0$ on M
f strictly convex on M	Condition D
f truly convex on M	$D^2 f \succ 0$ on M
f strongly convex on M	$\lambda_{\min}(D^2 f) \geq c > 0$ on M

Proof. We use Lemma 2 and show the equivalence of Conditions A and D. By Theorem 2, the convexity of f on M is equivalent to $D^2 f \succeq 0$ on M.

For any line segment $[\bar{x}, \bar{y}] \subset M$ of positive length, the linearity of f on $[\bar{x}, \bar{y}]$ is characterized by $\varphi'' = 0$ on $[0,1]$, with φ from (2). In view of (3), the latter is equivalent to $D^2 f(x)|_{\text{span}(\bar{y}-\bar{x})} = 0$ for all $x \in [\bar{x}, \bar{y}]$, so that the equivalence of Conditions A and D is shown. •

Note that Condition D collapses to Condition B in the case $n = 1$.

Example 5 *Consider once more the function f on $M(t)$ from Example 3 with $t \in (-1, 1)$. In the following we will show that it satisfies Condition D. In fact, assume that there is some line segment $[\bar{x}, \bar{y}] \subset M(t)$ such that*

$$D^2 f(x)|_{\text{span}(\bar{y}-\bar{x})} \;=\; (\bar{y}-\bar{x})^{\mathsf{T}} \begin{pmatrix} 12x_1^2 & 0 \\ 0 & 2 \end{pmatrix} (\bar{y}-\bar{x})$$

vanishes for all $x \in [\bar{x}, \bar{y}]$. In view of $D^2 f(x) \succeq 0$ and $(\bar{y}-\bar{x}) \neq 0$ this can only happen for singular matrices $D^2 f(x)$, that is, for $x_1 = 0$. Hence we necessarily have $[\bar{x}, \bar{y}] \subset \{x \in \mathbb{R}^2 |\ x_1 = 0\}$, $\text{span}(\bar{y}-\bar{x}) = \{0\} \times \mathbb{R}$ and, thus, $D^2 f(x)|_{\text{span}(\bar{y}-\bar{x})} > 0$. This contradiction shows that Condition D is satisfied.

6. Final remarks

In this paper we derived C^2-characterizations of plain, strict, as well as strong convexity on full dimensional convex sets. Table 1 summarizes these C^2-characterizations, where the characterization of true convexity actually just repeats its definition.

Note that the characterization of strict convexity on full dimensional convex sets simplifies to Condition B in the univariate case, so that Theorem 5 clarifies a remark after [4, Th. 2.13]. Condition C yields a weaker sufficient condition for strict convexity than true convexity, even without the full dimensionality assumption.

We finally point out that, in optimization, the set M often is given in functional form as

$$M \;=\; \{x \in \mathbb{R}^n |\ g_i(x) \leq 0,\ i \in I,\ h_j(x) = 0,\ j \in J\}$$

with finite index sets I and J, convex functions $g_i : \mathbb{R}^n \to \mathbb{R}$, $i \in I$, and affine functions $h_j : \mathbb{R}^n \to \mathbb{R}$, $j \in J$. A simple sufficient condition for the full dimensionality of M then is $J = \emptyset$ and the existence of a Slater point $x^\star$ in the sense that $g_i(x^\star) < 0$ holds for all $i \in I$.

Acknowledgements. The author is grateful to two anonymous referees whose comments lead to a substantially improved version of this article.

7. References

[1] Hiriart-Urruty J.-B., Lemaréchal C.; *Fundamentals of Convex Analysis*, Springer 2001.

[2] Jongen H.Th., Schmidt P.G.; *Analysis*, Lecture Notes, Aachener Beiträge zur Mathematik 19, Wissenschaftsverlag, Mainz 1998.

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

[4] Rockafellar R.T., Wets R.; *Variational Analysis*, Springer 1998.

[5] Shikhman V., Stein O.; *On jet convex functions and their tensor products*, Optimization 61, 2012, pp. 717–731.