

CHARACTERIZATIONS OF UNIFORM CONVEXITY FOR DIFFERENTIABLE FUNCTIONS

*Dedicated to Academician Professor Gradimir Milovanović
on the occasion of his 70th birthday.*

*Osama Alabdali, Allal Guessab and Gerhard Schmeisser**

We consider convex functions in d real variables. For applications, for example in optimization, various strengthened forms of convexity have been introduced. Among them, *uniform convexity* is one of the most general, defined by involving a so-called modulus ϕ . Inspired by three classical characterizations of ordinary convexity, we aim at characterizations of uniform convexity by conditions in terms of the gradient or the Hessian matrix of the considered function for certain classes of moduli ϕ .

1. INTRODUCTION

Let d be a positive integer. By $\langle \cdot, \cdot \rangle$, we denote the standard inner product in \mathbb{R}^d and by $\| \cdot \|$ the Euklidean norm, i.e., $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$ for $\mathbf{x} \in \mathbb{R}^d$. Furthermore, we use the notation $\mathbb{R}_+ := [0, +\infty[$.

1.1 Characterizations of convexity

We start by recalling the following standard definition: Let C be a nonempty, open, convex subset of \mathbb{R}^d . A function $f : C \rightarrow \mathbb{R}$ is said to be *convex* on C if for any

*Corresponding author. Gerhard Schmeisser

2010 Mathematics Subject Classification. 26B25.

Keywords and Phrases. Convexity, uniform convexity, characterizations for differentiable functions.

$\mathbf{x}, \mathbf{y} \in C$ and all $t \in [0, 1]$, we have

$$(1) \quad f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}).$$

For differentiable functions f , there exist the following characterizations of convexity which are very convenient in applications. For references see, e.g., [1, §17.2] or [2, §3.1.3–§3.1.4].

Theorem A [First order characterization] *Let $C \subset \mathbb{R}^d$ be a nonempty, open, convex set and let $f : C \rightarrow \mathbb{R}$ be a differentiable function. Then f is convex on C if and only if for any $\mathbf{x}, \mathbf{y} \in C$, we have*

$$(2) \quad f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle.$$

Another characterization of convexity in terms of ∇f uses a generalized notion of monotonicity. It may be introduced as follows.

Definition 1.1. *Let $C \subset \mathbb{R}^d$ be a nonempty, open, convex set and let $G : C \rightarrow \mathbb{R}^d$. Then G is said to be monotone on C if for any $\mathbf{x}, \mathbf{y} \in C$, we have*

$$(3) \quad \langle G(\mathbf{y}) - G(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq 0.$$

Theorem B [Monotone gradient characterization] *Let $C \subset \mathbb{R}^d$ be a nonempty, open, convex set and let $f : C \rightarrow \mathbb{R}$ be a continuously differentiable function. Then f is convex on C if and only if ∇f is monotone on C .*

Finally we mention a characterization of convexity in terms of $\nabla^2 f(\mathbf{x})$ which is the Hessian matrix of f at \mathbf{x} .

Theorem C [Second order characterization] *Let $C \subset \mathbb{R}^d$ be a nonempty, open, convex set and let $f : C \rightarrow \mathbb{R}$ be a twice continuously differentiable function. Then f is convex on C if and only if $\nabla^2 f(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in C$.*

1.2 Refinements of convexity

In applications, in particular in optimization, several strengthened forms of convexity have gained attention. Three of them may be introduced as follows; see, e.g., [1, §8.1, §10.2] or [2, §3.1, §9.1.2].

Definition 1.2. *Let C be as in the previous section. A function $f : C \rightarrow \mathbb{R}$ is said to be*

- (i) strictly convex if strict inequality holds in (1) whenever $\mathbf{x} \neq \mathbf{y}$ and $0 < t < 1$;
- (ii) strongly convex with constant $\mu > 0$ if $f - (\mu/2)\|\cdot\|^2$ is convex;
- (iii) uniformly convex with modulus ϕ if $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is an increasing function vanishing only at 0 such that

$$(4) \quad f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}) - t(1-t)\phi(\|\mathbf{y} - \mathbf{x}\|)$$

for any $\mathbf{x}, \mathbf{y} \in C$ and all $t \in [0, 1]$.

In (iii) the attribute *uniformly* is justified by the fact that for fixed t the term in (4) that strengthens (1) does not depend on the special location of the points $\mathbf{x}, \mathbf{y} \in C$ but only on their distance.

It is easily verified that a function is strongly convex with constant μ if and only if it is uniformly convex with modulus $\phi = (\mu/2)(\cdot)^2$; also see [1, Proposition 10.6]. As such, strong convexity is a special case of uniform convexity. The following implications are obvious:

$$\text{strong convexity} \implies \text{uniform convexity} \implies \text{strict convexity} \implies \text{convexity}$$

It is known that none of these implications is reversible. For the reader's convenience, we show this by examples. It follows from [1, Example 10.14] that $\|\cdot\|^4$ is uniformly convex on \mathbb{R}^d . Now assume that it is also strongly convex. Then there exists $\mu > 0$ such that $g := \|\cdot\|^4 - (\mu/2)\|\cdot\|^2$ is convex. However, for

$$\mathbf{x} := (-\sqrt{\mu}/2, 0, \dots, 0), \mathbf{y} := (\sqrt{\mu}/2, 0, \dots, 0) \in \mathbb{R}^d,$$

we have

$$g(\mathbf{x}) = g(\mathbf{y}) = -\frac{\mu^2}{16} < 0 \quad \text{and} \quad g\left(\frac{\mathbf{x} + \mathbf{y}}{2}\right) = 0$$

contradicting convexity of g . Hence the first implication is not reversible. Next, take $f := \|\cdot\|^p$ with $p \in]1, 2[$. It is known to be strictly convex on \mathbb{R}^d ; see [1, Example 8.21]. For positive integers n consider the points

$$\mathbf{x}_n := (n - 1, 0, \dots, 0), \mathbf{y}_n := (n + 1, 0, \dots, 0) \in \mathbb{R}^d$$

and choose $t = 1/2$. If f were uniformly convex, we would have

$$f\left(\frac{\mathbf{x}_n + \mathbf{y}_n}{2}\right) \leq \frac{f(\mathbf{x}_n) + f(\mathbf{y}_n)}{2} - \frac{1}{4}\phi(\|\mathbf{x}_n - \mathbf{y}_n\|)$$

and so

$$n^p \leq \frac{1}{2}[(n - 1)^p + (n + 1)^p] - \frac{1}{4}\phi(2).$$

This implies that

$$\frac{1}{2}\phi(2) \leq n^p \left[\left(1 - \frac{1}{n}\right)^p + \left(1 + \frac{1}{n}\right)^p - 2 \right].$$

The left-hand side is a positive constant while the right-hand side approaches 0 for $n \rightarrow \infty$ as is seen by Taylor expansion with respect to $1/n$. This is a contradiction. Thus the second implication is not reversible. Finally, it is obvious that an affine function is convex but not strictly convex, and so the third implication is not reversible as well.

It is easy to characterize strong convexity analogous to the statements of § 1.1. We just have to apply Theorems A, B and C to the function $g := f - (\mu/2)\|\cdot\|^2$ and rewrite the resulting criteria in terms of f .

In the case of strict convexity, there exist simple analogues of the results in terms of the gradient; see, e.g., [1, Proposition 17.13], [2, §3.1]. As regards Theorem A, we just have to require strict inequality in (2) whenever $\mathbf{x} \neq \mathbf{y}$. For an analogue of Theorem B, we need strict monotonicity of ∇f , which is defined by strict inequality in (3) whenever $\mathbf{x} \neq \mathbf{y}$. However, no characterization analogous to that of Theorem C is known. One might think of requiring $\nabla^2 f(\mathbf{x})$ to be positive definite for all $\mathbf{x} \in C$, but this is a sufficient condition only. It is not necessary as the example $f := \|\cdot\|^4$ shows. This function is strictly convex on \mathbb{R}^d , but its Hessian matrix $\nabla^2 f(\mathbf{x})$ is not positive definite at $\mathbf{x} = 0$.

It seems that characterizations in the spirit of Theorems A, B and C have not yet been given for uniformly convex functions. The purpose of this article are attempts to fill this gap, at least for certain classes of moduli ϕ which include $\phi = (\mu/2)(\cdot)^2$, and so strong convexity is covered.

2. THE RESULTS AND THEIR PROOFS

First we want to show that for any admissible modulus ϕ , uniform convexity implies a strengthened first order condition in case of differentiable functions. Since we shall use this result repeatedly, we state it as a lemma.

Lemma 2.1. *Let $C \subset \mathbb{R}^d$ be a nonempty, open, convex subset and let $f : C \rightarrow \mathbb{R}$ be differentiable. If f is uniformly convex on C with modulus ϕ , then, for any $\mathbf{x}, \mathbf{y} \in C$, there holds*

$$(5) \quad f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \phi(\|\mathbf{y} - \mathbf{x}\|).$$

Proof. If f is uniformly convex with modulus ϕ , then (4) holds. For $t \in [0, 1]$, we may rewrite this inequality as

$$f(\mathbf{y}) - f(\mathbf{x}) \geq \frac{f(t\mathbf{x} + (1-t)\mathbf{y}) - f(\mathbf{x})}{1-t} + t\phi(\|\mathbf{y} - \mathbf{x}\|).$$

Since f is differentiable at \mathbf{x} , the quotient on the right-hand side approaches $\langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle$ as $t \rightarrow 1$. Thus (5) holds. \square

We cannot expect that (5) will imply (4) for *any* admissible modulus ϕ . From now on, we therefore want to restrict the moduli under consideration by imposing an additional condition, which is realistic in applications.

For $p > 0$, we denote by \mathcal{U}_p the class of all functions ϕ with the following properties:

- (i) $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is increasing and vanishes at 0 only;
- (ii) for any $r \in \mathbb{R}_+$ and $\varepsilon \in [0, 1]$, we have $\phi(\varepsilon r) \geq \varepsilon^p \phi(r)$.

Obviously, if $\phi \in \mathcal{U}_p$, then $c\phi \in \mathcal{U}_p$ for each $c > 0$. Also note that if $\phi \in \mathcal{U}_p$, then it follows from (ii) that $\phi(\varepsilon r) \geq \varepsilon^q \phi(r)$ for all $q \geq p$, $r \in \mathbb{R}_+$ and $\varepsilon \in [0, 1]$. Hence $\mathcal{U}_p \subseteq \mathcal{U}_q$ for $0 < p \leq q$. The example $\phi = (\cdot)^q$ shows that

$$\mathcal{U}_p \subsetneq \mathcal{U}_q \quad \text{for } 0 < p < q.$$

Finally we mention that the function $\phi = (\mu/2)(\cdot)^2$ with $\mu > 0$, which occurs in connection with strong convexity as a special case of uniform convexity, belongs to \mathcal{U}_p for all $p \geq 2$.

One may ask if for each $p > 0$ the class \mathcal{U}_p contains a function ϕ which is modulus of some uniformly convex function f . The answer is no. The following result shows that we have to restrict p additionally.

Proposition 2.2. *Let C be as in Lemma 2.1 and let $f : C \rightarrow \mathbb{R}$ be uniformly convex with modulus $\phi \in \mathcal{U}_p$. Then necessarily $p \geq 2$.*

Proof. Set

$$Q(t, \mathbf{x}, \mathbf{y}) := \frac{tf(\mathbf{x}) + (1-t)f(\mathbf{y}) - f(t\mathbf{x} + (1-t)\mathbf{y})}{t(1-t)}$$

and define the function

$$\varphi : \begin{cases} \mathbb{R}_+ & \longrightarrow \mathbb{R}_+, \\ r & \longmapsto \inf\{Q(t, \mathbf{x}, \mathbf{y}) : \mathbf{x}, \mathbf{y} \in C, t \in]0, 1[, \|\mathbf{y} - \mathbf{x}\| = r\}. \end{cases}$$

Clearly, if ϕ is a modulus of uniform convexity of f , we must have

$$(6) \quad \phi(r) \leq \varphi(r) \quad \text{for all } r \in \mathbb{R}_+.$$

It is shown in [1, Proposition 10.10] that $\varphi(0) = 0$ and

$$\varphi(\gamma r) \geq \gamma^2 \varphi(r) \quad \text{for all } r \in \mathbb{R}_+, \gamma \in [1, \infty[.$$

This inequality is equivalent to

$$(7) \quad \varphi(\varepsilon r) \leq \varepsilon^2 \varphi(r) \quad \text{for all } r \in \mathbb{R}_+, \varepsilon \in]0, 1].$$

The properties of \mathcal{U}_p in conjunction with (6) and (7) yield

$$\varepsilon^p \phi(r) \leq \phi(\varepsilon r) \leq \varphi(\varepsilon r) \leq \varepsilon^2 \varphi(r),$$

and so $\phi(r) \leq \varepsilon^{2-p} \varphi(r)$.

Now assume that $0 < p < 2$. Then, keeping $r > 0$ fixed and letting $\varepsilon \rightarrow 0$, we deduce that $\phi(r) = 0$. This is a contradiction since ϕ vanishes at 0 only. \square

Next we want to show by examples that for $p \geq 2$, there do exist functions $\phi \in \mathcal{U}_p$ which are modulus of a uniformly convex function f . First consider the function $f_2 = \|\cdot\|^2$ defined on \mathbb{R}^d . A simple calculation yields

$$tf_2(\mathbf{x}) + (1-t)f_2(\mathbf{y}) - f_2(t\mathbf{x} + (1-t)\mathbf{y}) = t(1-t)\|\mathbf{y} - \mathbf{x}\|^2.$$

This shows that f_2 is uniformly convex with modulus $\phi = (\cdot)^2$. In fact, this is the largest modulus for f_2 since equality is attained in (4) everywhere. Here, obviously, $\phi \in \mathcal{U}_p$ for each $p \geq 2$.

The previous example may be considered as a trivial case. One may ask if for $p > 2$ there exists a $\phi \in \mathcal{U}_p$, which does not belong to any \mathcal{U}_q with $2 \leq q < p$ and is modulus of uniform convexity of some function f . The answer is affirmative. Indeed, consider the function $f_p = \|\cdot\|^p = f_2^{p/2}$ for $p > 2$. It can be deduced from [1, Proposition 10.13] that f_p is uniformly convex with modulus $\phi = c_p(\cdot)^p$, where

$$c_p = 2^{1-3p/2} \min\{p, 2^{p/2-1}\}.$$

Obviously, this function ϕ belongs to \mathcal{U}_p but $\phi \notin \mathcal{U}_q$ for $2 \leq q < p$.

2.1 A first order characterization of uniform convexity

The following result may be seen as an analogue of Theorem A for uniform convexity.

Theorem 2.3. *Let C be as in Lemma 2.1 and let $f : C \rightarrow \mathbb{R}$ be differentiable. Then f is uniformly convex on C with modulus $\phi \in \mathcal{U}_p$, where $p \geq 2$, if and only if for some constant $c > 0$ and any $\mathbf{x}, \mathbf{y} \in C$ there holds*

$$(8) \quad f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + c\phi(\|\mathbf{y} - \mathbf{x}\|).$$

Moreover, for $p = 2$ the preceding statement holds with $c = 1$.

Proof. If f is uniformly convex with modulus ϕ , then, by Lemma 2.1, inequality (8) holds with $c = 1$ and, of course with any $c \in]0, 1]$ as well.

Conversely, suppose that (8) holds with some $c > 0$. Take any $\mathbf{x}, \mathbf{y} \in C$, $t \in]0, 1[$ and write $\mathbf{z} := t\mathbf{x} + (1-t)\mathbf{y}$ for short. Since $\mathbf{z} \in C$ due to the convexity of C , we may use (8) twice to conclude that

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{z}) &\geq (1-t)\langle \nabla f(\mathbf{z}), \mathbf{x} - \mathbf{y} \rangle + c\phi(\|(1-t)(\mathbf{x} - \mathbf{y})\|), \\ f(\mathbf{y}) - f(\mathbf{z}) &\geq t\langle \nabla f(\mathbf{z}), \mathbf{y} - \mathbf{x} \rangle + c\phi(\|t(\mathbf{x} - \mathbf{y})\|). \end{aligned}$$

Since $\phi \in \mathcal{U}_p$, it follows from property (ii) of this class that

$$\begin{aligned} f(\mathbf{x}) - f(\mathbf{z}) &\geq (1-t)\langle \nabla f(\mathbf{z}), \mathbf{x} - \mathbf{y} \rangle + c(1-t)^p\phi(\|\mathbf{x} - \mathbf{y}\|), \\ f(\mathbf{y}) - f(\mathbf{z}) &\geq t\langle \nabla f(\mathbf{z}), \mathbf{y} - \mathbf{x} \rangle + ct^p\phi(\|\mathbf{x} - \mathbf{y}\|). \end{aligned}$$

Next, multiplying the first inequality by t and the second by $1 - t$ and adding the results, we obtain

$$(9) \quad tf(\mathbf{x}) + (1 - t)f(\mathbf{y}) \geq f(\mathbf{z}) + ct(1 - t)[t^{p-1} + (1 - t)^{p-1}]\phi(\|\mathbf{x} - \mathbf{y}\|).$$

Now we want to estimate $t^{p-1} + (1 - t)^{p-1}$ from below. If $p = 2$, then

$$t^{p-1} + (1 - t)^{p-1} \geq t + (1 - t) = 1,$$

and so (4) follows from (9) for $c = 1$. This completes already the proof of the last assertion.

If $p > 2$, then by the convexity of $g(t) := t^{p-1}$ on $]0, 1[$, we have

$$\frac{1}{2}t^{p-1} + \frac{1}{2}(1 - t)^{p-1} \geq \left(\frac{t + (1 - t)}{2}\right)^{p-1} = \frac{1}{2^{p-1}}.$$

Hence (4) follows from (9) for $c = 2^{p-2}$. This completes the proof of the first statement as well. \square

The second statement of Theorem 2.3 as well as Theorem A allows us to characterize strong convexity as follows.

Corollary 2.4. *Let C be as in Lemma 2.1 and let $f : C \rightarrow \mathbb{R}$ be differentiable. Then f is strongly convex on C with constant $\mu > 0$ if and only if for any $\mathbf{x}, \mathbf{y} \in C$ there holds*

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{\mu}{2} \|\mathbf{y} - \mathbf{x}\|^2.$$

2.2 A monotone gradient characterization of uniform convexity

In this subsection we want to establish an analogue of Theorem B for uniform convexity. For this purpose, we need a strengthened form of monotonicity.

Definition 2.5. *Let C be as in Lemma 2.1 and let $G : C \rightarrow \mathbb{R}^d$. Then G is said to be uniformly monotone on C with modulus ϕ if for any $\mathbf{x}, \mathbf{y} \in C$ we have*

$$\langle G(\mathbf{y}) - G(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq \phi(\|\mathbf{y} - \mathbf{x}\|).$$

Theorem 2.6. *Let C be as in Lemma 2.1 and let $f : C \rightarrow \mathbb{R}$ be continuously differentiable. Then f is uniformly convex on C with modulus $\phi \in \mathcal{U}_p$, where $p \geq 2$, if and only if for some constant $c > 0$ the gradient ∇f is uniformly monotone on C with modulus $c\phi$, i.e.,*

$$(10) \quad \langle \nabla f(\mathbf{y}) - \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq c\phi(\|\mathbf{y} - \mathbf{x}\|)$$

for all $\mathbf{x}, \mathbf{y} \in C$. Moreover, for $p = 2$ the preceding statement holds with $c = 2$.

Proof. If f is uniformly convex with modulus ϕ , then, by Lemma 2.1, inequality (5) holds and, of course, it also holds when we interchange the roles of \mathbf{x} and \mathbf{y} . Thus we obtain the inequalities

$$\begin{aligned} f(\mathbf{y}) &\geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \phi(\|\mathbf{y} - \mathbf{x}\|), \\ f(\mathbf{x}) &\geq f(\mathbf{y}) + \langle \nabla f(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle + \phi(\|\mathbf{x} - \mathbf{y}\|). \end{aligned}$$

Adding them, we deduce that

$$\langle \nabla f(\mathbf{y}) - \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq 2\phi(\|\mathbf{y} - \mathbf{x}\|).$$

This shows that (10) holds with $c = 2$.

Conversely, suppose that (10) holds. For $n \in \mathbb{N}$, define $t_i := i/(n+1)$ and set $\mathbf{z}_i = \mathbf{x} + t_i(\mathbf{y} - \mathbf{x})$, where $i = 0, \dots, n+1$. Note that along with \mathbf{x} and \mathbf{y} all of these points lie in C . Since f is continuously differentiable, we conclude with the help of the mean-value theorem that there exist points $\xi_i \in]t_i, t_{i+1}[$ such that

$$f(\mathbf{z}_{i+1}) - f(\mathbf{z}_i) = (t_{i+1} - t_i) \langle \nabla f(\mathbf{x} + \xi_i(\mathbf{y} - \mathbf{x})), \mathbf{y} - \mathbf{x} \rangle \quad (i = 0, \dots, n).$$

Summing up, we find that

$$\begin{aligned} f(\mathbf{y}) - f(\mathbf{x}) &= \sum_{i=0}^n (f(\mathbf{z}_{i+1}) - f(\mathbf{z}_i)) \\ &= \sum_{i=0}^n (t_{i+1} - t_i) \langle \nabla f(\mathbf{x} + \xi_i(\mathbf{y} - \mathbf{x})), \mathbf{y} - \mathbf{x} \rangle \\ &= \sum_{i=0}^n (t_{i+1} - t_i) \langle \nabla f(\mathbf{x} + \xi_i(\mathbf{y} - \mathbf{x})) - \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \\ &\quad + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \\ &= \sum_{i=0}^n \frac{(t_{i+1} - t_i)}{\xi_i} \langle \nabla f(\mathbf{x} + \xi_i(\mathbf{y} - \mathbf{x})) - \nabla f(\mathbf{x}), \xi_i(\mathbf{y} - \mathbf{x}) \rangle \\ &\quad + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle. \end{aligned}$$

Using (10) for estimating the inner products from below and recalling property (ii) of the class \mathcal{U}_p , we may continue as follows:

$$\begin{aligned} f(\mathbf{y}) - f(\mathbf{x}) &\geq c \sum_{i=0}^n \frac{(t_{i+1} - t_i)}{\xi_i} \phi(\xi_i \|\mathbf{x} - \mathbf{y}\|) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \\ &\geq c \phi(\|\mathbf{x} - \mathbf{y}\|) \sum_{i=0}^n (t_{i+1} - t_i) \xi_i^{p-1} + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle. \end{aligned}$$

On the right-hand side we have a Riemann sum which yields

$$\lim_{n \rightarrow +\infty} \sum_{i=0}^n (t_{i+1} - t_i) \xi_i^{p-1} = \int_0^1 x^{p-1} dx = \frac{1}{p}.$$

Thus, letting n tend to infinity, we obtain

$$(11) \quad f(\mathbf{y}) - f(\mathbf{x}) \geq \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{c}{p} \phi(\|\mathbf{y} - \mathbf{x}\|).$$

Now it follows from Theorem 2.3 and its proof that for an appropriate choice of $c > 0$, namely $c = p2^{p-2}$, inequality (11) implies (4).

For $p = 2$ the quotient in (11) reduces to 1 by setting $c = 2$. Then the second statement of Theorem 2.3 guarantees that (4) holds. This completes the proof. \square

The second statement of Theorem 2.6 as well as Theorem B allows us to characterize strong convexity as follows.

Corollary 2.7. *Let C be as in Lemma 2.1 and let $f : C \rightarrow \mathbb{R}$ be continuously differentiable. Then f is strongly convex on C with constant $\mu > 0$ if and only if*

$$\langle \nabla f(\mathbf{y}) - \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \geq \mu \|\mathbf{y} - \mathbf{x}\|^2$$

for all $\mathbf{x}, \mathbf{y} \in C$.

2.3 A second order characterization of uniform convexity

As regards Theorem C, there is a situation that resembles the case of strict convexity discussed in the introduction. We could not achieve a second order characterization of uniform convexity with modulus $\phi \in \mathcal{U}_p$ if $p > 2$. But we can present a sufficient second order condition for these values of p .

Proposition 2.8. *Let C be as in Lemma 2.1 and let $f : C \rightarrow \mathbb{R}$ be twice continuously differentiable. If $\phi \in \mathcal{U}_p$ with $p \geq 2$ and*

$$(12) \quad \langle \mathbf{w}, \nabla^2 f(\mathbf{x}) \mathbf{w} \rangle \geq 2^{p-1} \phi(\|\mathbf{w}\|)$$

for all $\mathbf{x} \in C$ and $\mathbf{w} \in \mathbb{R}^d$, then f is uniformly convex on C with modulus ϕ .

Proof. For $\mathbf{x}, \mathbf{y} \in C$, we have by Taylor’s theorem that

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, \nabla^2 f(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x}) \rangle$$

with some $\theta \in [0, 1]$. The convexity of C guarantees that $\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}) \in C$. Therefore (12) implies that

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + 2^{p-2} \phi(\|\mathbf{y} - \mathbf{x}\|).$$

This is (8) with $c = 2^{p-2}$. Now Theorem 2.3 and its proof show that (4) holds, and so f is uniformly convex with modulus ϕ . \square

For $p = 2$ condition (12) is necessary and sufficient for f to be uniformly convex with modulus ϕ . Moreover, several other characterizations hold which imply that f is strongly convex, as the following theorem shows.

Theorem 2.9. *Let C be as in Lemma 2.1 and let $f : C \rightarrow \mathbb{R}$ be twice continuously differentiable. For $\phi \in \mathcal{U}_2$ set $\kappa := \sup_{r>0} \phi(r)/r^2$. Then the following statements are equivalent:*

- (i) f is uniformly convex with modulus ϕ ;
- (ii) for any $\mathbf{x} \in C$ and all $\mathbf{w} \in \mathbb{R}^d$, there holds

$$(13) \quad \langle \mathbf{w}, \nabla^2 f(\mathbf{x})\mathbf{w} \rangle \geq 2\phi(\|\mathbf{w}\|);$$

- (iii) for each $\mathbf{x} \in C$ the eigenvalues of the Hessian $\nabla^2 f(\mathbf{x})$ lie in the interval $[2\kappa, +\infty[$;
- (iv) for any $\mathbf{x} \in C$ and all $\mathbf{w} \in \mathbb{R}^d$, there holds

$$\langle \mathbf{w}, \nabla^2 f(\mathbf{x})\mathbf{w} \rangle \geq 2\kappa\|\mathbf{w}\|^2;$$

- (v) f is strongly convex with constant 2κ .

Proof. It is enough to prove the following implications:

$$(i) \implies (ii) \implies (iii) \implies (iv) \implies (v) \implies (i)$$

Suppose that (i) holds. For any $\mathbf{x} \in C$ and all $\mathbf{y} \in C$ which are sufficiently close to \mathbf{x} , we have by Taylor expansion

$$(14) \quad \begin{aligned} f(\mathbf{y}) &= f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \\ &+ \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, \nabla^2 f(\mathbf{x})(\mathbf{y} - \mathbf{x}) \rangle + e(\mathbf{x}, \mathbf{y}) \end{aligned}$$

with a remainder e satisfying

$$\lim_{\mathbf{y} \rightarrow \mathbf{x}} \frac{e(\mathbf{x}, \mathbf{y})}{\|\mathbf{y} - \mathbf{x}\|^2} = 0.$$

We may take $\mathbf{y} = \mathbf{x} + \varepsilon\mathbf{w}$ with an arbitrary $\mathbf{w} \in \mathbb{R}^d$ and any sufficiently small $\varepsilon > 0$. Employing Lemma 2.1 and recalling that $\phi \in \mathcal{U}_2$, we deduce from (14) that

$$\frac{\varepsilon^2}{2} \langle \mathbf{w}, \nabla^2 f(\mathbf{x})\mathbf{w} \rangle + e(\mathbf{x}, \mathbf{x} + \varepsilon\mathbf{w}) \geq \phi(\|\varepsilon\mathbf{w}\|) \geq \varepsilon^2 \phi(\|\mathbf{w}\|).$$

Dividing by $\varepsilon^2/2$ and excluding $\mathbf{w} = 0$ for the moment, we find that

$$\langle \mathbf{w}, \nabla^2 f(\mathbf{x})\mathbf{w} \rangle + 2\|\mathbf{w}\|^2 \frac{e(\mathbf{x}, \mathbf{x} + \varepsilon\mathbf{w})}{\|\varepsilon\mathbf{w}\|^2} \geq 2\phi(\|\mathbf{w}\|).$$

Now, letting $\varepsilon \rightarrow 0$, we arrive at (13) for all $\mathbf{w} \neq 0$. For $\mathbf{w} = 0$ relation (13) holds trivially as an equality.

Next suppose that (ii) holds. Let $\mathbf{x} \in C$ be arbitrary but fixed. Since the Hessian $\nabla^2 f(\mathbf{x})$ is a real, symmetric matrix, its eigenvalues are all real. Let λ_{\min} be its smallest eigenvalue. For an arbitrary $r > 0$, let \mathbf{v} be an associated eigenvector of Euklidean length r . Then

$$\langle \mathbf{v}, \nabla^2 f(\mathbf{x})\mathbf{v} \rangle = \langle \mathbf{v}, \lambda_{\min} \mathbf{v} \rangle = \lambda_{\min} \|\mathbf{v}\|^2 = \lambda_{\min} r^2.$$

Thus, as a consequence of (13),

$$\lambda_{\min} \geq \frac{2}{r^2} \phi(\|\mathbf{v}\|) = \frac{2}{r^2} \phi(r).$$

Since this is valid for arbitrary $r > 0$, we conclude that $\lambda_{\min} \geq 2\kappa$, and so statement (iii) holds.

For any $\mathbf{x} \in C$ and $\mathbf{w} \in \mathbb{R}^d$ with $\mathbf{w} \neq 0$, consider the so-called Rayleigh quotient

$$\frac{\langle \mathbf{w}, \nabla^2 f(\mathbf{x})\mathbf{w} \rangle}{\|\mathbf{w}\|^2}.$$

It is bounded from below by the smallest eigenvalue of $\nabla^2 f(\mathbf{x})$, the bound being attained when \mathbf{w} is an associated eigenvector; see [3, § 8.1–§ 8.2]. Hence statement (iii) implies that

$$\frac{\langle \mathbf{w}, \nabla^2 f(\mathbf{x})\mathbf{w} \rangle}{\|\mathbf{w}\|^2} \geq 2\kappa.$$

Together with the trivial case of $\mathbf{w} = 0$, this shows that statement (iv) holds.

Next we want to show that (iv) implies (v). For $\mathbf{x}, \mathbf{y} \in C$, we have by Taylor's theorem that

$$f(\mathbf{y}) = f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \frac{1}{2} \langle \mathbf{y} - \mathbf{x}, \nabla^2 f(\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}))(\mathbf{y} - \mathbf{x}) \rangle$$

with some $\theta \in [0, 1]$. Clearly, $\mathbf{x} + \theta(\mathbf{y} - \mathbf{x}) \in C$ due to the convexity of C . Therefore statement (iv) allows us to deduce that

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle + \kappa \|\mathbf{y} - \mathbf{x}\|^2.$$

Now Corollary 2.4 guarantees that f is strongly convex with constant 2κ .

Finally, suppose that (v) holds. As we mentioned in the introduction, this implies that f is uniformly convex with modulus $\kappa(\cdot)^2$. In other words, we have

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}) - t(1-t)\kappa\|\mathbf{y} - \mathbf{x}\|^2$$

for any $\mathbf{x}, \mathbf{y} \in C$ and all $t \in [0, 1]$. Now the definition of κ shows that

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)f(\mathbf{y}) - t(1-t)\frac{\phi(r)}{r^2} \|\mathbf{y} - \mathbf{x}\|^2$$

for any $\mathbf{x}, \mathbf{y} \in C$, all $t \in [0, 1]$ and all $r > 0$. In particular, this inequality holds for $r = \|\mathbf{y} - \mathbf{x}\|$ provided that $\mathbf{y} \neq \mathbf{x}$. Thus, considering the trivial case of $\mathbf{y} = \mathbf{x}$ separately, we arrive at (i). \square

Remark 2.10. The definition of κ in Theorem 2.9 does not guarantee that $\kappa < +\infty$. However, if one of statements (i)–(v) holds, then all are true. Among them statements (iii) and (iv) show that κ must be finite.

Theorem 2.9 includes the following characterization of strong convexity, which is also a simple consequence of Theorem C.

Corollary 2.11. *Let C be as in Lemma 2.1 and let $f : C \rightarrow \mathbb{R}$ be twice continuously differentiable. Then f is strongly convex with constant $\mu > 0$ if and only if*

$$\langle \mathbf{w}, \nabla^2 f(\mathbf{x}) \mathbf{w} \rangle \geq \mu \|\mathbf{w}\|^2$$

for all $\mathbf{x} \in C$ and $\mathbf{w} \in \mathbb{R}^d$.

REFERENCES

1. H. H. BAUSCHKE, P. L. COMBETTES: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, New York, 2011.
2. S. BOYD, L. VANDENBERGHE: *Convex Optimization*. Cambridge University Press, 2004.
3. P. LANCASTER, M. TISMENETSKY: *The Theory of Matrices (2nd Edition)*. Academic Press, Orlando, 1985.

Osama Alabdali

Department of Mathematics
College of Education for Pure Sciences
University Of Anbar
Ramadi
Iraq
E-mail: *alsltan.2006@yahoo.com*

(Received 22.03.2019)

(Revised 20.07.2019)

Allal Guessab

Université de Pau et des Pays de l'Adour,
E2S UPPA, CNRS, LMAP
64000 Pau
France
E-mail: *allal.guessab@univ-pau.fr*

Gerhard Schmeisser

Department Mathematik
Friedrich-Alexander-Universität Erlangen-Nürnberg
91058 Erlangen
Germany
E-mail: *schmeisser@mi.uni-erlangen.de*