fractal and fractional

MDPI

Article

A q-Gradient Descent Algorithm with Quasi-Fejér Convergence for Unconstrained Optimization Problems

Shashi Kant Mishra [1,†], Predrag Rajković [2,†], Mohammad Esmael Samei [3,4,*,†], Suvra Kanti Chakraborty [5,†], Bhagwat Ram [6,7,†] and Mohammed K. A. Kaabar [8,9,*,†]

1 Department of Mathematics, Institute of Science, Banaras Hindu University, Varanasi 221005, India; bhu.skmishra@gmail.com
2 Faculty of Mechanical Engineering, University of Niš, 18000 Niš, Serbia; pedja.rajk@gmail.com
3 Department of Mathematics, Faculty of Basic Science, Bu-Ali Sina University, Hamedan 65178, Iran
4 Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 40402, Taiwan
5 Department of Mathematics, Ramakrishna Mission Vidyamandira Belur Math, Howrah 711202, India; suvrakanti89@gmail.com
6 Centre for Digital Transformation, Indian Institute of Management Ahmadabad (IIMA), Vastrapur 380015, India; bhagwatr@iima.ac.in or bhagwatram14@gmail.com
7 DST-Centre for Interdisciplinary Mathematical Sciences, Institute of Science, Banaras Hindu University, Varanasi 221005, India
8 Institute of Mathematical Sciences, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia
9 Department of Mathematics and Statistics, Washington State University, Pullman, WA 99163, USA
* Correspondence: mesamei@basu.ac.ir; mesamei@gmail.com (M.E.S.); mohammed.kaabar@wsu.edu (M.K.A.K.)
† These authors contributed equally to this work.

Abstract: We present an algorithm for solving unconstrained optimization problems based on the q-gradient vector. The main idea used in the algorithm construction is the approximation of the classical gradient by a q-gradient vector. For a convex objective function, the quasi-Fejér convergence of the algorithm is proved. The proposed method does not require the boundedness assumption on any level set. Further, numerical experiments are reported to show the performance of the proposed method.

Keywords: descent methods; q-calculus; iterative methods; inexact line searches

MSC: 41A60; 05A30; 65F08

check for updates

Citation: Mishra, S.K.; Rajković, P.; Samei, M.E.; Chakraborty, S.K.; Ram, B.; Kaabar, M.K.A. A q-Gradient Descent Algorithm with Quasi-Fejér Convergence for Unconstrained Optimization Problems. *Fractal Fract.* **2021**, *5*, 110. https://doi.org/10.3390/fractalfract5030110

Academic Editors: Vojislav V. Mitic, Cristina Serpa and Hans-Jörg Fecht

Received: 7 June 2021
Accepted: 27 August 2021
Published: 3 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The descent direction plays a central role in the development of optimization algorithms. The classical gradient descent method was first proposed by Cauchy [1] in 1847. The optimization problem is a significant mathematical model in a wide class of disciplines [2]. In applications such as image processing [3], data analysis [4], and machine learning [5], in which one needs to quickly provide an approximate solution, several gradient-based algorithms [6–9] have been proposed based on the iterative technique of the gradient descent method [2,10]. Quantum calculus is the modern field for the investigation of calculus without limits. Quantum calculus, or q-calculus, began with the work of Jackson in the early twentieth century [11], but similar kinds of calculus had already been developed by Euler and Jacobi in the eighteenth and nineteenth centuries, respectively. Recently it has come under deep interest due to the high demand of mathematics that models quantum computing. Besides appearing as a connection between mathematics and physics, q-calculus has many applications in different mathematical areas such as operator theory [12], combinatorics [13], orthogonal polynomials [14], basic hyper-geometric

functions [15], and other sciences like quantum theory [16,17], mechanics [18], and fractional calculus [19–25]. For more recent studies about fractional calculus analysis and applications, refer to [26–31].

The q-Taylor formula for functions of several variables with mean value theorems in q-calculus was first used to develop a new method for solving systems of equations [32]. The advantage of q-calculus is shown in the next example [32] where the a scheme involving q-derivatives finds the solution, but the classical Newton–Kantorovich method fails to do so

$$\begin{cases} |x_1^2 - 4| + e^{7x_2 - 36} = 2, \\ \log_{10}\left|\frac{12x_1^2}{x_2} - 6\right| + x_1^4 = 9. \end{cases}$$

For $q = 0.9$, the iterations converge to the exact solution as

$$(x_1, x_2)^T = \left(\sqrt{3}, \frac{36}{7}\right)^T.$$

The concept of the q-analogy of the gradient was first introduced in [33] for solving systems of equations. It has some advantages with respect the classical method when the functions are not differentiable. Further, the same concept of q-gradient was introduced in the steepest descent method [34] to optimize single objective functions [35]. The parameter q was generated by a Gaussian distribution with standard deviation σ that decreases by an iterative process as

$$\sigma^{(k+1)} = \beta\sigma^{(k)},$$

with a starting standard deviation σ^0 and the reduction factor β. The step length was generated using the golden section search method [34]. However, the convergence properties of the steepest descent method with inexact line searches have been studied under several strategies for the choice of the step length α_k [36–38]. Recently, several modified unconstrained optimization algorithms using the q-gradient have been proposed to solve unconstrained optimization problems [10,19,39–41].

In this paper, we propose a q-gradient line search scheme that provides a q-descent direction at every kth iteration. For this, a sequence $\left\{q^{(k)}\right\}$ [39] is taken to generate the values of q, and the backtracking technique is utilized to find the step length without requiring the bounded level sets or Lipschitz condition on the gradient of the function. We also provide the convergence proof theoretically when step length is fixed without any hypothesis on the level sets of the objective function. The advantage of using q-gradient is shown by comparing our method with the method given in [36] based on the number of iterations and function evaluations.

The paper is organized as follows. In the next section, some notations and definitions for q-calculus and other prerequisites are provided, which are used throughout the paper. In Section 3, the q-gradient descent algorithm is given, and its convergence analysis is provided in Section 4. Numerical experiments are performed in Section 4, which is followed by a section of concluding remarks.

2. Essential Preliminaries

We assume that $\mathbb{R}_+$ stands for the nonnegative real line, $q \in (0,1)$ is a real number, and the q-integer $[n]_q$ is given as:

$$[n]_q = \begin{cases} \dfrac{1-q^n}{1-q}, & q \neq 1, \\ n, & q = 1, \end{cases} \tag{1}$$

for all $n \in \mathbb{N}$. The expansion of $(1+x)_q^n$ is:

$$(1+x)_q^n = \begin{cases} 1, & n = 0, \\ (1+x)(1+qx)\cdots(1+q^{n-1}x), & n \geq 1. \end{cases} \tag{2}$$

The q-derivative of x^n with respect to x is:

$$D_q x^n = [n]_q x^{n-1}.$$

The q-derivative of a function $\psi : \mathbb{R} \to \mathbb{R}$ is given by

$$D_q \psi(x) = \frac{\psi(qx) - \psi(x)}{qx - x}, \qquad (x \neq 0).$$

In the special case, $D_q \psi(0) = \frac{\mathrm{d}\psi(0)}{\mathrm{d}x}$.
If ψ is differentiable, then

$$\lim_{q \to 1} D_q \psi(x) = \lim_{q \to 1} \frac{\psi(qx) - \psi(x)}{(q-1)x} = \frac{\mathrm{d}\psi(x)}{\mathrm{d}x}.$$

The higher order q-derivatives of ψ are:

$$D_q^0 \psi := \psi, \qquad D_q^n := D_q(D_q^{n-1}\psi), \ (n = 1, 2, 3, \dots).$$

The q-derivative of a function is a linear operator [42] for any constants c_1 and c_2 as:

$$D_q\{c_1\psi_1(x) + c_2\psi_2(x)\} = c_1 D_q \psi_1(x) + c_2 D_q \psi_2(x).$$

Let $\psi(x)$ be a continuous function on $[a, b]$, where $a, b \in \mathbb{R}$. Then, there exists $\hat{q} \in (0, 1)$ and $x \in (a, b)$ [43] such that

$$\psi(b) - \psi(a) = D_q \psi(x)(b - a), \tag{3}$$

for all $q \in (\hat{q}, 1) \cup (1, \hat{q}^{-1})$. The q-partial derivative of function $\psi : \mathbb{R}^n \to \mathbb{R}$ at $x \in \mathbb{R}^n$ with respect to x_i is defined as (see ([35])):

$$D_{q_i, x_i} \psi(x) = \begin{cases} \dfrac{1}{(1-q_i)x_i}\big[\psi(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) \\ \qquad -\psi(x_1, x_2, \dots, x_{i-1}, q_i x_i, x_{i+1}, \dots, x_n)\big], & x_i \neq 0, q_i \neq 1, \\ \dfrac{\partial}{\partial x_i}\psi(x_1, x_2, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n), & x_i = 0, \\ \dfrac{\partial}{\partial x_i}\psi(x_1, x_2, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n), & q_i = 1. \end{cases}$$

A function is called q-differentiable at a point if its q-partial derivatives are continuous. A continuously q-differentiable function ψ is a function whose q-derivative function is also continuous at the point. We now choose the parameter q as a vector; that is,

$$q = (q_1, \dots, q_i, \dots, q_n)^T \in \mathbb{R}^n.$$

Then, the q-gradient vector [35] of $\psi : \mathbb{R}^n \to \mathbb{R}$ is:

$$\nabla_q \psi(x)^T = \begin{bmatrix} D_{q_1, x_1}\psi(x) & \dots & D_{q_i, x_i}\psi(x) & \dots & D_{q_n, x_n}\psi(x) \end{bmatrix}. \tag{4}$$

Let $\{q_i^{(k)}\}$ [10,41] be a real sequence defined by

$$q_i^{(k+1)} = 1 - \frac{q_i^{(k)}}{(k+1)^2}, \tag{5}$$

for each $i = 1, \ldots, n$, where $k = 0, 1, 2 \ldots$ and $0 < q_i^{(0)} < 1$. Of course, the sequence $\{q_i^{(k)}\}$ converges to $(1, \ldots, 1)^T$ as $k \to \infty$, so that the q-gradient reduces to the classic gradient vector.

Proposition 1. *If $\psi(x) = a_0 + x^T a$, where $a_0 \in \mathbb{R}$ and $a \in \mathbb{R}^n$, then for any x, $q \in \mathbb{R}^n$,*

$$\nabla_{q^{(k)}} \psi(x) = \nabla \psi(x) = a.$$

Example 1. *Consider a function $\psi : \mathbb{R}^2 \to \mathbb{R}$ such that $\psi(x) = \frac{1}{x_1 x_2}$. Then, the q-gradient is given as*

$$\nabla_{q^{(k)}} \psi(x) = -\frac{1}{x_1 x_2} \begin{bmatrix} \frac{1}{q_1^{(k)} x_1} \\ \frac{1}{q_2^{(k)} x_2} \end{bmatrix}.$$

Definition 1 (q–Integral [42])**.** *The q-analog of the integral is given by*

$$\int_0^b \psi(x) \, \mathrm{d}_q x = b(1-q) \sum_{k=0}^{\infty} \psi(bq^k) q^k. \tag{6}$$

In the special case, it reduces to the classical integral $\int_0^b \psi(x) \, dx$, when $q \to 1$. It is right only in the case [44]

$$\int_0^1 x \, d_q x = (1-q) \sum_{k=0}^{\infty} q^k \cdot q^k = \frac{1}{1+q}.$$

Definition 2 (q-Newton–Leibniz formula [42])**.** *The q-anti-derivative of $\psi(x)$ is*

$$F_q(x) = \int_0^x \psi(t) \, \mathrm{d}_q t. \tag{7}$$

In that manner, the q-Newton–Leibnitz formula is

$$\int_a^b \psi(t) \, \mathrm{d}_q t = F_q(b) - F_q(a). \tag{8}$$

Definition 3 (Quasi-Fejér Convergence [36])**.** *A sequence $\{x^{(k)}\}$ is quasi-Fejèr convergent to a set $\mathcal{U} \subseteq \mathbb{R}^n$ if for every $u \in \mathcal{U}$, there exists a non-negative, summable sequence $\{\epsilon_k\} \subseteq \mathbb{R}$ such that $\epsilon_k \geq 0$, $\sum_{k=0}^{\infty} \epsilon_k < \infty$ and*

$$\left\| x^{(k+1)} - u \right\|^2 \leq \left\| x^{(k)} - u \right\|^2 + \epsilon_k,$$

for all k.

In the next section, the q-gradient descent algorithm is presented for solving unconstrained optimization problems.

3. A q–Gradient Descent Algorithm

We consider the unconstrained optimization problem:

$$\min \psi(x), \qquad x \in \mathbb{R}^n, \tag{9}$$

where $\psi : \mathbb{R}^n \to \mathbb{R}$ is a continuously q-differentiable convex function. Note that min is same as the max of $-\psi(x)$. We choose a starting point $x^{(0)} \in \mathbb{R}^n$. The general iterative scheme to solve (9) using the q-gradient is of the following form:

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}_{q^{(k)}}, \tag{10}$$

where $x^{(k+1)}$ is a new iterative point, $x^{(k)}$ is the previous iterative point, and $d^{(k)}_{q^{(k)}}$ is a q-descent direction given as:

$$d^{(k)}_{q^{(k)}} = -\nabla_{q^{(k)}} \psi\left(x^{(k)}\right),$$

and $\alpha_k \in \mathbb{R}$ is called the step length can be computed by two main line search strategies: exact line search and inexact line search. The exact optimal step length theoretically cannot be found for practical computation, and it is expensive to generate the value of α_k. Therefore, the most frequently used technique in practice is inexact line searching for finding the descent direction. When inexact line searches are performed then α_k is assigned a given predetermined value or through some finite iterative method. Note that existence of a solution to the minimization problem (9) is implicitly assumed. The condition for the existence of a solution [2] in the context of q-calculus is:

$$x^{(k)} \to x^* \in \Omega := \left\{ x : \nabla_{q^{(k)}} \psi(x) = 0 \right\}.$$

The following result presents the first-order necessary condition in the light of q-calculus.

Theorem 1. *Let Ω be a subset of $\mathbb{R}^n$ and ψ be a first-order q-differentiable real valued function on Ω. If x^* is a local minimizer of ψ over Ω, then for any feasible q-direction $d^{(k)}_{q^{(k)}}$ at x^*, we have*

$$\left(d^{(k)}_{q^{(k)}}\right)^T \nabla_{q^{(k)}} \psi(x^*) \geq 0.$$

Proof. We consider $x(\alpha) := x^* + \alpha d^{(k)}_{q^{(k)}} \in \Omega$. For $\alpha = 0$, we obtain $x(0) = x^*$. Define the composite function $f(\alpha) = \psi(x(\alpha))$. Applying q-Taylor's theorem to $\psi(\alpha)$ and taking $\alpha = 0$, we obtain

$$\psi(\alpha) = \psi(0) + \alpha\psi'(0) + O(\alpha).$$

We can write this as

$$\begin{aligned} \psi(\alpha) - \psi(0) &= \alpha\psi'(0) + O(\alpha) \\ \psi(x^* + \alpha d) - \psi(x^*) &= \alpha\psi'(0) + O(\alpha), \end{aligned}$$

and

$$\begin{aligned} \psi'(\alpha) = \frac{\mathrm{d}_q}{\mathrm{d}_q\alpha}\psi(x(\alpha)) &= \frac{\mathrm{d}_q}{\mathrm{d}_q\alpha}\psi\left(x^* + \alpha d^{(k)}_{q^{(k)}}\right) \\ &= \left(d^{(k)}_{q^{(k)}}\right)^T \nabla_{q^{(k)}} \psi\left(x^* + \alpha d^{(k)}_{q^{(k)}}\right). \end{aligned}$$

At $\alpha = 0$ and when we have $x^* = x(0)$, then

$$\psi'(0) = \left(d^{(k)}_{q^{(k)}}\right)^T \nabla_{q^{(k)}} \psi(x(0)).$$

Therefore,

$$\psi\left(x^* + \alpha d_{q^{(k)}}^{(k)}\right) - \psi(x^*) = \alpha\left(d_{q^{(k)}}^{(k)}\right)^T \nabla_{q^{(k)}}\psi(x(0)) + O(\alpha), \tag{11}$$

where $\alpha \geq 0$. Since x^* is a local minimizer of ψ over Ω, for α sufficiently small $\alpha > 0$, we can write

$$\psi(x^*) \leq \psi\left(x^* + \alpha d_{q^{(k)}}^{(k)}\right). \tag{12}$$

From (11) and (12),

$$\alpha\left(d_{q^{(k)}}^{(k)}\right)^T \nabla_{q^{(k)}}\psi(x(0)) = \psi\left(x^* + \alpha d_{q^{(k)}}^{(k)}\right) - \psi(x^*) \geq 0,$$

that is,

$$\alpha\left(d_{q^{(k)}}^{(k)}\right)^T \nabla_{q^{(k)}}\psi(x(0)) \geq 0.$$

Since $\alpha > 0$, then

$$\left(d_{q^{(k)}}^{(k)}\right)^T \nabla\psi(x(0)) \geq 0.$$

Since $x(0) = x^*$, then

$$\left(d_{q^{(k)}}^{(k)}\right)^T \nabla_{q^{(k)}}\psi(x^*) \geq 0.$$

This completes the proof. □

We present the following result for the interior case in the context of q-calculus.

Corollary 1. *Let Ω be a subset of $\mathbb{R}^n$ and ψ be a first-order q-differentiable real valued function on Ω. If x^* is a local minimizer of ψ over Ω and if x^* is an interior point of Ω, then $\nabla_{q^{(k)}}\psi(x^*) = 0$.*

Proof. Since x^* is a local minimizer of ψ over Ω, then for any feasible q-direction, we have

$$\left(d_{q^{(k)}}^{(k)}\right)^T \nabla_{q^{(k)}}\psi(x^*) \geq 0. \tag{13}$$

Since x^* is an interior point of Ω, then every q-direction is a feasible direction, therefore

$$-\left(d_{q^{(k)}}^{(k)}\right)^T \nabla_{q^{(k)}}f(x^*) \geq 0. \tag{14}$$

From (13) and (14), we obtain

$$\left(d_{q^{(k)}}^{(k)}\right)^T \nabla_{q^{(k)}}\psi(x^*) = 0.$$

For all $d_{q^{(k)}}^{(k)} \in \mathbb{R}^n$, we obtain $\nabla_{q^{(k)}}\psi(x^*) = 0$. This completes the proof. □

Before writing an algorithm for the q-gradient descent method, we need the following assumptions.

Assumption 1. *We consider the two following assumptions:*

1. *Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be convex and continuously q-differentiable.*

2. *The q-gradient of ψ with constant $\mathcal{L} > 0$ satisfies the following condition:*

$$\left\|\nabla_q\psi(x) - \nabla_q\psi(y)\right\| \leq \mathcal{L}\|x - y\|, \tag{15}$$

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

Assumption 2. *Let $\phi : \mathbb{R}_+ \to \mathbb{R}$ be such that:*

1. *ϕ is convex and continuously q-differentiable on $[0, \infty)$,*
2. *$\phi(0) = 0$ and $\phi'(0) < 1$,*
3. *$\lim_{v \to 0^+} \frac{\phi(v)}{v^2} > 0$.*

Note that (3) implies that $\phi'(0) \geq 0$, thus from (2), we obtain $0 \leq \phi'(0) < 1$ and from (1), ϕ is convex and continuously q-differentiable. Therefore, ϕ is non-decreasing. The statement of the theorem given in [45] can be presented in the light of q-calculus as given below.

Theorem 2. *Let $\mathcal{F} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ such that*

1. *There exists $(x^\circ, u^\circ) \in \mathbb{R}^n \times \mathbb{R}$ such that $\mathcal{F}(x^\circ, u^\circ) = 0$,*
2. *$\mathcal{F}$ is continuous in a neighborhood of (x°, u°),*
3. *$\mathcal{F}$ is q-differentiable with respect to the variable u in (x°, u°) and*

$$\frac{\partial_q \mathcal{F}}{\partial_q u}(x^\circ, u^\circ) \neq 0.$$

Then, there exists a neighborhood $\mathcal{U}(x^\circ)$ of x° and at least one function $\varrho : \mathcal{U}(x^\circ) \to \mathbb{R}$ such that $\varrho(x^\circ) = u^\circ$ and

$$\mathcal{F}(x, \varrho(x)) = 0, \tag{16}$$

for any $x \in \mathcal{U}(x^\circ)$.

4. *If $\frac{\partial_q \mathcal{F}}{\partial_q u}(.,.)$ is continuous at (x°, u°), then the function ϱ is the only one that satisfies (16) and is continuous at (x°, u°).*

Note that the proof q-analogue of the above implicit function theorem is out of the scope of the present research. For developing the algorithm, we need the following proposition whose proof is given in the light of [36].

Proposition 2. *Let*

$$\mathcal{G} = \left\{ x \in \mathbb{R}^n \mid \nabla_{q^{(k)}} \psi(x) \neq 0 \right\},$$

and ϕ satisfy Assumption 2. Then,

1. *For all $x \in \mathcal{G}$, there exists a unique $\varrho(x) > 0$ such that*

$$\psi(x) - \psi\left(x - \varrho(x) \nabla_{q^{(k)}} \psi(x)\right) = \phi(\varrho(x)) \left\| \nabla_{q^{(k)}} \psi(x) \right\|^2, \tag{17}$$

and

$$\psi\left(x - u \nabla_{q^{(k)}} \psi(x)\right) + \phi(u) \left\| \nabla_{q^{(k)}} \psi(x) \right\|^2 \leq \psi(x), \tag{18}$$

if and only if $0 \leq u \leq \varrho(x)$.

2. *$\varrho : \mathcal{G} \to \mathbb{R}_+$ is continuous in $\mathcal{G}$.*

Proof.

1. We first prove (1). Fix $x \in \mathcal{G}$, $u \in \mathbb{R}_+$ and define the function in the context of q-calculus as:

$$\mathcal{F}(x, u) = \psi\left(x - u \nabla_{q^{(k)}} \psi(x)\right) - \psi(x) + \phi(u) \left\| \nabla_{q^{(k)}} \psi(x) \right\|^2. \tag{19}$$

From (1) of Assumption 2, $\mathcal{F}(x,.)$ is convex and continuously q-differentiable, and when substituting $u = 0$ in (19), then we obtain

$$\mathcal{F}(x,0) = \phi(0)\left\|\nabla_{q^{(k)}}\psi(x)\right\|^2.$$

From (2) of Assumption 2, we have $\phi(0) = 0$, thus

$$\mathcal{F}(x,0) = 0. \tag{20}$$

Applying q-derivative with respect to u to (19), then

$$\frac{\partial_q \mathcal{F}}{\partial_q u} = \nabla_q\psi(x - u\nabla_q\psi(x))\big(-\nabla_q\psi(x)\big) + \phi'(0)\|\nabla_q\psi(x)\|^2, \tag{21}$$

Substituting $u = 0$ in the right hand side of above equation, we obtain

$$\frac{\partial_q \mathcal{F}}{\partial_q u} = \left\|\nabla_{q^{(k)}}\psi(x)\right\|^2(-1 + \phi'(0)), \tag{22}$$

Since $\phi'(0) < 1$, then

$$\frac{\partial_q \mathcal{F}}{\partial_q u} < 0. \tag{23}$$

In addition,

$$\mathcal{F}(x,u) \geq \psi^* - \psi(x) + \phi(u)\left\|\nabla_{q^{(k)}}\psi(x)\right\|^2, \tag{24}$$

where ψ^* is the minimum function value of ψ. From (20) and (23), we conclude that $\mathcal{F}(x,.)$ is negative in some interval to the right of zero, and from (24), (1) and (2) of Assumption 2, we obtain

$$\lim_{u\to\infty}\mathcal{F}(x,u) \geq +\infty. \tag{25}$$

From Theorem 2 it follows that there exists $\varrho(x) > 0$ such that $\mathcal{F}(x,\varrho(x)) = 0$. Using the above value in (19), we obtain (17). Since $\mathcal{F}(x,.)$ is convex, therefore there exists a uniqueness of $\varrho(x)$. Note that a convex function of a real variable can take a given value different from its minimum point at most two different points while

$$\mathcal{F}(x,.) = \mathcal{F}(x,\varrho(x)) = 0,$$

and from (20) and (23), the minimum point of $\mathcal{F}(x,.)$ is not zero. Thus, (1) of this proposition is proved.

2. Let $u^{(0)} = \varrho(x^{(0)})$ given by (1), for a given $x^{(0)} \in \mathcal{G}$. Then, we have that

$$\mathcal{F}(x^{(0)},u^{(0)}) = 0,$$

$F(.,.)$ is continuous in a neighborhood of $(x^{(0)},u^{(0)})$ and from (21)

$$\frac{\partial_q \mathcal{F}}{\partial_q u}\left(x^{(0)},u^{(0)}\right) = -\nabla_{q^{(k)}}\psi\left(x^{(0)} - u^{(0)}\nabla_{q^{(k)}}\psi(x^{(0)})\right)^T\nabla_{q^{(k)}}\psi(x^{(0)}) + \phi'(u^{(0)})\left\|\nabla_{q^{(k)}}\psi(x^{(0)})\right\|^2. \tag{26}$$

As $\mathcal{F}\left(x^{(0)},.\right)$ is strictly increasing at $u^{(0)}$, we have that

$$\frac{\partial_q \mathcal{F}}{\partial_q u}\left(x^{(0)},u^{(0)}\right) > 0.$$

From (26) we observe that $\frac{\partial_q \mathcal{F}}{\partial_q u}(.,.)$, is continuous at $\left(x^{(0)}, u^{(0)}\right)$ and all the hypotheses of Theorem 2 hold. Thus u is continuous at $x^{(0)}$.

□

We present the following Algorithm 1.

Algorithm 1: q-Gradient Descent (q-GD) Algorithm

Let objective function $\psi : \mathbb{R}^n \to \mathbb{R}$, $\phi : \mathbb{R}_+ \to \mathbb{R}$, the q-descent direction $d^{(0)}_{q^{(0)}}$ of the search at the starting point $x^{(0)} \in \mathbb{R}^n$;
fix $q_0^{(0)} \in (0,1)$ and $\epsilon = 10^{-6}$;
Set $s_0 \in (\delta_1, \delta_2)$ where numbers δ_1 and δ_2 are positive ;
for $k = 0, 1, 2, \ldots$ **do**
  Set $d^{(k)}_{q^{(k)}} = -\nabla_{q^{(k)}} \psi(x^{(k)})$;
  if $\left\| \nabla_{q^{(k)}} \psi(x^{(k)}) \right\| < \epsilon$ **then**
    $x^* = x^{(k)}$;
    $\psi^* = \psi\left(x^{(k)}\right)$;
    stop;
  else
    for $j = 0,1,2, \ldots$ **do**
      if $\psi\left(x^{(j)} + \alpha_k d^{(j)}_{q^{(j)}}\right) > \psi\left(x^{(j)}\right) + \phi(s_j)\left(d^{(j)}_{q^{(j)}}\right)^T g^{(j)}_{q^{(j)}}$ **then**
        Set $s_{j+1} = \frac{s_j}{1+q^{(j)}}$;
        else
          $\alpha_j = s_j$;
          break;
        end
      end
    end
    Set $x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}_{q^{(k)}}$;
  end
  for $i = 0, 1, \ldots, n$ **do**
    Set $q_i^{(k+1)} = 1 - \frac{q_i^{(k)}}{(k+1)^2}$;
  end
end
Result: minimizer x^* and minimum function value ψ^*.

Remark 1 ([36]). *Two important facts about computing α_k are presented below:*

1. *In Algorithm 1, the modified backtracking technique finds α_k using only one inequality instead of two inequalities required in [46]*
2. *We can find α_k by another technique; we take positive numbers δ_1 and δ_2 such that*

$$\frac{\mathcal{L}}{1 + \widehat{M}_k} \delta_1 < 1 - \delta_2, \tag{27}$$

where

$$\widehat{M}_k = \max_{1 \le i \le n} \left\{ q_i^{(k)} \right\},$$

so that step length α_k can be computed using

$$\delta_1 \le \alpha_k \le \frac{1}{\mathcal{L}} \left(1 + \widehat{M}_k\right)(1 - \delta_2), \tag{28}$$

here $\mathcal{L} > 0$.

Note that we start our Algorithm 1 by taking s_0 such that

$$\delta_1 < s_0 < \delta_2, \tag{29}$$

where δ_1 and δ_2 are two positive numbers. This proposed algorithm is well defined, then it must be established that the following inequality used in Algorithm 1,

$$\psi\left(x^{(k)} - s_k \nabla_{q^{(k)}} \psi\left(x^{(k)}\right)\right) \leq \psi(x^{(k)}) - \phi(s_k)\left\|\nabla_{q^{(k)}} \psi\left(x^{(k)}\right)\right\|^2, \tag{30}$$

is satisfied after some finite number of steps. Note that every accumulation point of $\{x^{(k)}\}$ is a minimizer of ψ. Since $\{\psi(x^{(k)})\}$ is non-increasing we have

$$\psi(x^*) \leq \psi\left(x^{(k)}\right),$$

for all k with fixed x^*. The content of Theorem 5 will be argued later. The analysis of backtracking used in Algorithm 1 to compute α_k is shown below due to [36].

Proposition 3. *The backtracking method of Algorithm 1 defined by Equations (29) and (30) stops after a finite number of iterations with*

$$\min\left\{\delta_1\,,\ \frac{\varrho\left(x^{(k)}\right)}{1+\widehat{M}_k}\right\} \leq \alpha_k \leq \min\left\{\delta_2\,,\ \varrho\left(x^{(k)}\right)\right\}. \tag{31}$$

Proof. We present two cases for the value of s_0:

1. $s_0 \in \left(0, \varrho\left(x^{(k)}\right)\right)$,
2. $s_0 \geq \varrho\left(x^{(k)}\right)$.

1. From Equations (29) and (30), and (1) of Proposition 2, $\alpha_k = s_0$ and iteration stops at $j = 0$. Since $s_0 < \delta_2$, and $s_0 < \varrho(x^{(k)})$, then

$$\alpha_k = s_0 < \min\left\{\delta_2\,,\ \varrho\left(x^{(k)}\right)\right\},$$

and since $s_0 > \delta_1$, then

$$s_0 = \alpha_k \geq \min\left\{\delta_1\,,\ \frac{\varrho\left(x^{(k)}\right)}{1+\widehat{M}_k}\right\}.$$

2. There exists a unique $t \in \mathbb{N}$, where $t \geq 1$ such that

$$\left(1+\widehat{M}_k\right)^{t-1}\varrho\left(x^{(k)}\right) < s_0 \leq \left(1+\widehat{M}_k\right)^{t}\varrho\left(x^{(k)}\right). \tag{32}$$

Then,

$$\frac{\varrho\left(x^{(k)}\right)}{1+\widehat{M}_k} < \frac{s_0}{\left(1+\widehat{M}_k\right)^{t}} \leq \varrho\left(x^{(k)}\right). \tag{33}$$

From Equations (29) and (30), we have

$$s_j = s_0\left(1+\widehat{M}_k\right)^{-j},$$

used in Algorithm 1 so the above inequality establishes that

$$\frac{\varrho\left(x^{(k)}\right)}{1+\widehat{M}_k} < s_t \leq \varrho\left(x^{(k)}\right). \tag{34}$$

We now claim that $\alpha_k = s_t$. From Equation (34), we obtain $s_t \leq \varrho(x^{(k)})$, and if we assume

$$s_{t-1} > \varrho(x^{(k)}),$$

then for $t = 1$, we obtain Case 2, so this assumption is true. Using Proposition 2, and $\alpha_k = s_t$, the inequality (30) is satisfied, but this does not satisfy if $\alpha_k = s_{t-1}$. Note that (31) follows from (34) and in fact, we have that $s_t \leq s_0 < \delta_2$.

The proof is complete. □

The following proposition states that the q-gradient descent method moves in orthogonal steps whose proof is very similar to the proof of [2] (Proposition 6.1).

Proposition 4. *If $\{x^{(k)}\}_{k=0}^{\infty}$ is a q-gradient descent sequence for a given function $\psi : \mathbb{R}^n \to \mathbb{R}$, then for every k, the vector*

$$x^{(k+1)} - x^{(k)},$$

is orthogonal to the vector $x^{(k+2)} - x^{(k+1)}$.

Proof. The iterative formula of the q-gradient descent is:

$$x^{(k+1)} - x^{(k)} = -\alpha_k \nabla_{q^{(k)}} \psi\left(x^{(k)}\right), \tag{35}$$

For $k = k + 1$, we obtain

$$x^{(k+2)} - x^{(k+1)} = -\alpha_{k+1} \nabla_{q^{(k)}} \psi\left(x^{(k+1)}\right). \tag{36}$$

From (35) and (36), we obtain

$$\left\langle x^{(k+1)} - x^{(k)}, x^{(k+2)} - x^{(k+1)} \right\rangle = \left\langle -\alpha_k \nabla_{q^{(k)}} \psi\left(x^{(k)}\right), -\alpha_{k+1} \nabla_{q^{(k)}} \psi\left(x^{(k+1)}\right) \right\rangle. \tag{37}$$

We need to show that

$$\left\langle \nabla_{q^{(k)}} \psi\left(x^{(k)}\right), \nabla_{q^{(k)}} \psi\left(x^{(k+1)}\right) \right\rangle = 0.$$

Since $\alpha_k \geq 0$ is a minimizer of

$$\phi_k(\alpha) = \psi\left(x^{(k)} + \alpha_k d_{q^{(k)}}^{(k)}\right) - \psi\left(x^{(k)}\right) - \phi(s_k)\left(d_{q^{(k)}}^{(k)}\right)^T g_{q^{(k)}}^{(k)},$$

used in Algorithm 1, then from the first-order necessary condition, we have

$$\frac{\mathrm{d}_q \phi_k(\alpha)}{\mathrm{d}_q \alpha} = 0,$$

that is,

$$-\nabla_{q^{(k)}} \psi\left(x^{(k)} - \alpha \nabla_q \psi\left(x^{(k)}\right)\right) \nabla_{q^{(k)}} \psi\left(x^{(k)}\right) = 0,$$

that is,

$$-\nabla_{q^{(k)}} \psi\left(x^{(k+1)}\right) \left(\nabla_{q^{(k)}} \psi\left(x^{(k)}\right)\right)^T = 0,$$

We obtain the desired result after putting the value of the above equation in (37). □

The above theorem implies that $\nabla_{q^{(k)}}\psi(x^{(k)})$ is parallel to the tangent plane to the level set

$$\left\{\psi(x) = \psi\left(x^{(k+1)}\right)\right\},$$

at $x^{(k+1)}$ when $q^{(k)} \to (1,\ldots,1)^T$ as $k \to \infty$. Note that each approximate point generated by the q-gradient descent algorithm decreases the corresponding objective function ψ.

Proposition 5. *If* $\left\{x^{(k)}\right\}_{k=0}^{\infty}$ *is the q-gradient descent sequence for* $\psi : \mathbb{R}^n \to \mathbb{R}$ *and*

$$\nabla_{q^{(k)}}\psi\left(x^{(k)}\right) \neq 0,$$

then

$$\psi\left(x^{(k+1)}\right) < \psi\left(x^{(k)}\right).$$

Proof. We know that

$$x^{(k+1)} = x^{(k)} - \alpha_k \nabla_{q^{(k)}}\psi\left(x^{(k)}\right),$$

where $\alpha_k \geq 0$ is the minimizer of

$$\phi_k(\alpha) = \psi\left(x^{(k)} + \alpha_k d^{(k)}_{q^{(k)}}\right) - \psi\left(x^{(k)}\right) - \phi(s_k)\left(d^{(k)}_{q^{(k)}}\right)^T g^{(k)}_{q^{(k)}}.$$

Therefore, for $\alpha \geq 0$, we have $\phi_k(\alpha_k) \leq \phi_k(\alpha)$ for all α. For $\alpha = 0$, we have

$$\frac{d_q\phi_k}{d_q\alpha} = -\left\|\nabla_{q^{(k)}}\psi\left(x^{(k)}\right)\right\|^2.$$

Since $\nabla_{q^{(k)}}\psi\left(x^{(k)}\right) \neq 0$, then $-\left\|\nabla_{q^{(k)}}\psi\left(x^{(k)}\right)\right\|^2 < 0$. Thus,

$$\frac{\mathrm{d}_q\phi_k}{\mathrm{d}_q\alpha} < 0.$$

There exists an $\overline{\alpha} > 0$ such that $\phi_k(0) > \phi_k(\alpha)$, for each $\alpha \in (0, \overline{\alpha}]$. That is,

$$\psi\left(x^{(k+1)}\right) = \phi_k(\alpha_k) \leq \phi_k(\alpha) < \phi_k(0) = \psi\left(x^{(k)}\right).$$

Thus,

$$\psi\left(x^{(k+1)}\right) < \psi\left(x^{(k)}\right).$$

This completes the proof. □

4. Convergence Analysis

In this section, we present the convergence analysis of the proposed method with the inexact line searches and the concept of quasi Fejér convergence.

Theorem 3. *[36] If* $\left\{x^{(k)}\right\}$ *is quasi-Fejér convergence to a nonempty set* $\mathcal{U} \subseteq \mathbb{R}^n$*, then* $\{x^{(k)}\}$ *is bounded. Furthermore, if an accumulation point* x *of* $\{x^{(k)}\}$ *belongs to* $\mathcal{U}$ *then* $\lim_{k\to\infty} x^{(k)} = x$.

Proof. The proof of the above theorem can be seen in [36]. □

Theorem 4. *For Algorithm 1 it holds that*

1. *There exists $\gamma > 0$ such that*

$$\psi\left(x^{(k+1)}\right) \leq \psi\left(x^{(k)}\right) - \gamma\left\|x^{(k+1)} - x^{(k)}\right\|^2,$$

for all $k \in \mathbb{N}$,

2. *The $\{\psi(x^{(k)})\}$ is nonincreasing and convergent,*

3. $\sum_{k=0}^{\infty}\left\|x^{(k+1)} - x^{(k)}\right\|^2 < \infty.$

Proof.

1. For Algorithm 1, we compute α_k using inequalities (27) and (29). Note that

$$\left\|x^{(k+1)} - x^{(k)}\right\|^2 = \alpha_k^2\left\|\nabla_q\psi\left(x^{(k)}\right)\right\|^2.$$

Since $\alpha_k > 0$, we also have

$$\frac{\phi(\alpha_k)}{\alpha_k^2}\left\|x^{(k+1)} - x^{(k)}\right\|^2 = \phi(\alpha_k)\left\|\nabla_q\psi\left(x^{(k)}\right)\right\|^2.$$

Since

$$\phi(\alpha_k)\left\|\nabla_q\psi\left(x^{(k)}\right)\right\|^2 \leq \psi\left(x^{(k)}\right) - \psi\left(x^{(k+1)}\right),$$

thus

$$\frac{\phi(\alpha_k)}{\alpha_k^2}\left\|x^{(k+1)} - x^{(k)}\right\|^2 \leq \psi\left(x^{(k)}\right) - \psi\left(x^{(k+1)}\right). \tag{38}$$

We take

$$0 < \xi < \lim_{v\to 0^+}\frac{\inf\phi(u)}{u^2}.$$

By definition of ξ, there exists $\theta > 0$ such that if for a general case, the value of step length $\alpha \in (0, \theta)$, then

$$\xi < \frac{\phi(\alpha)}{\alpha^2}. \tag{39}$$

However, for every k, we take the following two cases for choosing the step length as:

(i) If $\alpha_k \in (0, \theta)$, then from (39), we have $\frac{\phi(\alpha_k)}{\alpha_k^2}$.

(ii) If $\alpha_k \geq \theta$. In this case, by Proposition 3, we have that

$$\alpha_k \leq \min\left\{\delta_2\,,\, \varrho\left(x^{(k)}\right)\right\} \leq \delta_2,$$

and it follows from (1) and (2) of Assumption 2 that ϕ is increasing, implying $\phi(\theta) \leq \phi(\alpha_k)$. Thus, we have

$$\frac{\phi(\theta)}{\delta_2^2} \leq \frac{\phi(\alpha_k)}{\alpha_k^2}. \tag{40}$$

From (38), we have

$$\psi\left(x^{(k+1)}\right) \leq \psi\left(x^{(k)}\right) - \frac{\phi(\alpha_k)}{\alpha_k^2}\left\|x^{(k+1)} - x^{(k)}\right\|^2.$$

From Equations (39) and (40) and using above inequality, we obtain

$$\psi\left(x^{(k+1)}\right) \leq \psi\left(x^{(k)}\right) - \gamma\left\|x^{(k+1)} - x^{(k)}\right\|^2,$$

where

$$\gamma = \min\left\{\xi, \frac{\phi(\theta)}{\delta_2^2}\right\}.$$

Thus, (1) of this theorem is proved. If we use Equations (29) and (30) to compute the step length then from q-Newton–Leibniz formula [42]

$$\begin{aligned}\psi\left(x^{(k+1)}\right) = \psi\left(x^{(k)}\right) &- \alpha_k\left\|\nabla_{q^{(k)}}\psi\left(x^{(k)}\right)\right\|^2 \\ &- \alpha_k\int_0^1\Big[\nabla_{q^{(k)}}\psi\left(x^{(k)} - u\alpha_k\nabla_{q^{(k)}}\psi\left(x^{(k)}\right)\right) \\ &-\nabla_{q^{(k)}}\psi\left(x^{(k)}\right)\Big]^T\nabla_{q^{(k)}}\psi\left(x^{(k)}\right)\mathrm{d}_q u.\end{aligned}$$

From (2) of Assumption 1, we obtain

$$\begin{aligned}\psi\left(x^{(k+1)}\right) \leq \psi\left(x^{(k)}\right) &- \alpha_k\left\|\nabla_{q^{(k)}}\psi\left(x^{(k)}\right)\right\|^2 \\ &+ \mathcal{L}\alpha_k^2\left\|\nabla_{q^{(k)}}\psi\left(x^{(k)}\right)\right\|^2\int_0^1 u\,\mathrm{d}_q u.\end{aligned}$$

Using a special case of the q-integral [44], we obtain

$$\psi\left(x^{(k+1)}\right) \leq \psi\left(x^{(k)}\right) - \alpha_k\left(1 - \frac{\mathcal{L}\alpha_k}{1+\widehat{M}_k}\right)\left\|\nabla_{q^{(k)}}\psi\left(x^{(k)}\right)\right\|^2,$$

where we assume

$$\widehat{M}_k = \max_{1\leq i\leq n}\left\{q_i^{(k)}\right\}.$$

Since

$$\left\|\nabla_{q^{(k)}}\psi\left(x^{(k)}\right)\right\|^2 = \frac{1}{\alpha_k^2}\left\|x^{(k+1)} - x^{(k)}\right\|^2,$$

then

$$\psi\left(x^{(k+1)}\right) \leq \psi\left(x^{(k)}\right) - \frac{1}{\alpha_k}\left(1 - \frac{\mathcal{L}\alpha_k}{1+\widehat{M}_k}\right)\left\|x^{(k+1)} - x^{(k)}\right\|^2. \tag{41}$$

From (28), we have

$$\delta_1 \leq \alpha_k \leq \frac{1+\widehat{M}_k}{\mathcal{L}}(1-\delta_2),$$

then we obtain

$$\frac{1}{\alpha_k}\left(1 - \frac{\mathcal{L}\alpha_k}{1+\widehat{M}_k}\right) \geq \delta_2\frac{\mathcal{L}}{\left(1+\widehat{M}_k\right)(1-\delta_2)}.$$

From (41), we obtain

$$\psi\left(x^{(k+1)}\right) \leq \psi\left(x^{(k)}\right) - \delta_2\frac{\mathcal{L}}{\left(1+\widehat{M}_k\right)(1-\delta_2)}\left\|x^{(k+1)} - x^{(k)}\right\|^2.$$

where

$$\delta_2\frac{\mathcal{L}}{\left(1+\widehat{M}_k\right)(1-\delta_2)} = \gamma.$$

Thus, (1) is proved.

2. It follows from (1) using $\gamma > 0$.

3. By (1), there exist $\gamma > 0$ such that

$$\sum_{k=0}^{r} \left\| x^{(k+1)} - x^{(k)} \right\|^2 \leq \frac{1}{\gamma} \left(\psi\left(x^{(0)}\right) - \psi\left(x^{(r)}\right) \right) \leq \frac{1}{\delta} \left(\psi\left(x^{(0)}\right) - \psi(x^*) \right).$$

Suppose that $r \to \infty$, we obtain

$$\sum_{k=0}^{\infty} \left\| x^{(k+1)} - x^{(k)} \right\|^2 < \infty.$$

This completes the proof. □

We also need to present the following proposition to prove the convergence of Algorithm 1.

Proposition 6. *Let*

$$\mathcal{T} = \left\{ \mathfrak{z} \in \mathbb{R}^n \, : \, \psi(\mathfrak{z}) \leq \lim_{k \to \infty} \inf \psi\left(x^{(k)}\right) \right\}. \tag{42}$$

Every point generated by Algorithm 1 is placed in T, making T nonempty. For any $\mathfrak{z} \in \mathcal{T}$*, we have*

$$\left\| x^{(k+1)} - \mathfrak{z} \right\|^2 \leq \left\| x^{(k)} - \mathfrak{z} \right\|^2 + \left\| x^{(k+1)} - x^{(k)} \right\|^2,$$

then the sequence $\{x^{(k)}\}$ *generated by Algorithm 1 is a quasi-Fejér convergence to a point* $x^* \in \mathcal{T}$ *with any* $\alpha_k > 0$*.*

Proof. Given that $\mathfrak{z} \in \mathcal{T}$, then

$$\left\| x^{(k+1)} - \mathfrak{z} \right\|^2 - \left\| x^{(k)} - \mathfrak{z} \right\|^2 - \left\| x^{(k+1)} - x^{(k)} \right\|^2 = -2\left(\mathfrak{z} - x^{(k)} \right)^T \left(x^{(k+1)} - x^{(k)} \right).$$

Since $x^{(k+1)} - x^{(k)} = -\alpha_k \nabla_{q^{(k)}} \psi(x^{(k)})$, then

$$\left\| x^{(k+1)} - \mathfrak{z} \right\|^2 - \left\| x^{(k)} - \mathfrak{z} \right\|^2 - \left\| x^{(k+1)} - x^{(k)} \right\|^2 = 2\alpha_k \left(\mathfrak{z} - x^{(k)} \right)^T \nabla_{q^{(k)}} \psi\left(x^{(k)} \right).$$

A function will be called a q-convex function if ψ satisfies the following inequality in the light of q-calculus:

$$\psi(\mathfrak{z}) - \psi\left(x^{(k)}\right) \geq \left(\mathfrak{z} - x^{(k)} \right)^T \nabla_{q^{(k)}} \psi\left(x^{(k)} \right).$$

With the above inequality and (42), we obtain

$$\left\| x^{(k+1)} - \mathfrak{z} \right\|^2 - \left\| x^{(k)} - \mathfrak{z} \right\|^2 - \left\| x^{(k+1)} - x^{(k)} \right\|^2 \leq 2\alpha_k \left(\psi(\mathfrak{z}) - \psi\left(x^{(k)}\right) \right) \leq 0,$$

and from (3) of Theorem 2, we have

$$\sum_{k=0}^{\infty} \left\| x^{(k+1)} - x^{(k)} \right\|^2 < \infty,$$

then we have that $\left\{ x^{(k)} \right\}$ is a quasi-Fejér convergence to $\mathcal{T} \subseteq \mathbb{R}^n$ with

$$\left\| x^{(k+1)} - x^{(k)} \right\|^2 = \epsilon_k,$$

and from Theorem 3, $\{x^{(k)}\}$ is bounded. Further, it has an accumulation point x of $\{x^{(k)}\}$, which is in $\mathcal{T}$. Thus, $\lim_{k\to\infty} x^{(k)} = x$. □

Theorem 5. *The sequence $\{x^{(k)}\}$ generated by Algorithm 1 converges in the sense of quasi-Fejér convergence to a minimizer of function $\psi : \mathbb{R}^n \to \mathbb{R}$.*

Proof. From Proposition 6, $\lim_{k\to\infty} x^{(k)} = x^* \in \mathcal{T}$, where T is a set of accumulation points that are responsible for decreasing the objective function in every iteration. However, we need to prove that $x^* \in \mathcal{T}^*$, where T^* is a set of minimizers that minimize the objective function. From Algorithm 1, suppose that $x^* \notin \mathcal{T}^*$, then from the convexity of ψ, $x^* \in \mathcal{G}$ and

$$\left\| \nabla_{q^{(k)}} \psi(x^*) \right\| > 0.$$

From Proposition 2, $\varrho(x^*) > 0$ and $\varrho\left(x^{(k)}\right)$ converges to $\varrho(x^*)$. Thus, there exists k_0 such that for all $k \geq k_0$, we have

$$v(x^{(k)}) \geq \frac{\varrho(x^*)}{1 + \widehat{M}_k}, \tag{43}$$

$$\left\| \nabla_q \psi\left(x^{(k)}\right) \right\|^2 \geq \frac{1}{1 + \widehat{M}_k} \left\| \nabla_{q^{(k)}} \psi(x^*) \right\|^2. \tag{44}$$

Let

$$\sigma = \left(\min\left\{ \delta_1 , \frac{\varrho(x^*)}{1 + \widehat{M}_k} \right\} \right)^2 \frac{\left\| \nabla_q \psi(x^*) \right\|^2}{1 + \widehat{M}_k}.$$

Then, for any $k \geq k_0$,

$$\left\| x^{(k+1)} - x^{(k)} \right\|^2 = \alpha_k^2 \left\| \nabla_{q^{(k)}} \psi\left(x^{(k)}\right) \right\|^2.$$

Since

$$\alpha_k^2 \geq \left(\min\left\{ \delta_1 , \frac{\varrho\left(x^{(k)}\right)}{1 + \widehat{M}_k} \right\} \right)^2,$$

then

$$\left\| x^{(k+1)} - x^{(k)} \right\|^2 \geq \left(\min\left\{ \delta_1 , \frac{\varrho(x^{(k)})}{1 + \widehat{M}_k} \right\} \right)^2 \left\| \nabla_{q^{(k)}} \psi\left(x^{(k)}\right) \right\|^2. \tag{45}$$

From Equations (43) and (45), we obtain

$$\left\| x^{(k+1)} - x^{(k)} \right\|^2 \geq \left(\min\left\{ \delta_1 , \frac{\varrho(x^*)}{\left(1 + \widehat{M}_k\right)^2} \right\} \right)^2 \frac{\left\| \nabla_{q^{(k)}} \psi(x^*) \right\|^2}{\widehat{M}_k} = \sigma > 0.$$

This contradicts (3) of Theorem 4. Thus it is proved that $x^* \in \mathcal{T}^*$. If we choose α_k using (27) and (28), then we have

$$\left\| x^{(k+1)} - x^{(k)} \right\|^2 = \alpha_k^2 \left\| \nabla_{q^{(k)}} \psi\left(x^{(k)}\right) \right\|^2.$$

Since $\alpha_k^2 \geq \delta_1^2$, then

$$\left\| x^{(k+1)} - x^{(k)} \right\|^2 \geq \delta_1^2 \left\| \nabla_{q^{(k)}} \psi\left(x^{(k)}\right) \right\|^2.$$

From (3) of Theorem 4, the continuity of $\nabla_{q^{(k)}} f(.)$, and $\delta_1 > 0$, used in the above inequality, we obtain $\nabla_{q^{(k)}} \psi(x^{(k)}) \to 0$ for $k \to \infty$. Moreover, as $q^{(k)} \to (1, \ldots, 1)^T$, this implies that

$$\nabla_{q^{(k)}} \psi\left(x^{(k)}\right),$$

is a good approximation of $\nabla \psi\left(x^{(k)}\right)$ [2]. Further, see the proof of Proposition 2.2 in [47] for an affine function where the classical gradient and q-gradient of ψ are the same. For other functions, Example 3.2 and Remark 3.3 can be seen again in [47]. Hence, the accumulation point $x^* = x^{(k)}$ is a minimizer of the function. □

5. Experimental Results

We compared the numerical performance of our algorithm with the methodology used in [36]. The stopping criteria were set as

$$\left\| \nabla_{q^{(k)}} \psi\left(x^{(k)}\right) \right\| \leq \epsilon,$$

where $\epsilon = 10^{-6}$ to terminate both the algorithms. Numerical results were compared based on the number of iterations and number of function evaluations. The iteration was stopped either when it satisfied the stopping criteria or when iteration counts were 500. All problems were taken from [48] and computer codes were written in R language.

The numerical experiments were performed on an Intel Core i5-3210M CPU with 2.5 GHz, 4 GB of RAM and a 64-bit XXX (Intel, Santa Clara, CA, USA) to solve the unconstrained minimization problems.

Example 2 ([49]). *Consider a function $\psi : \mathbb{R}^3 \to \mathbb{R}$ is given by*

$$\psi(x) = \frac{1}{2} x^T Q x - b^T x,$$

where

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \tau & 0 \\ 0 & 0 & \tau \end{bmatrix}, \qquad b = -\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

We apply a q-gradient descent algorithm with a starting point

$$x^{(0)} = (-3, -2, -1)^T,$$

on different values of $\tau = 2, 5, 10, 20, 50$. We compared our method with the methodology used in [36] and the numerical results are described in Table 1.

It is worth mentioning that the method given in this paper generates the least number of iterations, and minimizer x^* and minimum function value $\psi(x^*) = \psi^*$ are almost the same for both methods.

Table 1. Numerical results of Example 2.

τ	$\bar{x}$	q-Gradient Descent		
		$\psi(\bar{x})$	GN	IT
2	$(-1.001E+00, -5.000E-01, -2.501E-01)^T$	$-8.750E-01$	$3.683E-07$	16
5	$(-1.002E+00, -2.000E-01, -4.019E-02)^T$	$-6.200E-01$	$8.270E-07$	56
10	$(-1.003E+00, -1.000E-01, -9.991E-03)^T$	$-5.550E-01$	$8.859E-07$	97
20	$(-1.004E+00, -5.000E-02, -2.499E-03)^T$	$-5.262E-01$	$9.984E-07$	139
50	$(-1.004E+00, -2.000E-02, -3.999E-04)^T$	$-5.102E-01$	$9.913E-07$	353
		Classical Gradient Descent [36]		
		$\psi(\bar{x})$	GN	IT
2	$(-1.001E+00, -5.000E-01, -2.498E-01)^T$	$-8.750E-01$	$4.068E-07$	15
5	$(-1.003E+00, -2.000E-01, -3.993E-02)^T$	$-6.200E-01$	$8.345E-07$	71
10	$(-1.004E+00, -1.000E-01, -1.012E-02)^T$	$-5.550E-01$	$9.782E-07$	134
20	$(-1.006E+00, -5.000E-02, -2.497E-03)^T$	$-5.262E-01$	$9.640E-07$	215
50	$(-1.010E+00, -2.000E-02, -4.272E-04)^T$	$-5.101E-01$	$9.843E-07$	588

Example 3 ([50]). *Consider a function* $\psi : \mathbb{R}^2 \to \mathbb{R}$ *such that*

$$\psi(x) = x_1^4 - 2x_1^2x_2 + x_1^2 + x_2^2 - 2x_1 + 1.$$

With a starting point

$$x^{(0)} = (-2, -2)^T,$$

Algorithm 1 shows improvement over the algorithm given in [36]. The numerical results are shown in the following Tables 2 and 3, with abbreviations used in columns as 'gn' norm of gradient, 'nf' number of function evaluations, and step length computed using the backtracking technique. We observe that our proposed algorithm converges to the solution point in 28 iterations while the algorithm used in [36] converges to the same solution point in 40 iterations. The graphs for both methods in terms of fe versus the logarithm of the function value are provided in Figures 1 and 2. The three-dimensional pictorial representation of Example 3 is shown in Figure 3.

Table 2. Numerical output of Example 3 using Algorithm 1.

	q-Gradient Descent		
It	ψ	$g2n$	nf
0	$4.50000E+01$	$5.53114E+01$	1
1	$3.05531E+00$	$3.83381E+00$	4
2	$1.44344E+00$	$2.45969E+00$	7
3	$5.37577E-01$	$1.48052E+00$	10
4	$3.32689E-01$	$8.22955E-01$	12
5	$2.09224E-01$	$7.95222E-01$	15
6	$1.50867E-01$	$5.00475E-01$	17
7	$1.10048E-01$	$5.35543E-01$	20
8	$8.48946E-02$	$3.57301E-01$	22
9	$6.60685E-02$	$3.64658E-01$	24
10	$5.33656E-02$	$2.90936E-01$	26
11	$4.33762E-02$	$2.65104E-01$	28
12	$3.57854E-02$	$2.51345E-01$	30
13	$2.96786E-02$	$1.99254E-01$	32
14	$2.45786E-02$	$2.26142E-01$	34
15	$2.04572E-02$	$1.51693E-01$	36
16	$1.66695E-02$	$2.10239E-01$	38
17	$1.36586E-02$	$1.14514E-01$	40
18	$1.05164E-02$	$2.01242E-01$	42
19	$8.15978E-03$	$8.24237E-02$	44
20	$5.09294E-03$	$2.12029E-01$	45
21	$2.88677E-03$	$4.64793E-02$	47
22	$4.05728E-04$	$8.80875E-02$	50
23	$6.74352E-05$	$7.16943E-03$	52
24	$1.98233E-05$	$1.75357E-02$	55
25	$5.78355E-06$	$2.33610E-03$	57
26	$3.56220E-06$	$6.21160E-03$	60
27	$1.61834E-06$	$1.41408E-03$	62
28	$1.32663E-06$	$3.28145E-03$	65

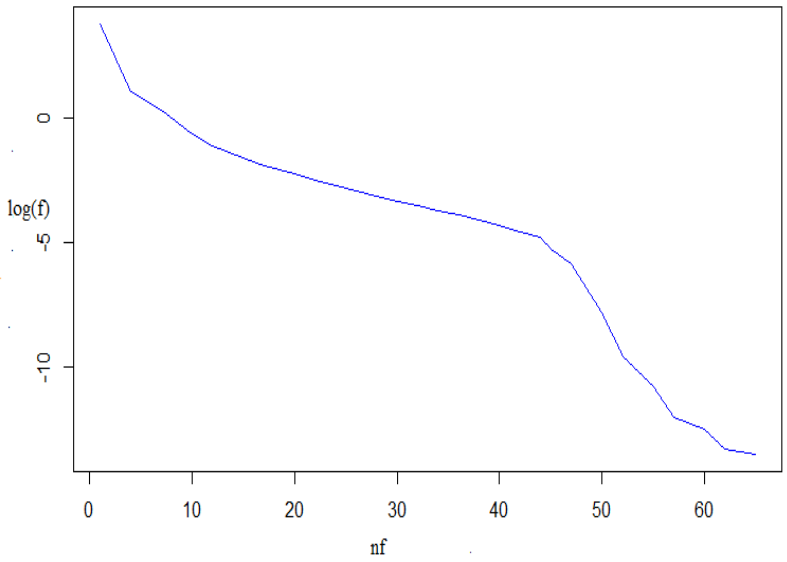

Figure 1. Number of function evaluations due to q-gradient descent algorithm for Example 3.

Table 3. Numerical results of Example 3 using Classical Gradient Descent [36].

	Classical Gradient Descent		
It	ψ	$g2n$	nf
0	$4.50000E+01$	$5.53173E+01$	1
1	$3.05623E+00$	$3.83810E+00$	4
2	$1.44520E+00$	$2.46074E+00$	7
3	$5.38450E-01$	$1.48336E+00$	10
4	$3.33100E-01$	$8.23444E-01$	12
5	$2.09391E-01$	$7.95583E-01$	15
6	$1.51000E-01$	$5.00572E-01$	17
7	$1.10457E-01$	$4.84543E-01$	19
8	$8.61656E-02$	$3.86837E-01$	21
9	$6.78803E-02$	$3.30047E-01$	23
10	$5.44785E-02$	$3.30855E-01$	25
11	$4.40901E-02$	$2.35325E-01$	27
12	$3.51020E-02$	$3.02819E-01$	29
13	$2.81958E-02$	$1.68607E-01$	31
14	$2.07562E-02$	$2.94811E-01$	33
15	$1.54894E-02$	$1.13592E-01$	35
16	$6.31703E-03$	$2.74534E-01$	37
17	$2.67696E-03$	$4.64201E-02$	39
18	$1.88478E-03$	$1.04102E-01$	41
19	$1.35417E-03$	$3.27922E-02$	43
20	$9.76185E-04$	$7.34138E-02$	45
21	$7.13476E-04$	$2.36981E-02$	47
22	$5.22365E-04$	$5.29427E-02$	49
23	$3.86176E-04$	$1.73832E-02$	51
24	$2.85774E-04$	$3.87568E-02$	53
25	$2.12958E-04$	$1.28821E-02$	55
26	$1.58791E-04$	$2.86713E-02$	57
27	$1.19008E-04$	$9.61596E-03$	59
28	$8.92260E-05$	$2.13709E-02$	61
29	$6.71499E-05$	$7.21554E-03$	63
30	$5.05489E-05$	$1.60174E-02$	65
31	$3.81590E-05$	$5.43511E-03$	67
32	$2.88110E-05$	$1.20540E-02$	69
33	$2.17989E-05$	$4.10562E-03$	71
34	$1.64954E-05$	$9.09882E-03$	73
35	$1.25020E-05$	$3.10790E-03$	75
36	$9.47621E-06$	$6.88380E-03$	77
37	$7.19130E-06$	$2.35637E-03$	79
38	$5.45769E-06$	$5.21692E-03$	81
39	$4.14574E-06$	$1.78870E-03$	83
40	$3.14931E-06$	$3.95878E-03$	85

Dolan and Moré [51] presented an appropriate technique to demonstrate the performance profiles, which is a statistical process. The performance ratio is presented as:

$$\rho_{p,s} = \frac{r_{(p,s)}}{\min\left\{r_{(p,s)} : 1 \leq r \leq n_s\right\}}, \tag{46}$$

where $r_{(p,s)}$ refers to the iteration and function evaluations for solver s spent on problem p, and n_s refers to the number of problems in the model test. The cumulative distribution function is given as:

$$P_s(\tau) = \frac{1}{n_p}\text{size}\left\{p \in \rho_{(p,s)} \leq \tau\right\}, \tag{47}$$

where $P_s(\tau)$ is the probability that a performance ratio $\rho_{(p,s)}$ is within a factor of $\tau \in \mathbb{R}$. That is, for a subset of the methods being analyzed, we plot the fraction $\rho_s(\tau)$ of problems for which any given method is within a factor τ of the best. We use this tool to show the performance of Algorithm 1. First, we solved 28 test problems with different starting points and stored the number of iterations and number of function evaluation values in Table 4. In fact, Figures 4 and 5 show that the q-Gradient Descent method solves about 86% and 79% of 28 test problems [48] with the least number of iterations and function evaluations, respectively. We can conclude that the proposed method is superior.

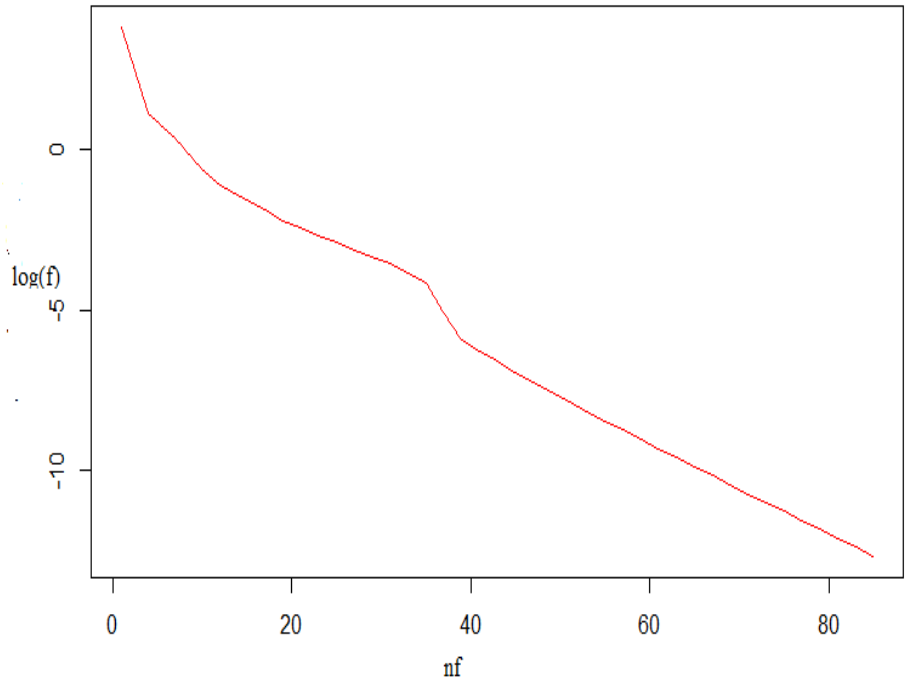

Figure 2. Number of function evaluations due to Classical Gradient Descent algorithm for Example 3.

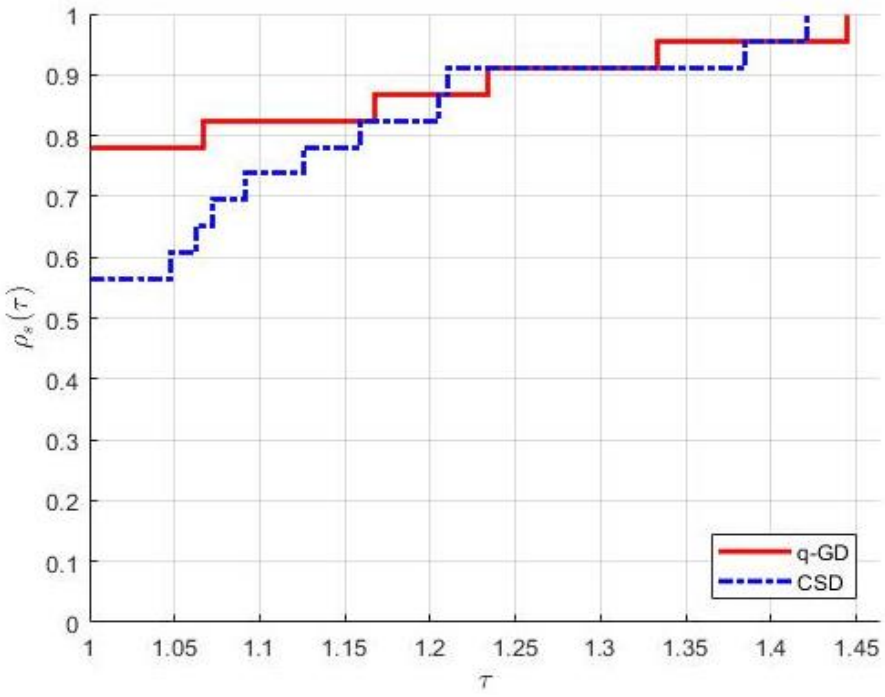

Figure 3. Performance profile for number of function evaluations based on Table 4.

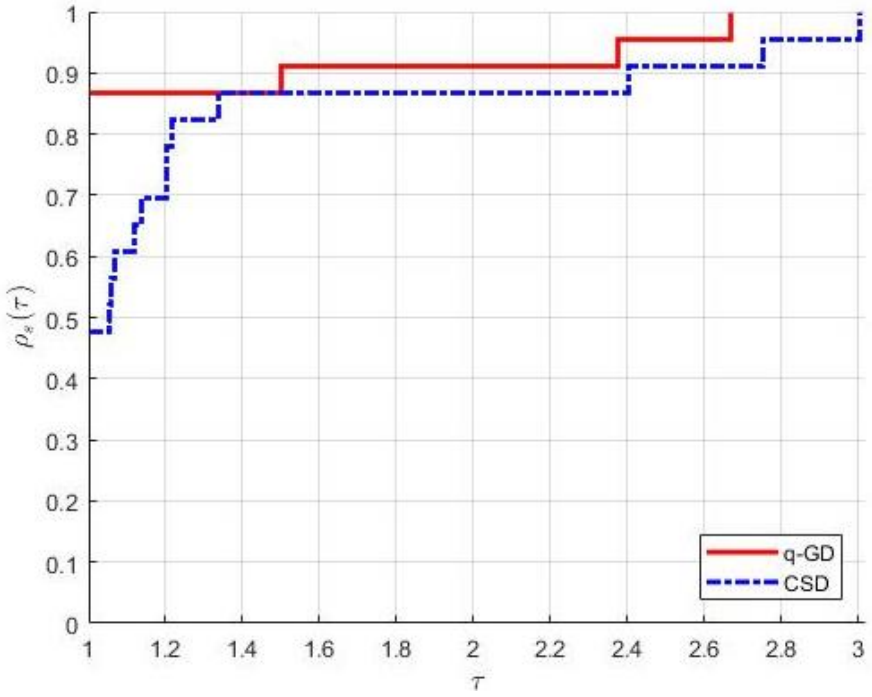

Figure 4. Performance profile for number of iterations based on Table 3.

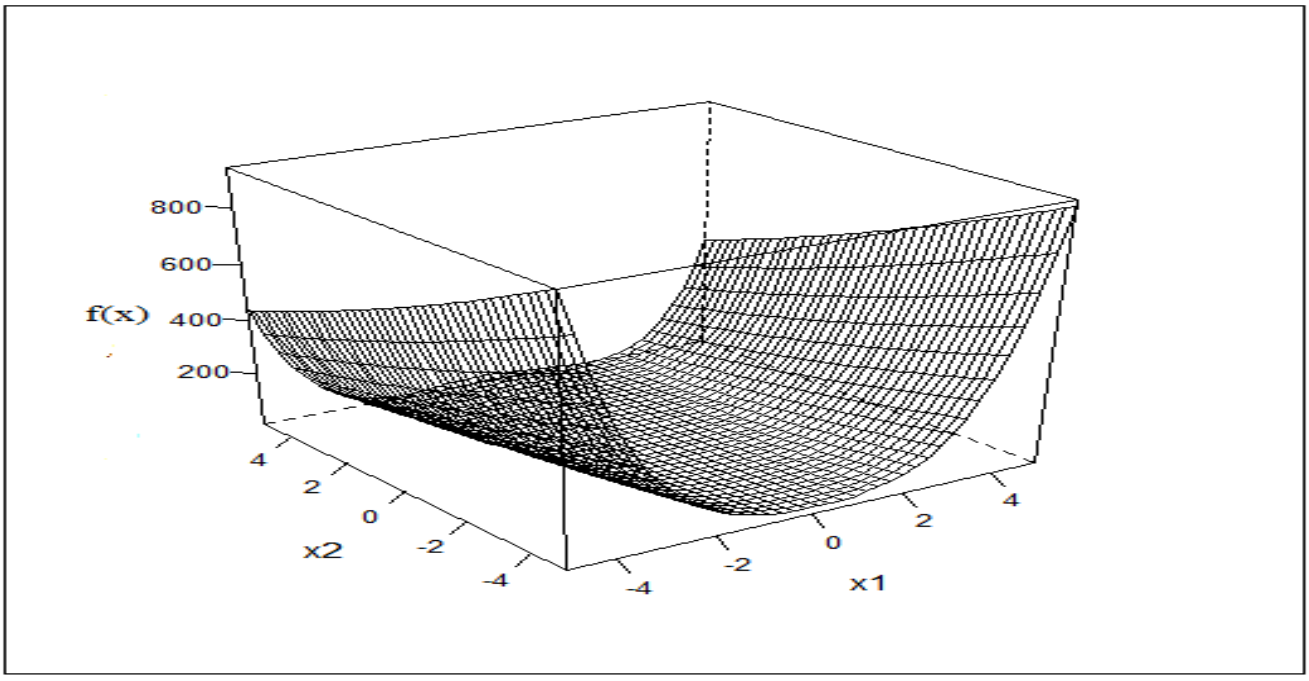

Figure 5. Graphics of Example 3.

Table 4. Numerical results of 28 test problems.

Sl. No.	Problem Name	Starting Point	q-Gradient Descent (q-GD)		Classical Gradient Descent (CSD) [36]	
			NI	*FE*	*NI*	*FE*
1	Booth	$(3.000E-01, 6.000E+00)^T$	7	16	7	15
2		Aluffi Pentini $(-1.000E+00, 7.000E-01)^T$	3	8	4	9
3		Bohachevsky $(8.000E-01, 5.000E-01)^T$	9	30	9	30
4		Branin $(1.800E+00, 1.500E+00)^T$	19	43	20	52
5	Colville	$(1.000E+00, 1.000E+00, 1.000E+00, 8.000E-01)^T$	163	347	498	1012
6	Csendes	$(-1.300E+00, 2.500E+00)^T$	5	19	6	27
7	Ackley2	$(2.300E+00, 1.600E+00)^T$	2	26	2	36
8	Csendes	$(-1.600E+00, 1.600E+00)^T$	3	14	3	15
9	Cubic	$(1.300E+00, -1.300E+00)^T$	61	197	251	645
10	Deckkers Aarts	$(2.000E+00, 1.200E+00)^T$	8	48	73	348
11	Dixon Price	$(-1.000E+00, -2.000E+00)^T$	19	44	24	53
12	Himmelblau	$(1.000E+00, 4.000E+00)^T$	10	52	11	36
13	Leon	$(8.000E-01, 1.400E+00)^T$	64	153	284	595
14	diagonal4	$(-1.300E+00, -1.500E+00, 1.300E+00, 1.500E+00)^T$	4	8	3	6
15	Zakharov	$(2.300E+00, -1.500E+00)^T$	8	18	8	18
16	FH1	$(1.000E-02, 1.000E-02)^T$	19	51	33	76
17	Zakharov	$(-2.000E+00, 2.500E+00)^T$	8	18	8	18
18	Three Hump Camel	$(3.000E+00, 5.000E+00)^T$	18	45	19	45
19	Six Hump Camel	$(-1.000E+00, -3.000E+00)^T$	9	24	9	24
20	Matyas	$(-1.000E+00, -3.000E+00)^T$	3	7	2	6
21	FH2	$(1.010E+00, 1.000E-02)^T$	22	48	25	51
22	Raydan 1	$(1.000E+00, 1.000E+00)^T$	3	9	9	9
23	Raydan 2	$(1.000E+00, -1.000E+00)^T$	4	7	4	7
24	Hager	$(1.000E+00, -1.000E+00)^T$	5	11	6	12
25	Generalized Tridiagonal 1	$(1.200E+00, 1.700E+00)^T$	26	74	29	60
26	Extended Tridiagonal 1	$(1.200E+00, -2.700E+00)^T$	28	63	34	73
27	BDEXP	$(1.000E+00, 1.000E+00, 1.000E+00, 1.000E+00)^T$	4	16	4	16
28	BDQRTIC	$(1.000E+00, 1, 1.2, 1, 1.3, 1)^T$	15	42	16	44

6. Conclusions

A q-gradient descent optimization algorithm was presented to solve unconstrained optimization problems. The approach was based on the q-gradient vector in place of the classical gradient to present the algorithm. The quasi-Fejér convergence of the algorithm was proved. The algorithm sometimes declares a fixed point that is not the solution, but it is very close to it because of using the q-derivative. Further, examples and numerical results demonstrate the improvement of the proposed method over the classical method. Our future work will include applying the q-gradient to the accelerated gradient descent method for unconstrained optimization.

Author Contributions: Conceptualization, S.K.M., P.R., S.K.C., M.K.A.K., B.R. and M.E.S.; methodology, S.K.M., P.R., S.K.C., M.K.A.K., B.R. and M.E.S.; software, M.E.S.; validation, M.K.A.K., B.R. and M.E.S.; formal analysis, S.K.M., P.R., S.K.C. and B.R.; investigation, S.K.M., P.R., S.K.C., M.K.A.K., B.R. and M.E.S.; resources, S.K.M., P.R., S.K.C., M.K.A.K., B.R. and M.E.S.; data curation, M.E.S.; writing—original draft preparation, S.K.M., P.R., S.K.C., M.K.A.K., B.R. and M.E.S.; writing—review and editing, M.K.A.K., B.R. and M.E.S.; visualization, M.E.S.; supervision, M.K.A.K. and M.E.S.; project administration, M.K.A.K. and M.E.S.; funding acquisition, M.K.A.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Acknowledgments: The first author was supported by the Science and Engineering Research Board (Grant No. DST-SERBMTR-2018/000121) and the research grant for faculty (Institute of Eminence-BHU Grant No. 3254). The second author was supported by the Ministry Education, Science and Technological Development, Grant No. 174013. The third author was supported by Bu-Ali Sina University. The authors would like to thank the editors and the reviewers for their helpful comments and suggestions which have led to the improvement of the earlier version of this article.

Conflicts of Interest: The authors declare that they have no competing interest.

References

1. Cauchy, A. Méthode générale pour la résolution des systemes d'équations simultanées. *Comp. Rend. Sci. Paris* **1847**, *25*, 536–538.
2. Mishra, S.K.; Ram, B. *Introduction to Unconstrained Optimization with R*; Springer Nature: Gateway East, Singapore, 2019.
3. El Mouatasim, A. Fast gradient descent algorithm for image classification with neural networks. *Signal Image Video Process.* **2020**, *14*, 1565–1572.. [CrossRef]
4. Chen, Y.; Miao, D. Granular regression with a gradient descent method. *Inf. Sci.* **2020**, *537*, 246–260. [CrossRef]
5. Pan, H.; Niu, X.; Li, R.; Dou, Y.; Jiang, H. Annealed gradient descent for deep learning. *Neurocomputing* **2020**, *380*, 201–211. [CrossRef]
6. Koshak, W.J.; Krider, E.P. A linear method for analyzing lightning field changes. *J. Atmos. Sci.* **1994**, *51*, 473–488. [CrossRef]
7. Liao, L.Z.; Qi, L.; Tam, H.W. A gradient-based continuous method for large-scale optimization problems. *J. Glob. Optim.* **2005**, *31*, 271–286. [CrossRef]
8. Nesterov, Y. Universal gradient methods for convex optimization problems. *Math. Program.* **2015**, *152*, 381–404. [CrossRef]
9. Nezhadhosein, S. A Modified Descent Spectral Conjugate Gradient Method for Unconstrained Optimization. *Iran. J. Sci. Technol. Trans. A Sci.* **2021**, *45*, 209–220. [CrossRef]
10. MishraSamei Mishra, S.K.; Samei, M.E.; Chakraborty, S.K.; Ram, B. On q-variant of Dai–Yuan conjugate gradient algorithm for unconstrained optimization problems. *Nonlinear Dyn.* **2021**, *104*, 2471–2496. [CrossRef]
11. Jackson, F.H. q-Difference equations. *Am. J. Math.* **1910**, *32*, 305–314. [CrossRef]
12. Aral, A.; Gupta, V.; Agarwal, R.P. *Applications of q-Calculus in Operator Theory*; Springer: New York, NY, USA, 2013.
13. Ernst, T. The different tongues of q-calculus. *Proc. Est. Acad. Sci.* **2008**, *57*, 81–99. [CrossRef]
14. Ernst, T. A method for q-calculus. *J. Nonlinear Math. Phys.* **2003**, *10*, 487–525. [CrossRef]
15. Srivastava, H. Operators of basic (or q-) calculus and fractional q-calculus and their applications in geometric function theory of complex analysis. *Iran. J. Sci. Technol. Trans. A Sci.* **2020**, *44*, 327–344. [CrossRef]
16. Piejko, K.; Sokol, J.; Trkabka-Wikeclaw, K. On q-Calculus and Starlike Functions. *Iran. J. Sci. Technol. Trans. A Sci.* **2019**, *43*, 2879–2883. [CrossRef]
17. Mursaleen, M.; Ansari, K.J.; Nasiruzzaman, M. Approximation by q-analogue of Jakimovski–Leviatan operators involving q-Appell polynomials. *Iran. J. Sci. Technol. Trans. A Sci.* **2017**, *41*, 891–900. [CrossRef]
18. Dimakis, A.; Müller-Hoissen, F. Quantum mechanics on a lattice and q-deformations. *Phys. Lett. B* **1992**, *295*, 242–248. [CrossRef]
19. Lai, K.K.; Mishra, S.K.; Panda, G.; Chakraborty, S.K.; Samei, M.E.; Ram, B. A limited memory q-BFGS algorithm for unconstrained optimization problems. *J. Appl. Math. Comput.* **2020**, 1–20. [CrossRef]
20. Samei, M.E.; Hedayati, V.; Rezapour, S. Existence results for a fraction hybrid differential inclusion with Caputo-Hadamard type fractional derivative. *Adv. Differ. Equ.* **2019**, *2019*, 163. [CrossRef]
21. Rezapour, S.; Imran, A.; Hussain, A.; Martínez, F.; Etemad, S.; Kaabar, M.K.A. Condensing Functions and Approximate Endpoint Criterion for the Existence Analysis of Quantum Integro-Difference FBVPs. *Symmetry* **2021**, *13*, 469. [CrossRef]
22. Samei, M.E.; Ghaffari, R.; Yao, S.W.; Kaabar, M.K.A.; Martínez, F.; Inc, M. Existence of Solutions for a Singular Fractional q-Differential Equations under Riemann–Liouville Integral Boundary Condition. *Symmetry* **2021**, *13*, 1235. [CrossRef]
23. Ntouyas, S.K.; Samei, M.E. Existence and uniqueness of solutions for multi-term fractional q-integro-differential equations via quantum calculus. *Adv. Differ. Equ.* **2019**, *2019*, 475. [CrossRef]
24. Samei, M.E.; Hedayati, V.; Ranjbar, G.K. The existence of solution for k-dimensional system of Langevin Hadamard-type fractional differential inclusions with 2k different fractional orders. *Mediterr. J. Math.* **2020**, *17*, 37. [CrossRef]
25. Etemad, S.; Rezapour, S.; Samei, M.E. On fractional hybrid and non-hybrid multi-term integro-differential inclusions with three-point integral hybrid boundary conditions. *Adv. Differ. Equ.* **2020**, *2020*, 161. [CrossRef]
26. Kaabar, M.K.A.; Martínez, F.; Gómez-Aguilar, J.F.; Ghanbari, B.; Kaplan, M.; Günerhan, H. New approximate analytical solutions for the nonlinear fractional Schrödinger equation with second-order spatio-temporal dispersion via double Laplace transform method. *Math. Methods Appl. Sci.* **2021**, *44*, 11138–11156. [CrossRef]
27. Alzabut, J.; Selvam, A.; Dhineshbabu, R.; Kaabar, M.K.A. The Existence, Uniqueness, and Stability Analysis of the Discrete Fractional Three-Point Boundary Value Problem for the Elastic Beam Equation. *Symmetry* **2021**, *13*, 789. [CrossRef]
28. Etemad, S.; Souid, M.S.; Telli, B.; Kaabar, M.K.A.; Rezapour, S. Investigation of the neutral fractional differential inclusions of Katugampola-type involving both retarded and advanced arguments via Kuratowski MNC technique. *Adv. Differ. Equ.* **2021**, *2021*, 1–20. [CrossRef]
29. Mohammadi, H.; Kaabar, M.K.A.; Alzabut, J.; Selvam, A.G.M.; Rezapour, S. A Complete Model of Crimean-Congo Hemorrhagic Fever (CCHF) Transmission Cycle with Nonlocal Fractional Derivative. *J. Funct. Spaces* **2021**, *2021*, 1–12. [CrossRef]
30. Matar, M.M.; Abbas, M.I.; Alzabut, J.; Kaabar, M.K.A.; Etemad, S.; Rezapour, S. Investigation of the p-Laplacian nonperiodic nonlinear boundary value problem via generalized Caputo fractional derivatives. *Adv. Differ. Equ.* **2021**, *2021*, 1–18. [CrossRef]
31. Alam, M.; Zada, A.; Popa, I.L.; Kheiryan, A.; Rezapour, S.; Kaabar, M.K.A. A fractional differential equation with multi-point strip boundary condition involving the Caputo fractional derivative and its Hyers–Ulam stability. *Bound. Value Probl.* **2021**, *2021*, 1–18. [CrossRef]
32. Rajković, P.M.; Marinković, S.D.; Stanković, M.S. On q-Newton–Kantorovich method for solving systems of equations. *Appl. Math. Comput.* **2005**, *168*, 1432–1448. [CrossRef]

33. Rajković, P.M.; Marinković, S.D.; Stanković, M. The q-gradient Method. In Proceedings of the International Symposium "Geometric Function Theory and Applications", Sofia, Bulgaria, 27–31 August 2010; pp. 240–244.
34. Kiefer, J. Sequential minimax search for a maximum. *Proc. Am. Math. Soc.* **1953**, *4*, 502–506. [CrossRef]
35. Soterroni, A.C.; Galski, R.L.; Ramos, F.M. The q-gradient vector for unconstrained continuous optimization problems. In *Operations Research Proceedings 2010*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 365–370.
36. Burachik, R.; Graña Drummond, L.; Iusem, A.N.; Svaiter, B. Full convergence of the steepest descent method with inexact line searches. *Optimization* **1995**, *32*, 137–146. [CrossRef]
37. Kiwiel, K.C.; Murty, K. Convergence of the steepest descent method for minimizing quasiconvex functions. *J. Optim. Theory Appl.* **1996**, *89*, 221–226. [CrossRef]
38. Yuan, Y.X. A new stepsize for the steepest descent method. *J. Comput. Math.* **2006**, *24*, 149–156.
39. Mishra, S.K.; Panda, G.; Ansary, M.A.T.; Ram, B. On q-Newton's method for unconstrained multiobjective optimization problems. *J. Appl. Math. Comput.* **2020**, 1–20. [CrossRef]
40. Lai, K.K.; Mishra, S.K.; Ram, B. On q-Quasi-Newton's Method for Unconstrained Multiobjective Optimization Problems. *Mathematics* **2020**, *8*, 616. [CrossRef]
41. Chakraborty, S.K.; Panda, G. Newton like line search method using q-calculus. In *International Conference on Mathematics and Computing*; Springer: Singapore, 2017; pp. 196–208.
42. Kac, V.; Cheung, P. *Quantum Calculus*; Springer Science & Business Media: New York, NY, USA, 2001.
43. Rajković, P.; Stanković, M.; Marinković, D.S. Mean value theorems in q-calculus. *Mat. Vesn.* **2002**, *54*, 171–178.
44. Andrews, G.E. *q-Series: Their Development and Application in Analysis, Number Theory, Combinatorics, Physics and Computer Algebra: Their Development and Application in Analysis, Number Theory, Combinatorics, Physics, and Computer Algebra*; Number 66; American Mathematical Soc.: Rhode Island, USA, 1986.
45. Pastor, J.R.; Calleja, P.P.; Trejo, C.A. *Análisis Matemático Volúmen: Cálculo Infinitesimal de Varias Variables Aplicaciones*; McGraw-Hill: Madrid, Spain, 1963.
46. Dennis, J.E., Jr.; Schnabel, R.B. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*; SIAM: Philadelphia, PA, USA, 1996.
47. Lai, K.K.; Mishra, S.K.; Ram, B. A q-conjugate gradient algorithm for unconstrained optimization problems. *Pac. J. Optim.* **2021**, *17*, 57–76.
48. Andrei, N. An unconstrained optimization test functions collection. *Adv. Model. Optim.* **2008**, *10*, 147–161.
49. Meza, J.C. Steepest descent. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 719–722. [CrossRef]
50. Snyman, J.; Hay, A. The spherical quadratic steepest descent (SQSD) method for unconstrained minimization with no explicit line searches. *Comput. Math. Appl.* **2001**, *42*, 169–178. [CrossRef]
51. Dolan, E.D.; Moré, J.J. Benchmarking optimization software with performance profiles. *Math. Program.* **2002**, *91*, 201–213. [CrossRef]