

Article



On *q*-Quasi-Newton's Method for Unconstrained Multiobjective Optimization Problems

Kin Keung Lai ^{1,*,†}, Shashi Kant Mishra ^{2,†} and Bhagwat Ram ^{3,†}

- ¹ College of Economics, Shenzhen University, Shenzhen 518060, China
- ² Department of Mathematics, Institute of Science, Banaras Hindu University, Varanasi 221005, India; shashikant.mishra@bhu.ac.in
- ³ DST-Centre for Interdisciplinary Mathematical Sciences, Institute of Science, Banaras Hindu University, Varanasi 221005, India; bhagwat.ram2@bhu.ac.in
- * Correspondence: mskklai@outlook.com
- + These authors contributed equally to this work.

Received: 01 April 2020; Accepted: 13 April 2020; Published: 17 April 2020

Abstract: A parameter-free optimization technique is applied in Quasi-Newton's method for solving unconstrained multiobjective optimization problems. The components of the Hessian matrix are constructed using *q*-derivative, which is positive definite at every iteration. The step-length is computed by an Armijo-like rule which is responsible to escape the point from local minimum to global minimum at every iteration due to *q*-derivative. Further, the rate of convergence is proved as a superlinear in a local neighborhood of a minimum point based on *q*-derivative. Finally, the numerical experiments show better performance.

Keywords: multiobjective programming; methods of quasi-Newton type; Pareto optimality; *q*-calculus; rate of convergence

MSC: 90C29; 90C53; 58E17; 05A30; 41A25

1. Introduction

Multiobjective optimization is the method of optimizing two or more real valued objective functions at the same time. There is no ideal minimizer to minimize all objective functions at once, thus the optimality concept is replaced by the idea of Pareto optimality/efficiency. A point is called Pareto optimal or efficient if there does not exist an alternative point with the equivalent or smaller objective function values, such that there is a decrease in at least one objective function value. In many applications such as engineering [1,2], economic theory [3], management science [4], machine learning [5,6], and space exploration [7], etc., several multiobjective optimization techniques are used to make the desired decision. One of the basic approaches is the weighting method [8], where a single objective optimization problem is created by the weighting of several objective functions. Another approach is the ϵ -constraint method [9], where we minimize only the chosen objective function and keep other objectives as constraints. Some multiobjective algorithms require a lexicographic method, where all objective functions are optimized in their order of priority [10,11]. First, the most preferred function is optimized, then that objective function is transformed into a constraint and a second priority objective function is optimized. This approach is repeated until the last objective function is optimized. The user needs to choose the sequence of objectives. Two distinct lexicographic optimizations with distinct sequences of objective functions do not produce the same solution. The disadvantages of such approaches are the choice of weights, constraints, and importance of the functions, respectively, which are not known in advance and they have to be specified from the beginning. Some other techniques [12–14] that do not need any prior information are developed for solving unconstrained

multiobjective optimization problems (UMOP) with at most linear convergence rate. Other methods like heuristic approaches or evolutionary approaches [15] provide an approximate Pareto front but do not guarantee the convergence property.

Newton's method [16] that solves the single-objective optimization problems is extended for solving (UMOP), which is based on an a priori parameter-free optimization method [17]. In this case, the objective functions are twice continuously differentiable, no other parameter or ordering of the functions is needed, and each objective function is replaced with a quadratic model. The rate of convergence is observed as superlinear, and it is quadratic if the second-order derivative is Lipschitz continuous. Newton's method is also studied under the assumptions of Banach and Hilbert spaces for finding the efficient solutions of (UMOP) [18]. A new type of Quasi-Newton algorithm is developed to solve the nonsmooth multiobjective optimization problems, where the directional derivative of every objective function exists [19].

A necessary condition for finding the vector critical point of (UMOP) is introduced in the steepest descent algorithm [12], where neither weighting factors nor ordering information for the different objective functions are assumed to be known. The relationship between critical points and efficient points is discussed in [17]. If the domain of (UMOP) is a convex set and the objective functions are convex component-wise then every critical point is the weak efficient point, and if the objective functions are strictly convex component-wise, then every critical point is the efficient point. The new classes of vector invex and pseudoinvex functions for (UMOP) are also characterized in terms of critical points and (weak) efficient points [20] by using Fritz John (FJ) optimality conditions and Karush–Kuhn–Tucker (KKT) conditions. Our focus is on Newton's direction for a standard scalar optimization problem which is implicitly induced by weighting the several objective functions. The weighting values are a priori unknown and non-negative KKT multipliers, that is, they are not required to fix in advance. Every new point generated by the Newton algorithm [17] initiates such weights in the form of KKT multipliers.

Quantum calculus or *q*-calculus is also called calculus without limits. The *q*-analogues of mathematical objects can be again recaptured as $q \rightarrow 1$. The history of quantum calculus can be traced back to Euler (1707–1783), who first proposed the quantum *q* in Newton's infinite series. In recent years, many researchers have shown considerable interest in examining and exploring the quantum calculus. Therefore, it emerges as an interdisciplinary subject. Of course, the quantum analysis is very useful in numerous fields such as in signal processing [21], operator theory [22], fractional integral and derivatives [23], integral inequalities [24], variational calculus [25], transform calculus [26], sampling theory [27], etc. The quantum calculus is seen as the bridge between mathematics and physics. To study some recent developments in quantum calculus, interested researches should refer to [28–31].

The *q*-calculus was first studied in the area of optimization [32], where the *q*-gradient is used in steepest descent method to optimize objective functions. Further, global optimum was searched using *q*-steepest descent method and *q*-conjugate gradient method where a descent scheme is presented using *q*-calculus with the stochastic approach which does not focus on the order of convergence of the scheme [33]. The *q*-calculus is applied in Newton's method to solve unconstrained single objective optimization [34]. Further, this idea is extended to solve (UMOP) within the context of the *q*-calculus [35].

In this paper, we present the *q*-calculus in Quasi-Newton's method for solving (UMOP). We approximate the second *q*-derivative matrices instead of evaluating them. Using *q*-calculus, we present the convergence rate is superlinear.

The rest of this paper is organized as follows. Section 2 recalls the problem, notation, and preliminaries. Section 3 derives a *q*-Quasi-Newton direction search method solved by (KKT) conditions. Section 4 establishes the algorithms for convergence analysis. The numerical results are given in Section 5 and the conclusion is in the last section.

2. Preliminaries

Denote \mathbb{R} as the set of real numbers, \mathbb{N} as the set of positive integers, and \mathbb{R}_+ or (\mathbb{R}_-) as the set of strictly positive or (negative) real numbers. If a function is continuous on any interval excluding zero, then the function is called continuous *q*-differentiable. For a function $f : \mathbb{R} \to \mathbb{R}$, the *q*-derivative of f [36] denoted as $D_{q,x}f$, is given as

$$D_{q,x}f(x) = \begin{cases} \frac{f(x) - f(qx)}{(1 - q)x}, & x \neq 0, q \neq 1\\ f'(x), & x = 0. \end{cases}$$
(1)

Suppose $f : \mathbb{R}^n \to \mathbb{R}$, whose partial derivatives exist. For $x \in \mathbb{R}^n$, consider an operator $\epsilon_{q,i}$ on f as

$$(\varepsilon_{q,i})f(x) = f(x_1, x_2, \dots, qx_i, x_{i+1}, \dots, x_n).$$
(2)

The *q*-partial derivative of *f* at *x* with respect to x_i , indicated by $D_{q,x_i}f$, is [23]:

$$D_{q,x_i}f(x) = \begin{cases} \frac{f(x) - (\varepsilon_{q,i}f)(x)}{(1-q)x_i}, & x_i \neq 0, \ q \neq 1, \\ \frac{\partial f}{\partial x_i}, & x_i = 0. \end{cases}$$
(3)

We are interested to solve the following (UMOP):

$$\begin{array}{ll} \text{minimize} & F(x) & (4) \\ \text{subject to} & x \in X, \end{array}$$

where $X \subseteq \mathbb{R}^n$ is a feasible region and $F : X \to \mathbb{R}^m$. Note that the function $F = (f_1, f_2, ..., f_m)$ is a vector function whose components are real valued functions such as $f_j : X \to \mathbb{R}$, where j = 1, ..., m. In general, n and m are independent. For $x, y \in \mathbb{R}^n$, we present the vector inequalities as:

$$x = y \iff x_i = y_i; \forall i = 1, ..., n,$$

$$x \ge y \iff x_i \ge y_i \forall i = 1, ..., n,$$

$$x \ge y \iff x_i \ge y_i \text{ and } x \neq y,$$

$$x > y \iff x_i > y_i \forall i = 1, ..., n.$$

A point $x^* \in X$ is called Pareto optimal point such that there is no any point $x \in X$, for which $F(x) \leq F(x^*)$, and $F(x) \neq F(x^*)$. A point $x^* \in X$ is called weakly Pareto optimal point if there is no $x \in X$ for which $F(x) < F(x^*)$. Similarly, a point $x^* \in X$ is a local Pareto optimal if there exists a neighborhood $Y \subseteq X$ of x^* such that the point x^* is a Pareto optima for F restricted on Y. Similarly, a point x^* is a local weak Pareto optimal if there exists a neighborhood $Y \subseteq X$ of x^* such that the point x^* is a neighborhood $Y \subseteq X$ of x^* such that the point x^* is a weak Pareto optimal for F restricted on Y. The matrix $JF(x) \in \mathbb{R}^{m \times n}$ is the Jacobian matrix of f_j at x, i.e., the j-th row of JF(x) is $\nabla_q f_j(x)$ (q-gradient) for all $j = 1, \ldots, m$. Let $Wf_j(x)$ be the Hessian matrix of f_j at x for all $j = 1, \ldots, m$. Note that every Pareto optimal point is a weakly Pareto optimal point [37]. The directional derivative of f_j at x in the descent direction d_q is given as:

$$f'_j(x,d_q) = \lim_{\alpha \to 0} \frac{f_j(x+\alpha d_q) - f_j(x)}{\alpha}$$
(5)

The necessary condition to get the critical point for multiobjective optimization problems is given in [17]. For any $x \in \mathbb{R}^n$, ||x|| denotes the Euclidean norm in \mathbb{R}^n . Let $K(x^0, r) = \{x : ||x - x^0|| \le r\}$ with a center $x^0 \in \mathbb{R}^n$ and radius $r \in \mathbb{R}_+$. Norm of the matrix $A \in \mathbb{R}^{n \times n}$ is $||A|| = \max_{x \in \mathbb{R}^{n \times n}} \frac{||Ax||}{||x||}$, $x \ne 0$. The following proposition indicates that when f(x) is a linear function, then the *q*-gradient is similar to the classical gradient. **Proposition 1** ([33]). If $f(x) = a + p^T x$, where $a \in \mathbb{R}$ and $p \in \mathbb{R}^n$, then for any $x \in \mathbb{R}^n$, and $q \in (0, 1)$, we have $\nabla_q f(x) = \nabla f(x) = p$.

All the quasi-Newton methods approximate the Hessian of function f as $W^k \in \mathbb{R}^{n \times n}$, and update the new formula based on previous approximation [38]. Line search methods are imperative methods for (UMOP) in which a search direction is first computed and then along this direction a step-length is chosen. The entire process is an iterative.

3. The *q*-Quasi-Newton Direction for Multiobjective

The most well-known quasi-Newton method for single objective function is the BFGS (Broyden, Fletcher, Goldfarb, and Shanno) method. This is a line search method along with a descent direction d_q^k within the context of *q*-derivative, given as:

$$d_q^k = -(W^k)^{-1} \nabla_q f(x^k), \tag{6}$$

where *f* is a continuously *q*-differentiable function, and $W^k \in \mathbb{R}^{n \times n}$ is a positive definite matrix that is updated at every iteration. The new point is:

$$x^{k+1} = x^k + \alpha_k d_q^k. \tag{7}$$

In the case of the Steepest Descent method and Newton's method, W^k is taken to be an Identity matrix and exact Hessian of f, respectively. The quasi-Newton BFGS scheme generates the next W^{k+1} as

$$W^{k+1} = W^k - \frac{W^k s^k (s^k)^T W^k}{(s^k)^T W^k s^k} + \frac{y^k (y^k)^T}{(s^k)^T y^k},$$
(8)

where $s^k = x^{k+1} - x^k = \alpha_k d_q^k$, and $y^k = \nabla_q f(x^{k+1}) - \nabla_q f(x^k)$. In Newton's method, second-order differentiability of the function is required. While calculating W^k , we use *q*-derivative which behaves like a Hessian matrix of f(x). W^{k+1} may not be a positive definite, which can be modified to be a positive definite through the symmetric indefinite factorization [39]. The *q*-Quasi-Newton's direction $d_q(x)$ is an optimal solution of the following modified problem [40] as:

$$\min_{d_q \in \mathbb{R}^n} \max_{j=1,\dots,m} \nabla_q f_j(x) d_q + \frac{1}{2} d_q^T W_j(x)^T d_q,$$
(9)

where $W_i(x)$ is computed as (8). The solution and optimal value of (9) are:

$$\psi(x) = \min_{d_q \in \mathbb{R}^n} \max_{j=1,\dots,m} \nabla_q f_j(x)^T d_q + \frac{1}{2} d_q^T W_j(x) d_q,$$
(10)

and

$$d_q(x) = \arg\min_{d_q \in \mathbb{R}^n} \max_{j=1,\dots,m} \nabla f_j(x)^T d_q + \frac{1}{2} d_q^T W_j(x) d_q.$$
(11)

The problem (9) becomes a convex quadratic optimization problem (CQOP) as follows:

minimize
$$h(t, d_q) = t$$
,
subject to $\nabla_q f_j(x)^T d_q + \frac{1}{2} d_q^T W_j(x) d_q - t \le 0, \ j = 1, \dots, m$, (12)
where $(t, d_q) \in \mathbb{R} \times \mathbb{R}^n$.

The Lagrangian function of (CQOP) is:

$$L((t, d_q), \lambda) = t + \sum_{j=1}^{m} \lambda_j \bigg(\nabla_q f_j(x)^T d_q + \frac{1}{2} d_q^T W_j(x) d_q - t \bigg).$$
(13)

For $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)^T$, we obtain the following (KKT) conditions [40]:

$$\sum_{j=1}^{m} \lambda_j \big(\nabla_q f_j(x) + W_j(x) d_q \big) = 0, \tag{14}$$

$$\lambda_j \ge 0, \ j = 1, \dots, m, \tag{15}$$

$$\sum_{j=1}^{m} \lambda_j = 1, \tag{16}$$

$$\nabla_q f_j(x)^T d_q + \frac{1}{2} d_q^T W_j(x) d_q \le t, \ j = 1, \dots, m,$$
(17)

$$\lambda_j \left(\nabla_q f_j(x)^T d_q + \frac{1}{2} d_q^T W_j(x) d_q - t \right) = 0, \ j = 1, \dots, m.$$
(18)

The solution $(d_q(x), \psi(x))$ is unique, and set $\lambda_j = \lambda_j(x)$ for all j = 1, ..., m with $d_q = d_q(x)$ and $t = \psi(x)$ for satisfying (14)–(18). From (14), we obtain

$$d_q(x) = -\left(\sum_{j=1}^m \lambda_j(x) W_j(x)\right)^{-1} \sum_{j=1}^m \lambda_j(x) \nabla_q f_j(x).$$
⁽¹⁹⁾

This is a so-called *q*-Quasi-Newton's direction for solving (UMOP). We present the basic result for relating the stationary condition at a given point x to its q-Quasi-Newton direction $d_q(x)$ and function ψ .

Proposition 1. Let $\psi : X \to \mathbb{R}$ and $d_q : X \to \mathbb{R}^n$ be given by (10) and (11), respectively, and $W_j(x) \ge 0$ for all $x \in X$. Then,

- $\psi(x) \leq 0$ for all $x \in X$. 1.
- 2. The conditions below are equivalent:
 - *The point x is non stationary.* (a)
 - (b)
 - (c)
 - $d_q(x) \neq 0$ $\psi(x) < 0.$ $d_q(x)$ is a descent direction. (*d*)
- *The function* ψ *is continuous.* 3.

Proof. Since $d_q = 0$, then from (10), we have

$$\psi(x) \leq \min_{d_q \in \mathbb{R}^n} \max_{j=1,...,m} \nabla_q f_j(x)^T 0 + \frac{1}{2} d_q^T W_j(x) 0 = 0,$$

thus $\psi(x) \leq 0$. It means that $JF(x^*)d_q(x) \in \mathbb{R}^m_-$. Thus, the given point $x \in \mathbb{R}^n$ is non-stationary. Since $W_i(x)$ is positive definite, and from (10) and (11), we have

$$\nabla_q f_j(x)^T d_q(x) < \nabla f_j(x)^T d_q(x) + \frac{1}{2} d_q(x)^T W_j(x)^T d_q(x) = \psi(x) \le 0.$$

Since $\psi(x)$ is the optimal value of (CQOP), and it is negative, thus solution of (CQOP) can never be $d_q(x) = 0$. It is sufficient to show that the continuity [41] of ψ in set $Y \subset X$. Since $\psi(x) \leq 0$, then

$$\nabla_q f_j(x)^T d_q(x) \le -\frac{1}{2} d_q(x)^T W_j(x) d_q(x),$$
(20)

for all j = 1, ..., m, and $W_j(x)$, where j = 1, ..., m are positive definite for all $x \in Y$. Thus, the eigenvalues of Hessian matrices $W_j(x)$, where j = 1, ..., m are uniformly bounded away from zero on Y so there exists $R, S \in \mathbb{R}_+$ such that

$$R = \max_{x \in Y, j=1,\dots,m} \|\nabla_q f_j(x)\|,\tag{21}$$

and

$$S = \min_{x \in Y, \|e\| = 1, j = 1, \dots, m} e^T W_j(x) e.$$
 (22)

From (20) and using Cauchy-Schwarz inequality, we get

$$\|\nabla_q f_j(x)\| \|d_q(x)\| \le \frac{1}{2}S \|d_q(x)\|^2 \le R \|d_q(x)\|,$$

that is,

$$d_q(x) \le 2\frac{R}{S},$$

for all $x \in Y$, that is, Newton's direction is uniformly bounded on Y. We present the family of function $\{\aleph_{x,j}\}_{x \in Y, j=1,...,m}$, where

 $\aleph_{x,j}: Y \to \mathbb{R},$

and

$$z o
abla_q f(z)^T d_q(x) + rac{1}{2} d_q(x)^T W_j(x) d_q(x).$$

We shall prove that this family of functions is uniformly equicontinuous. For small value $\epsilon_z \in \mathbb{R}_+$ there exists $\delta_z \in \mathbb{R}_+$, and for $y \in K(z, \delta_z)$, we have

$$\|W_j(y)-\nabla_q^2 f_j(z)\| < \frac{\epsilon_z}{2},$$

and

$$\|\nabla_q^2 f_j(y) - \nabla_q^2 f_j(z)\| < \frac{\epsilon_z}{2},$$

for all j = 1, ..., m. because of *q*-continuity of Hessian matrices, the second inequality is true. Since *Y* is compact space, then there exists a finite sub-cover.

$$\psi_{x,j}(z) = \nabla_q f_j(z)^T d_q(x) + \frac{1}{2} d_q(x)^T W_j(x) d_q(x),$$

that is

$$\psi_{x,j}(z) = \nabla_q f_j(z)^T d_q(x) + \frac{1}{2} d_q(x)^T \nabla^2 f_j(z) d_q(x) + \frac{1}{2} d_q(x)^T (W_j(z) - \nabla_q^2 f_j(z) d_q(x)).$$

To show the *q*-continuous of last term, set $y_1, y_2 \in Y$ such that $||y_1 - y_2|| < \delta$ for small $\delta \in \mathbb{R}_+$, then

$$\begin{split} |\frac{1}{2}d_q(x)^T W_j(y_1) - \nabla_q^2 f_j(y_1) d_q(x) - \frac{1}{2}d_q(x)^T W_j(y_2) - \nabla_q^2 f_j(y_2) d_q(x)| \\ &\leq \frac{1}{2} \|d_q(x)\|^2 (\|B_j(y_1) - \nabla^2 f_j(z_1))\| + \|\nabla_q^2 f_j(z_2) \\ &- \nabla_q^2 f_j(z_1 + \|B_j(y_2) - \nabla^2 f_j(z_{21}))\| + \|\nabla_q^2 f_j(z_2) - \nabla_q^2 f_j(z_{21}\|) \\ &\leq \frac{1}{2} \|d_q(x)\|^2 (\varepsilon_{z1} + \varepsilon_{z2}). \end{split}$$

 $\psi_{x,j}$ is uniformly continuous [40] for all $x \in Y$ and for all j = 1, ..., m. There exists $\delta \in R_+$ such that for all $y, z \in Y$, $||y - z|| < \delta$ implies $|\psi(y) - \psi(z)| < \epsilon$ for all $x \in Y$. Thus, $||y - z|| < \delta$.

$$egin{aligned} \psi(z) &\leq \max_{j=1,\dots,m}
abla f_j(z)^T d_q(y) + rac{1}{2} d_q(y)^T W_j(z) d_q(y) = \phi_y(z) \ &\leq \phi_y(y) + |\phi_y(z) - \phi_y(y)| < \psi(y) + \epsilon. \end{aligned}$$

Thus, $\psi(z) - \psi(y) < \epsilon$. If we interchange *y* and *z*, then $|\psi(z) - \psi(y)| < \epsilon$. It proves the continuity of ψ . \Box

The following modified lemma is due to [17,42].

Lemma 1. Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be continuously q-differentiable. If $x^* \in X$ is not a critical point for $\nabla_q(x)d_q < 0$, where $d_q \in \mathbb{R}^n$, $\sigma \in (0, 1]$, and $\varepsilon > 0$. Then,

$$x + \alpha d_q(x) \in X$$
 and $F(x + \alpha d_q(x)) < F(x) + \alpha \gamma \psi(x)$,

for any $\alpha \in (0, \sigma]$ and $\gamma \in (0, \varepsilon]$.

Proof. Since x^* is not a critical point, then $\psi(x) < 0$. Let r > 0 such that $B(x, r) \subset X$ and $\alpha \in (0, \sigma]$. Therefore,

$$F(x + \alpha d_q(x)) - F(x) = \alpha \nabla_q F(x)^T d_q(x) + o_j(\alpha d_q(x), x)$$

Since $\nabla_q(x)d_q(x) < \psi(x)$, for $\alpha \in (0, \sigma]$, then

$$F(x + \alpha d_q(x)) - F(x) = \alpha \gamma \psi(x) + \alpha (1 - \sigma) \psi(x) + o_j(\alpha d_q(x), x).$$

The last term in the right-hand side of the above equation is non-positive because $\psi(x) \le \frac{\psi(x^*)}{2} < 0$, for $\alpha \in [0, \sigma]$. \Box

4. Algorithm and Convergence Analysis

We first present the following Algorithm 1 [43] to find the gradient of the function using q-calculus. The higher-order q-derivative of f can be found in [44].

Algorithm 1 *q*-Gradient Algorithm

1: Input
$$q \in (0, 1), f(x), x \in \mathbb{R}, z$$
.
2: if $x = 0$ then
3: Set $g \leftarrow \lim \left(\frac{f(z) - f(q * z)}{(z - q * z)}, z, 0\right)$.
4: else
5: Set $g \leftarrow \frac{f(x) - f(q * x)}{(x - q * x)}$.
6: Print $\nabla_q f(x) \leftarrow g$.

Example 1. Given that
$$f : \mathbb{R}^2 \to \mathbb{R}$$
 defined by $f(x_1, x_2) = x_2^2 + 3x_1^3$. Then $\nabla_q f(x) = \begin{bmatrix} 3x_1^2(1+q+q^2) \\ x_2(1+q) \end{bmatrix}$.

We are now prepared to write the unconstrained q-Quasi-Newton's Algorithm 2 for solving (UMOP). At each step, we solve the (CQOP) to find the *q*-Quasi-Newton direction. Then, we obtain the step length using the Armijo line search method. In every iteration, the new point and Hessian approximation are generated based on historical values.

Algorithm 2 q-Quasi-Newton's Algorithm for Unconstrained Multiobjective (q-QNUM)

1: Choose $q \in (0,1)$, $x^0 \in X$, symmetric definite matrix $W^0 \in \mathbb{R}^{n \times n}$, $c \in (0,1)$, and a small tolerance value $\epsilon > 0$. 2: **for** k=0,1,2,... **do** 3: Solve (CQOP). Compute d_q^k and ψ^k . 4: if $\psi^k > -\epsilon$ then 5: Stop. 6: else 7: Choose α_k as the $\alpha \in (0, 1]$ such that $x^k + \alpha d_q^k \in X$ and $F(x^k + \alpha d_q^k) \leq F(x^k) + c\alpha \psi^k$. 8: Update $x^{k+1} \leftarrow x^k + \alpha_k d_a^k$. 9: Update W_i^{k+1} , where $j = 1, \ldots, m$ using (8). 10:

We now finally start to show that every sequence produced by the proposed method converges to a weakly efficient point. It does not matter how poorly the initial point is guessed. We assume that the method does not stop, and produces an infinite sequence of iterates. We now present the modified sufficient conditions for the superlinear convergence [17,40] within the context of q-calculus.

Theorem 1. Let $\{x^k\}$ be a sequence generated by (q-QNUM), and $Y \subset X$ be a convex set. Also, $\gamma \in (0,1)$ and $r, a, b, \delta, \epsilon > 0$, and

- *(a)*
- (b)
- $aI \leq W_j(x) \leq bI \text{ for all } x \in Y, j = 1, \dots, m,$ $\|\nabla_q^2 f_j(y) \nabla_q^2 f_j(x), \| < \frac{\varepsilon}{2} \text{ for all } x, y \in Y \text{ with } \|y x\| \in \delta,$ $\|(W_j^k \nabla_q^2 f_j(x^k))(y x^k)\| < \frac{\varepsilon}{2} \|y x^k\| \text{ for all } k \geq k_0, y \in Y, j = 1, \dots, m,$ (c)

(d)
$$\frac{\varepsilon}{a} \leq 1-c$$
,

 $\overset{\ddot{B}}{\|d_q(x^0)\| < \min\{\delta, r(1-\frac{\varepsilon}{a})\}. }$ (e) (f)

Then, for all $k \ge k_0$ *, we have that*

1.
$$\|x^k - x^{k_0}\| \le \|d_q(x^0)\| \frac{1 - (\frac{\varepsilon}{a})^{k-k_0}}{1 - (\frac{\varepsilon}{a})}$$

2.
$$\alpha_k = 1$$
,

 $\|d_a(x^k)\| < \|d_a(x^{k_0})\|(\frac{\varepsilon}{a})^{k-k_0},$ 3.

4.
$$||d_q(x^{k+1})|| \le ||d_q(x^k)||_{\overline{a}}^{\underline{\varepsilon}}$$
.

Then, the sequence $\{x^k\}$ converges to local Pareto points $x^* \in \mathbb{R}^m$, and the convergence rate is superlinear.

Proof. From part 1, part 3 of this theorem and triangle inequality,

$$||x^{k} + d_{q}(x^{k}) - x^{0}|| \le \frac{1 - \left(\frac{\epsilon}{a}\right)^{k+1}}{1 - \frac{\epsilon}{a}} ||d_{q}(x^{k_{0}})||.$$

From (d) and (f), we follow x^k , $x^k + d_q(x^k) \in K(x^{k_0}, r)$ and $x^k + d_q(x^k) - x^k < \delta$. We also have

$$f_j(x^k + d_q(x^k)) \le f_j(x^k) + d_q(x^k) \nabla_q f(x^k) + \frac{1}{2} d_q(x^k) (\nabla_q^2 f)(x^k) + \frac{1}{2} \|d_q(x^k)\|^2,$$

that is,

$$f_j(x^k + d_q(x^k)) \le f_j(x^k) + \psi(x^k) + \frac{\varepsilon}{2} \|d_q(x^k)\|^2$$

= $f_j(x^k) + \gamma \psi(x^k) + (1 - \gamma)\psi(x^k) + \frac{\varepsilon}{2} \|d_q(x^k)\|^2.$

Since $\psi \leq 0$ and $(1 - \gamma)\psi(x^k) + \frac{\epsilon}{2} ||d_q(x^k)||^2 \leq (\epsilon - a(1 - \gamma))\frac{||d_q(x^k)||^2}{2} \leq 0$, we get $f_j(x^k + d_q(x^k)) \leq f_j(x^k) + \gamma \psi(x^k)$,

for all j = 1, ..., m. The Armijo conditions holds for $\alpha_k = 1$. Part 1 of this theorem holds. We now set $x^k, x^{k+1} \in K(x^{k_0}, r)$, and $||x^{k+1} - x^k|| < \delta$. Thus, we get $x^{k+1} = x^k + d_q(x^k)$. We now define $v(x^{k+1}) = \sum_{j=1}^m \lambda_j^k \nabla_q f_j(x^{k+1})$. Therefore,

$$|\psi(x^{k+1})| \le \frac{1}{2a} \|v(x^{k+1})\|^2$$

We now estimate $||v(x^{k+1})||$. For $x \in X$, we define

$$G^k(x) := \sum_{j=1}^m \lambda_j^k f_j(x^{k+1})$$

and

$$H^k = \sum_{j=1}^m \lambda_j^k W_j(x^k),$$

where $\lambda_j^k \ge 0$, for all j = 1, ..., m, are KKT multipliers. We obtain following:

$$\nabla_q G^k(x) = \sum_{j=1}^m \lambda_j^k \nabla_q f_j(x),$$

and

$$\nabla_q^2 G^k(x) = \sum_{j=1}^m \lambda_j^k \nabla_q^2 f_j(x).$$

Then, $v^{k+1} = \nabla_q G^k(x^{k+1})$. We get

$$d_q(x^k) = -(H^k)^{-1} \nabla_q G^k(x^k).$$

From assumptions (b) and (c) of this theorem,

$$\|\nabla_q^2 G^k(y) - \nabla_q^2 G^k(x^k)\| < \frac{\varepsilon}{2},$$

$$\|(H^k - \nabla_q^2 G^k(x^k))(y - x^k)\| < \frac{\varepsilon}{2} \|y - x^k\|$$

hold for all $x, y \in Y$ with $||y - x|| < \delta$ and $k \ge k_0$. We have

$$\nabla_q G^k(x^k + d_q(x^k)) - (\nabla_q G^k(x^k) + H^k d_q(x^k)) \| < \epsilon \|d_q(x^k)\|.$$

Since $\nabla_q G^k(x^k) + H^k d_q(x^k) = 0$, then

$$\|v(x^{k+1})\| = \|\nabla_q G^k(x^{k+1})\| < \epsilon \|d_q(x^k)\|,$$

Mathematics 2020, 8, 616

and

$$|\psi^{k+1}| \leq rac{1}{2a} \|v(x^{k+1})\|^2 < rac{\epsilon^2}{2a} \|d_q(x^k)\|^2.$$

We have

$$\frac{a}{2} \|d_q(x^{k+1})\|^2 < \frac{\epsilon^2}{2a} \|d_q(x^k)\|^2$$

 $\|d_q(x^{k+1})\| < \frac{\epsilon}{q} \|d_q(x^k)\|$

Thus,

Thus, part 4 is proved. We finally prove superlinear convergence of $\{x^k\}$. First we define

$$r^k = \|d^0_q\| \frac{\frac{\varepsilon}{a}^{k-k_0}}{1-\frac{\varepsilon}{a}},$$

and

$$\delta^k = \|d_q^{k_0}\| \left(\frac{\varepsilon}{a}\right)^{k-k_0}.$$

From triangle inequality, assumptions (e), (f) and part 1, we have $K(x^k, r^k) \subset K(x^{k_0}, r) \subset V$. Choose any $\tau \in \mathbb{R}_+$, and define

$$\bar{\varepsilon} = \min\{a\frac{\tau}{1+2\tau}, \varepsilon\}.$$

For $k \ge k_0$ inequalities

$$\|\nabla_q^2 f_j(y) - \nabla_q^2 f_j(x)\| < \frac{\bar{\varepsilon}}{2}$$

for all $x, y \in K(x^k, r^k)$ with $||y - x|| < \delta^k$, and

$$\|W_j(x^l) - \nabla_q^2 f_j(x^l)(y - x^l)\| < \frac{\overline{\varepsilon}}{2}$$

for all $y \in K(x^k, r^k)$ and $l \ge k$ holds both for j = 1, ..., m. Assumptions (a)–(f) are satisfied for $\overline{\epsilon}, r^k, \delta^k$, and x^k instead of ϵ, r, δ , and x^0 , respectively. We have

$$||x^{l} - x^{k}|| \le ||d_{q}(x^{k})|| \frac{1 - (\frac{\overline{\epsilon}}{a})^{l-k}}{1 - \frac{\overline{\epsilon}}{a}}$$

Let $l \to \infty$ and we get $||x^* - x^k|| \le ||d_q(x^k)|| \frac{1}{1-\frac{\varepsilon}{a}}$. Using the last inequality, and part 4, we have

$$\|x^* - x^{k+1}\| \le \|d_q(x^{k+1})\| \frac{1}{1 - \frac{\tilde{\varepsilon}}{a}} \le \|d_q(x^k)\| \frac{\frac{\tilde{\varepsilon}}{a}}{1 - \frac{\tilde{\varepsilon}}{a}}$$

From above and triangle inequality, we have

$$||x^* - x^{k+1}|| \ge ||x^{k+1} - x^k|| - ||x^* - x^{k+1}||$$

that is,

$$\|x^* - x^{k+1}\| \ge \|d_q(x^k)\| - \|d^k\| \frac{\bar{\varepsilon}}{1 - \frac{\bar{\varepsilon}}{\bar{a}}} = \|d_q^k\| \frac{1 - 2\frac{\bar{\varepsilon}}{\bar{a}}}{1 - \frac{\bar{\varepsilon}}{\bar{a}}}.$$
(23)

Since $1 - 2\frac{\overline{\tilde{e}}}{\overline{a}} > 0$, and $1 - 2\frac{\overline{\tilde{e}}}{\overline{a}} > 0$, then we get

$$\|x^* - x^{k+1}\| \le \tau \|x^* - x^k\|$$

where $\tau \in \mathbb{R}_+$ is chosen arbitrarily. Thus, the sequence $\{x^k\}$ converges superlinearly to x^* . \Box

5. Numerical Results

The proposed algorithm (*q*-QNUM), i.e., Algorithm 2, presented in Section 4 is implemented in MATLAB (2017a) and tested on some test problems known from the literature. All tests were run under the same conditions. The box constraints of the form $lb \le x \le ub$ are used for each test problem. These constraints are considered under the direction search problem (CQOP) such that the newly generated point always lies in the same box, that is, $lb \le x + d_q \le ub$ holds. We use the stopping criteria at x^k as: $\psi(x^k) > -\epsilon$ where $\epsilon \in \mathbb{R}_+$. All test problems given in Table 1 are solved 100 times. The starting points are randomly chosen from a uniform distribution between lb and ub. The first column in the given table is the name of the test problem. We use the abbreviation of author's names and number of the problem in the corresponding paper. The second column indicates the source of the paper. The third column is for lower bound and upper bound. We compare the results of (q-QNUM) with (QNMO) of [40] in the form of a number of iterations (*iter*), number of objective functions evaluation (*obj*), and number of gradient evaluations (*grad*), respectively. From Table 1, we can conclude that our algorithm shows better performance.

Example 2. Find the approximate Pareto front using (q-QNUM) and (QNMO) for the given (UMOP) [45]:

Minimize
$$f_1(x_1, x_2) = (x_1 - 1)^2 + (x_1 - x_2)^2$$
,
Minimize $f_2(x_1, x_2) = (x_2 - 3)^2 + (x_1 - x_2)^2$,

where $-3 \le x_1, x_2 \le 10$.

The number of Pareto points generated due to (*q*-QNUM) with Algorithm 1 and (QNMO) is shown in Figure 1. One can observe that the number of iterations as *iter* = 200 in (*q*-QNUM) and *iter* = 525 in (QNMO) are responsible for generating the approximate Pareto front of above (UMOP).



Figure 1. Approximate Pareto Front of Example 1.

12 of	14

	Source		(a ONUM)		(ONMO)			
Problem		[lb,ub]	iter	obj	grad	iter	obj	grad
BK1	[46]	[-5, 10]	200	200	200	200	200	200
MOP5	[46]	[-30, 30]	141	965	612	333	518	479
MOP6	[46]	[0, 1]	250	2177	1712	181	2008	2001
MOP7	[46]	[-400, 400]	200	200	200	751	1061	1060
DG01	[47]	[-10, 13]	175	724	724	164	890	890
IKK1	[47]	[-50, 50]	170	170	170	253	254	253
SP1	[45]	[-3, 2]	200	200	200	525	706	706
SSFYY1	[45]	[-2, 2]	200	200	200	200	300	300
SSFYY2	[45]	[-100, 100]	263	277	277	263	413	413
SK1	[48]	[-10, 10]	139	1152	1152	87	732	791
SK2	[48]	[-3, 11]	154	1741	1320	804	1989	1829
VU1	[49]	[-3,3]	316	1108	1108	11,361	19 <i>,</i> 521	11,777
VU2	[49]	[-3,7]	99	1882	1882	100	1900	1900
VFM1	[50]	[-2, 2]	195	195	195	195	290	290
VFM2	[50]	[-4, 4]	200	200	200	524	693	678
VFM3	[50]	[-3, 3]	161	1130	601	690	1002	981

Table 1. Numerical Results of Test Problems.

6. Conclusions

The *q*-Quasi-Newton method converges superlinearly to the solution of (UMOP) if all objective functions are strongly convex within the context of *q*-derivative. In a neighborhood of this solution, the algorithm uses a full Armijo steplength. The numerical performance of the proposed algorithm is faster than their actual evaluation.

Author Contributions: K.K.L. gave reasonable suggestions for this manuscript; S.K.M. gave the research direction of this paper; B.R. revised and completed this manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Science and Engineering Research Board (Grant No. DST-SERB-MTR-2018/000121) and the University Grants Commission (IN) (Grant No. UGC-2015-UTT-59235).

Acknowledgments: The authors are grateful to the anonymous reviewers and the editor for the valuable comments and suggestions to improve the presentation of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Eschenauer, H.; Koski, J.; Osyczka, A. *Multicriteria Design Optimization: Procedures and Applications*; Springer: Berlin, Germany, 1990.
- 2. Haimes, Y.Y.; Hall, W.A.; Friedmann, H.T. *Multiobjective Optimization in Water Resource Systems, The Surrogate Worth Trade-off Method*; Elsevier Scientific: Amsterdam, The Netherlands, 1975.
- Nwulu, N.I.; Xia, X. Multi-objective dynamic economic emission dispatch of electric power generation integrated with game theory based demand response programs. *Energy Convers. Manag.* 2015, *89*, 963–974. [CrossRef]
- 4. Badri, M.A.; Davis, D.L.; Davis, D.F.; Hollingsworth, J. A multi-objective course scheduling model: Combining faculty preferences for courses and times. *Comput. Oper. Res.* **1998**, 25, 303–316. [CrossRef]

- Ishibuchi, H.; Nakashima, Y.; Nojima, Y. Performance evaluation of evolutionary multiobjective optimization algorithms for multiobjective fuzzy genetics-based machine learning. *Soft Comput.* 2011, 15, 2415–2434. [CrossRef]
- 6. Liu, S.; Vicente, L.N. The stochastic multi-gradient algorithm for multi-objective optimization and its application to supervised machine learning. *arXiv* **2019**, arXiv:1907.04472.
- 7. Tavana, M. A subjective assessment of alternative mission architectures for the human exploration of mars at NASA using multicriteria decision making. *Comput. Oper. Res.* **2004**, *31*, 1147–1164. [CrossRef]
- Gass, S.; Saaty, T. The computational algorithm for the parametric objective function. *Nav. Res. Logist. Q.* 1955, 2, 39–45. [CrossRef]
- 9. Miettinen, K. Nonlinear Multiobjective Optimization; Kluwer Academic: Boston, MA, USA, 1999.
- 10. Fishbum, P.C. Lexicographic orders, utilities and decision rules: A survey. *Manag. Sci.* **1974**, *20*, 1442–1471. [CrossRef]
- 11. Coello, C.A. An updated survey of GA-based multiobjective optimization techniques. *ACM Comput. Surv.* (*CSUR*) **2000**, *32*, 109–143. [CrossRef]
- 12. Fliege, J.; Svaiter, B.F. Steepest descent method for multicriteria optimization. *Math. Method. Oper. Res.* 2000, 51, 479–494. [CrossRef]
- 13. Drummond, L.M.G.; Iusem, A.N. A projected gradient method for vector optimization problems. *Comput. Optim. Appl.* **2004**, *28*, 5–29. [CrossRef]
- 14. Drummond, L.M.G.; Svaiter, B.F. A steepest descent method for vector optimization. *J. Comput. Appl. Math.* **2005**, *175*, 395–414. [CrossRef]
- 15. Branke, J.; Dev, K.; Miettinen, K.; Slowiński, R. (Eds.) *Multiobjective Optimization: Interactive and Evolutionary Approaches;* Springer: Berlin, Germany, 2008.
- 16. Mishra, S.K.; Ram, B. *Introduction to Unconstrained Optimization with R*; Springer Nature: Singapore, 2019; pp. 175–209.
- 17. Fliege, J.; Drummond, L.M.G.; Svaiter, B.F. Newton's method for multiobjective optimization. *SIAM J. Optim.* **2009**, *20*, 602–626. [CrossRef]
- Chuong, T.D. Newton-like methods for efficient solutions in vector optimization. *Comput. Optim. Appl.* 2013, 54, 495–516. [CrossRef]
- 19. Qu, S.; Liu, C.; Goh, M.; Li, Y.; Ji, Y. Nonsmooth Multiobjective Programming with Quasi-Newton Methods. *Eur. J. Oper. Res.* **2014**, 235, 503–510. [CrossRef]
- 20. Jiménez; M.A.; Garzón, G.R.; Lizana, A.R. (Eds.) *Optimality Conditions in Vector Optimization*; Bentham Science Publishers: Sharjah, UAE, 2010.
- 21. Al-Saggaf, U.M.; Moinuddin, M.; Arif, M.; Zerguine, A. The *q*-least mean squares algorithm. *Signal Process*. **2015**, *111*, 50–60. [CrossRef]
- 22. Aral, A.; Gupta, V.; Agarwal, R.P. *Applications of q-Calculus in Operator Theory*; Springer: New York, NY, USA, 2013.
- 23. Rajković, P.M.; Marinković, S.D.; Stanković, M.S. Fractional integrals and derivatives in *q*-calculus. *Appl. Anal. Discret. Math.* **2007**, *1*, 311–323.
- 24. Gauchman, H. Integral inequalities in q-calculus. Comput. Math. Appl. 2004, 47, 281–300. [CrossRef]
- 25. Bangerezako, G. Variational q-calculus. J. Math. Anal. Appl. 2004, 289, 650–665. [CrossRef]
- 26. Abreu, L. A *q*-sampling theorem related to the q-Hankel transform. *Proc. Am. Math. Soc.* **2005**, *133*, 1197–1203. [CrossRef]
- 27. Koornwinder, T.H.; Swarttouw, R.F. On *q*-analogues of the Fourier and Hankel transforms. *Trans. Am. Math. Soc.* **1992**, *333*, 445–461.
- 28. Ernst, T. *A Comprehensive Treatment of q-Calculus;* Springer: Basel, Switzerland; Heidelberg, Germany; New York, NY, USA; Dordrecht, The Netherlands; London, UK, 2012.
- 29. Noor, M.A.; Awan, M.U.; Noor, K.I. Some quantum estimates for Hermite-Hadamard inequalities. *Appl. Math. Comput.* **2015**, 251, 675–679. [CrossRef]
- 30. Pearce, C.E.M.; Pečarić, J. Inequalities for differentiable mappings with application to special means and quadrature formulae. *Appl. Math. Lett.* **2000**, *13*, 51–55. [CrossRef]
- 31. Ernst, T. A Method for q-Calculus. J. Nonl. Math. Phys. 2003, 10, 487–525. [CrossRef]

- 32. Sterroni, A.C.; Galski, R.L.; Ramos, F.M. The *q*-gradient vector for unconstrained continuous optimization problems. In *Operations Research Proceedings*; Hu, B., Morasch, K., Pickl, S., Siegle, M., Eds.; Springer: Heidelberg, Germany, 2010; pp. 365–370.
- 33. Gouvêa, E.J.C.; Regis, R.G.; Soterroni, A.C.; Scarabello, M.C.; Ramos, F.M. Global optimization using *q*-gradients. *Eur. J. Oper. Res.* **2016**, *251*, 727–738. [CrossRef]
- Chakraborty, S.K.; Panda, G. Newton like line search method using *q*-calculus. In *International Conference on Mathematics and Computing. Communications in Computer and Information Science*; Giri, D., Mohapatra, R.N., Begehr, H., Obaidat, M., Eds.; Springer: Singapore, 2017; Volume 655, pp. 196–208.
- 35. 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**. [CrossRef]
- 36. Jackson, F.H. On *q*-functions and a certain difference operator. *Earth Environ. Sci. Trans. R. Soc. Edinb.* **1908**, 46, 253–281. [CrossRef]
- 37. Bento, G.C.; Neto, J.C. A subgradient method for multiobjective optimization on Riemannian manifolds. *J. Optimiz. Theory App.* **2013**, *159*, 125–137. [CrossRef]
- 38. Andrei, N. A diagonal quasi-Newton updating method for unconstrained optimization. *Numer. Algorithms* **2019**, *81*, 575–590. [CrossRef]
- 39. Nocedal, J.; Wright, S.J. *Numerical Optimization*, 2nd ed.; Springer Series in Operations Research and Financial Engineering; Springer: New York, NY, USA, 2006.
- 40. Povalej, Z. Quasi-Newton's method for multiobjective optimization. J. Comput. Appl. Math. 2014, 255, 765–777. [CrossRef]
- 41. Ye, Y.L. D-invexity and optimality conditions. J. Math. Anal. Appl. 1991, 162, 242–249. [CrossRef]
- 42. Morovati, V.; Basirzadeh, H.; Pourkarimi, L. Quasi-Newton methods for multiobjective optimization problems. 40R-Q. J. Oper. Res. 2018, 16, 261–294. [CrossRef]
- 43. Samei, M.E.; Ranjbar, G.K.; Hedayati, V. Existence of solutions for equations and inclusions of multiterm fractional *q*-integro-differential with nonseparated and initial boundary conditions. *J. Inequal Appl.* **2019**, 273. [CrossRef]
- 44. Adams, C.R. The general theory of a class of linear partial difference equations. *Trans. Am. Math.Soc.* **1924**, 26, 183–312.
- 45. Sefrioui, M.; Perlaux, J. Nash genetic algorithms: Examples and applications. In Proceedings of the 2000 Congress on Evolutionary Computation, La Jolla, CA, USA, 16–19 July 2000; Volume 1, pp. 509–516.
- 46. Huband, S.; Hingston, P.; Barone, L.; While, L. A review of multiobjective test problems and a scalable test problem toolkit. *IEEE T. Evolut. Comput.* **2006**, *10*, 477–506. [CrossRef]
- Ikeda, K.; Kita, H.; Kobayashi, S. Failure of Pareto-based MOEAs: Does non-dominated really mean near to optimal? In Proceedings of the 2001 Congress on Evolutionary Computation, Seoul, Korea, 27–30 May 2001; Volume 2, pp. 957–962.
- 48. Shim, M.B.; Suh, M.W.; Furukawa, T.; Yagawa, G.; Yoshimura, S. Pareto-based continuous evolutionary algorithms for multiobjective optimization. *Eng Comput.* **2002**, *19*, 22–48. [CrossRef]
- Valenzuela-Rendón, M.; Uresti-Charre, E.; Monterrey, I. A non-generational genetic algorithm for multiobjective optimization. In Proceedings of the Seventh International Conference on Genetic Algorithms, East Lansing, MI, USA, 19–23 July 1997; pp. 658–665.
- 50. Vlennet, R.; Fonteix, C.; Marc, I. Multicriteria optimization using a genetic algorithm for determining a Pareto set. *Int. J. Syst. Sci.* **1996**, *27*, 255–260. [CrossRef]



© 2020 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 (http://creativecommons.org/licenses/by/4.0/).