The q-CG Method Applied to the Wave Annihilation Problem

Marluce C. Scarabello*[1], Érica J. C. Gouvêa[1], Aline C. Soterroni[1], Roberto L. Galski[2] and Fernando M. Ramos[3]

[1]Applied Computing Program, National Institute for Space Research, São José dos Campos, Brazil

[2]Satellite Tracking and Control Center, National Institute for Space Research, São José dos Campos, Brazil

[3]Laboratory of Computing and Applied Mathematics, National Institute for Space Research, São José dos Campos, Brazil

November 20, 2013

The history of q-calculus dates back to the beginnings of the last century when, based on pioneering works of Euler and Heine, the English reverend Frank Hilton Jackson developed it in a systematic way [1]. His work gave rise to generalizations of series, functions and special numbers. He reintroduced the concepts of the q-derivative [2] [3] (also known as Jackson's derivative) and introduced the q-integral [4].

Recently, based on Jackson's derivative, a generalization of the classical steepest descent method, called the q-gradient (q-G) method has been proposed for solving unconstrained continuous global optimization problems [5]. The main idea behind this new method is the use of the negative of the q-gradient of the objective function as the search direction. The geometric interpretation of the derivative of $f(x)$, a one-variable function, is simply the slope of the tangent line at a given point x, see Fig. 1. Similarly, the q-derivative has also a straightforward geometric interpretation as the slope of the secant line passing through the points $(x, f(x))$ and $(qx, f(qx))$. It is immediately evident that the q-derivative can be positive or negative depending on the value of the parameter q. For $q = 1$, the q-G method reduces to the classical steepest descent method. Thus, the use of Jackson's derivative can be used as an effective mechanism for escaping from local minima [5].

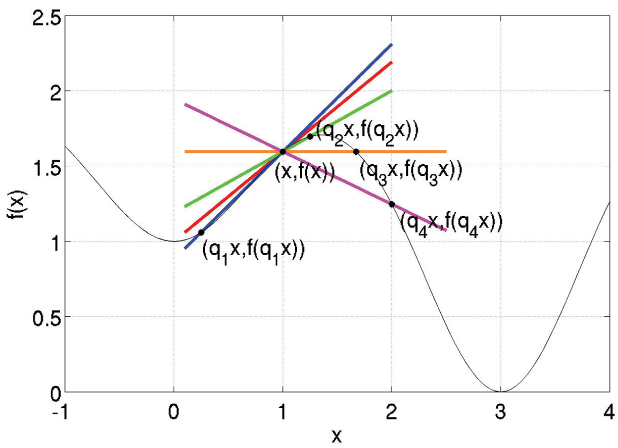

Figure 1: Geometric interpretation of the classical derivative and the q-derivative for different values of the parameter q.

Along the same line, Gouvêa *et al.* (2013) introduced the q-CG method, a generalization of q-derivative of the well-known Conjugate-Gradient method using the concept.

*marluce.scarabello@inpe.br

The q-CG method is a generalization of Fletcher-Reeves conjugate gradient method [6] in which the first search direction is the negative of the q-gradient vector, $-\nabla_{\mathbf{q}} f(\mathbf{x}^k)$, and the following directions are linear combinations of the negative of the q-gradient vector at the current point and previous directions. Values of q_i are drawn from a Gaussian probability density function (pdf), with a standard deviation that decreases as the iterative search proceeds. Starting from σ^0, the standard deviation of the pdf is decreased by $\sigma^{k+1} = \beta \cdot \sigma^k$, where $0 < \beta < 1$ is the reduction factor. As σ^k approaches zero, the values of q_i^k tend to unity and the q-CG method reduces to its classical version.

Gradient-based methods usually perform a linear search along the descent direction. However, depending on the value of q, the negative of the q-gradient may not point to the local descent direction. One way to circumvent this problem is to use a step length α^k that decreases with the iteration k. Here, the initial step length α^0 is reduced by $\alpha^{k+1} = \beta \cdot \alpha^k$, where, for the sake of simplicity, β is the same reduction factor used to compute σ^k. As the step length decreases (and the values of q_i^k, in parallel, tend to unity), a smooth transition to an increasingly local search process occurs.

To evaluate the performance this new approach we applied it to a notoriously difficult multimodal optimization problem, the wave annihilation problem with 56 dimensions; see [7] for details. We compared our results with those obtained with three optimization algorithms published in [7]: simulated parallel annealing within a neighborhood (SPAN), a parallel implementation of the DIRECT algorithm, and a new quase-Newton stochastic algorithm, QNSTOP. The global mininum of wave annihilation problem is zero. For comparison purposes, we use the same experimental setup as in [7]. The stopping criteria was maximum of 10^6 functions evaluations.

The corresponding values of the best parameters σ^0, α^0 and β used are 0.024, 0.102 and 0.9999, respectively. The values of σ^0 and α^0 are normalized by $L = \sqrt{\sum_{i=1}^{n}(\mathbf{x}_{max_i} - \mathbf{x}_{min_i})^2}$, the largest linear length of the search space. The initial standard deviation σ^0 determines how stochastic is the search. The reduction factor β controls the speed of the transition from stochastic to deterministic search. A β close to 1 reduces the risk of being trapped in a local minimum. And, the initial step length α^0, depends heavily on the topology of the search space requiring some empirical exploration. In the end, an appropriate specification of the three free parameters is strictly dependent on the objective function. Although a bad choice may lead to some deterioration in its performance, the q-CG method has shown to be sufficiently robust to still be capable of reaching the global minimum.

Table 1 display the minimum, maximum, first, second, and third quartile of the values objective function for each of the algorithms over the 50 independent experiments.

	Minimum	1 st quartile	2 nd quartile	3 rd quartile	Maximum
DIRECT	8.19×10^{-7}	1.02×10^{-3}	5.76×10^{-3}	5.74×10^{-2}	2.7×10^{-1}
SPAN	2.71	3.35	25.20	26.25	26.62
q-CG	0.06	0.20	0.44	0.74	27.52
QNSTOP	26.64	27.10	27.19	27.30	27.48

Table 1: Results for the wave annihilation problem for $n = 56$.

Althouth surpassed by the DIRECT algorithm, we observe that the q-CG method performs much better than the QNSTOP and the SPAN algorithms. Similarly to these two algorithms, the q-CG was sometimes trapped by a local minimum near 28, what suggests that the algorithm's three free parameters (specially, β) need further adjustments. In spite of this fact, the q-CG method was still able to find a reasonable estimate for the minimum in 80% of the runs (40 out of 50). Overall, these results are promissing and illustrate well the potential of the q-CG method in solving complex optimization problems.

References

[1] T. Ernst, *J. Nonlinear Math. Phys* **10**, 487–525 (2003).

[2] T. W. Chaundy, *J. London Math. Soc.*, **37**, 126–128 (1962).

[3] F. H. Jackson, *Trans. Roy. Soc. Edinburgh*, **46**, 253-281 (1908).

[4] F. H. Jackson, *Quart. J. Pure Appl. Math.*, **41**, 193-203 (1910).

[5] A. C. Soterroni, R. L. Galski and F. M. Ramos, “The q-gradient vector for unconstrained continuous optimization problem”, in *Operations Research Proceddings 2010*, Springer-Verlag Berlin Heidelberg, pp. 365–370 (2011).

[6] R. Fletcher and C. M. Reeves, “Function minimization by conjugate gradients”, in *Computer*, (1964)

[7] David R. Easterling, Layne T. Watson, Michael L. Madigan, Brent S. Castle and Michael W. Trosset “Parallel deterministic and stochastic global minimization of functions with very many minima”, in *Computational Optimization and Applications*, Springer US, New York (2013).

[8] W. W. Hager, R. Rostamian, D. Wang: The wave annihilation technique and the design of nonreflective coatings. *SIAM J. Appl. Math.* 60(4), 1388–1424 (2000)

[9] E. J. C. Gouvêa, A. C. Soterroni, M. C. Scarabello, I. Y. Sumida, F. R. Corrêa, R. L. Galski and F. M. Ramos “Generalization of the Conjugate Gradient Optimization method using q-calculus”, 113 p., In : IPDO 2013 : 4th Inverse problems, design and optimization symposium, 2013 June 26-28, Albi, ed. by O. Fudym, J.-L. Battaglia, G.S. Dulikravich et al., Albi ; Ecole des Mines d’Albi-Carmaux, 2013 (ISBN 979-10-91526-01-2)