KKT Proximity Measure for Testing Convergence in Smooth Multi-objective Optimization

Rupesh Tulshyan
tulshyan@iitk.ac.in

Kalyanmoy Deb
deb@iitk.ac.in

Sunith Bandaru
sunithb@iitk.ac.in

Kanpur Genetic Algorithms Laboratory
Indian Institute of Technology Kanpur
Kanpur 208016, India

ABSTRACT

An earlier study defined a KKT-proximity measure to test the convergence property of an evolutionary algorithm for solving single-objective optimization problems. In this paper, we extend this measure for testing convergence of a set of non-dominated solutions to the Pareto-optimal front in the case of smooth multi-objective optimization problems. Simulation results of NSGA-II on different two and three-objective test problems indicate the suitability of using the proximity measure as a convergence metric for terminating a simulation of an evolutionary multi-criterion optimization algorithm.

Categories and Subject Descriptors

G.1 [**Numerical Analysis**]: nonlinear programming, constrained optimization

General Terms

Algorithms, Performance, Verification

Keywords

Karush-Kuhn-Tucker conditions, multi-objective evolutionary algorithms, termination condition

1. INTRODUCTION

Evolutionary algorithms, including evolutionary multi objective optimization (EMO) algorithms, do not usually use any theoretically motivated termination criterion for stopping a simulation. Often, ad-hoc measures (fixed number of generations, threshold change in objective value, etc.) are used which require some experimentations for setting the appropriate parameter values. However, for smooth problems, there exist derivative based optimality conditions, the most popular being the Karush-Kuhn-Tucker (KKT) conditions which can be used to verify if a point is optimal [2]. An earlier study [3] has revealed the point nature of the KKT conditions for single-objective optimization problems stating that there is no apparent correlation between proximity of a solution from the optimum and the corresponding extent of satisfaction of KKT conditions. The study also suggested a KKT-proximity measure by relaxing the complementary slackness condition so that under certain constraint qualification conditions, a series of points that reduce the KKT-proximity measure in a piece-wise continual manner will eventually converge to a KKT point.

This paper extends the above KKT-proximity measure for multi-objective optimization problems. Since the optimality conditions for multi-objective optimization involve more parameters and additional constraints, a systematic procedure is developed for computing the KKT-proximity measure for each non-dominated point. Ideally, this measure should reduce as the points approach the Pareto-optimal front and should finally be zero when they are exactly on the Pareto-optimal front. Since a set of non-dominated points may have different KKT-proximity values, we propose to use a representative set and decide the termination based on the best, median and worst of these values.

Our primary goal in this paper is to study the reduction of this new multi-objective KKT-proximity measure for non-dominated points evolved over the generations of a popular EMO algorithm (NSGA-II) on some test problems. Thereafter, we explore the possibility of using a KKT-multiplier-driven local search on the non-dominated points of the last generation in cases where the reduction of the KKT-proximity error to zero does not occur.

2. FIRST-ORDER KKT OPTIMALITY CONDITIONS

A typical multi-objective optimization problem with inequality constraints can be formulated as:

$$\begin{array}{rl} \text{Minimize} & \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_k(\mathbf{x})\}, \quad \mathbf{x} \in \mathrm{R}^n \\ \text{Subject to} & g_j(\mathbf{x}) \leq 0, \quad j = 1, 2, \ldots, m. \end{array} \tag{1}$$

If the above multi-objective problem is convex, and if the objective and constraint functions at a feasible point $\bar{\mathbf{x}}$ are continuously differentiable, then the Karush-Kuhn-Tucker (KKT) sufficient condition for $\bar{\mathbf{x}}$ to be Pareto-optimal, is that there exist multipliers $\mathbf{0} < \boldsymbol{\lambda} \in \mathrm{R}^k$ ($0 < \lambda_i \ \forall i$) and $\mathbf{0} \leq \mathbf{u} \in \mathrm{R}^m$ such that:

$$\sum_{i=1}^{k} \lambda_i \nabla f_i(\bar{\mathbf{x}}) + \sum_{j=1}^{m} u_j \nabla g_j(\bar{\mathbf{x}}) = \mathbf{0}, \tag{2}$$

$$u_j g_j(\bar{\mathbf{x}}) = 0, \quad \forall j, \tag{3}$$

A point in the decision space satisfying the above conditions is called a KKT point. Eq. (2) is the *stationarity* condi-

GECCO'11, July 12–16, 2011, Dublin, Ireland.
ACM 978-1-4503-0690-4/11/07.

tion, which emphasizes that at the KKT point, the gradient vectors of the objective function and the constraints remain in an equilibrium, making the point stationary. Eq. (3) is the *complementary slackness* condition, which ensures participation of only the *active* constraints ($g_j(\mathbf{x}) = 0$) in the stationarity condition. While $u_j's$ are the Lagrange multipliers, $\lambda_i's$ can simply be called as the KKT multipliers. Note that in case of single-objective optimization, where $k = 1$, the above conditions reduce to the popular single-objective KKT optimality conditions by putting $\boldsymbol{\lambda} = \lambda_1$ and since $\boldsymbol{\lambda} > \mathbf{0}$, we can impose $\lambda_1 = 1$ without loss of generality. Interested readers can refer to [2] for further details.

3. KKT PROXIMITY MEASURE

Tulshyan et al. [3] discussed the point nature of the KKT conditions for a single-objective case. Calling the norm of the LHS of the Eq. (2) (for $k = 1$) as *KKT Error*, that is,

$$\text{KKT Error} = \|\nabla f_1(\bar{\mathbf{x}}) + \sum_{j=1}^{m} u_j \nabla g_j(\bar{\mathbf{x}})\|, \tag{4}$$

the study showed that the magnitude of the KKT Error at any arbitrary point cannot be related to its proximity to a KKT point, unless the KKT conditions are modified. It further proposed schemes for relaxing the KKT conditions, essentially the complementary slackness condition, in order to obtain a new proximity measure which shows a near-monotonic reducing behavior as points get closer to the optimum and is zero at the actual optimum. Extending the above approach and modifying it for the multi-objective case, we propose the following methodology for computing the KKT-Proximity measure at a *feasible* solution ($\mathbf{x}^k$):

$$\begin{array}{rl} \text{Minimize} & \epsilon_k, \\ \text{Subject to} & \|\sum_{i=1}^{k} \lambda_i \nabla f(\mathbf{x}^k) + \sum_{j=1}^{m} u_j \nabla g_j(\mathbf{x}^k)\|^2 \leq \epsilon_k, \\ & \sum_{j=1}^{m} u_j g_j(\mathbf{x}^k) \geq -\epsilon_k, \\ & \lambda_i \geq 1 \quad \forall i \text{ and } u_j \geq 0 \quad \forall j. \end{array}$$

Here, the variable vector is $(\epsilon_k, \boldsymbol{\lambda}, \mathbf{u})$. The value ϵ_k obtained after the optimization is the KKT-proximity measure at the point $\mathbf{x}^k$.

4. EXPERIMENTAL RESULTS

We analyze the proposed proximity measure with respect to its following properties: (1) it reduces smoothly to zero for a sequence of points approaching the Pareto-optimal front, and (2) it is zero at every point on the Pareto-optimal front. To study its behaviour, we solve two well-known test problems using NSGA-II [1]. For every generation of NSGA-II, we first find the non-dominated front of solutions (NF). A widely distributed sample set P ($P \subset NF$) representing the non-dominated front is obtained using k-means clustering (in objective space). The KKT-proximity measure is then computed for each solution belonging to P.

Figure 1 shows the best, median and worst values of the KKT-proximity measure calculated for the set P in each generation for 10 different NSGA-II runs. The inset shows the track of iterates with best value of the measure. Note how all points in P of the last generation lie exactly on the Pareto-optimal front. The error value also smoothly reduces to zero for all the 10 runs.

On a much harder three-objective DTLZ5 problem, it is observed that though the reduction of the proximity measure

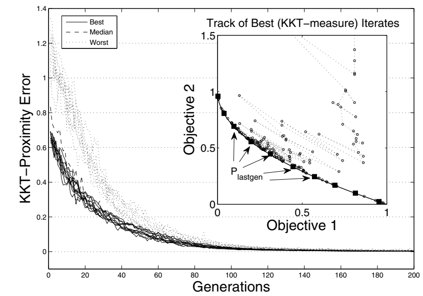

Figure 1: KKT-Proximity measure for ZDT1.

is more or less smooth, the actual values are not adequate for termination even after 100 generations as shown in Figure 2. For such cases, we propose the use of the KKT multipliers obtained at each point in the final non-dominated front for solving a corresponding single objective achievement scalarizing function [4] with $w_i = \lambda_i$. Figure 2 also shows the effect of this local search after the 100-th generation.

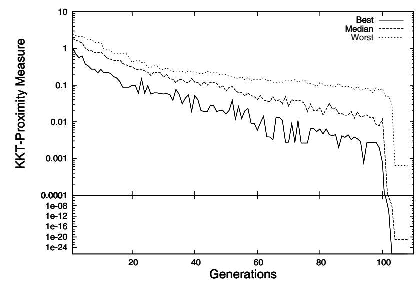

Figure 2: KKT-Proximity Measure for DTLZ5.

5. CONCLUSIONS

We have developed and demonstrated a KKT proximity measure for multi-objective optimization problems. The measure has a smoothly reducing property and converges to zero for most cases but requires a local search effort for the latter in harder problems. The use of these measures for terminating a simulation is considered to be part of future work. Further, this local search can be concurrently used with EMO algorithms for faster convergence.

6. REFERENCES

[1] K. Deb, S. Agarwal, A. Pratap, and T. Meyarivan. A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[2] K. Miettinen. *Nonlinear multiobjective optimization.* Kluwer Academic Publishers, Boston, 1999.

[3] R. Tulshyan, R. Arora, K. Deb, and J. Dutta. Investigating EA solutions for approximate KKT conditions in smooth problems. In *Proceedings of the 12th annual conference on Genetic and evolutionary computation*, GECCO '10, pages 689–696, New York, NY, USA, 2010. ACM.

[4] A. Wierzbicki. The use of reference objectives in multiobjective optimization. In *Multiple Criteria Decision Making: Theory and Applications*, pages 468–486. Spring-Verlag, 1980.