Non-linear Dimensionality Reduction Procedures for Certain Large-Dimensional Multi-objective Optimization Problems: Employing Correntropy and a Novel Maximum Variance Unfolding

Dhish Kumar Saxena and Kalyanmoy Deb

Kanpur Genetic Algorithms Laboratory (KanGAL)
Department of Mechanical Engineering
Indian Institute of Technology, PIN 208016, India
{dksaxena,deb}@iitk.ac.in
http://www.iitk.ac.in/kangal

Abstract. In our recent publication [1], we began with an understanding that many real-world applications of multi-objective optimization involve a large number (10 or more) of objectives but then, existing evolutionary multi-objective optimization (EMO) methods have primarily been applied to problems having smaller number of objectives (5 or less). After highlighting the major impediments in handling large number of objectives, we proposed a principal component analysis (PCA) based EMO procedure, for dimensionality reduction, whose efficacy was demonstrated by solving upto 50-objective optimization problems. Here, we are addressing the fact that, when the data points live on a non-linear manifold or that the data structure is non-gaussian, PCA which yields a smaller dimensional 'linear' subspace may be ineffective in revealing the underlying dimensionality. To overcome this, we propose two new non-linear dimensionality reduction algorithms for evolutionary multi-objective optimization, namely C-PCA-NSGA-II and MVU-PCA-NSGA-II. While the former is based on the newly introduced correntropy PCA [2], the later implements maximum variance unfolding principle [3,4,5], in a novel way. We also establish the superiority of these new EMO procedures over the earlier PCA-based procedure, both in terms of accuracy and computational time, by solving upto 50-objective optimization problems.

1 Introduction

In formulating a multi-objective optimization problem, designers and decision-makers prefer to put every performance index related to the problem as an objective, rather than as a constraint, thereby totalling a large number of objectives. However, evolutionary multi-objective optimization (EMO) methods which find a representative set of solutions in the Pareto-optimal front [6], are, in general, found to be vulnerable to large-objective optimization problems. In [1], we had illustrated this 'curse of dimensionality' on the elitist non-dominated sorting GA or NSGA-II [7]. While solving DTLZ2(10), we had shown that approximately only 4% solutions could come to the Pareto-optimal front. When large number of objectives exist, the probability of having any two arbitrary

S. Obayashi et al. (Eds.): EMO 2007, LNCS 4403, pp. 772–787, 2007.

solutions to be non-dominated to each other increases, as there are many objectives in which a trade-off (one is better in one objective but worse in any other objective) can occur. While dealing with a finite-sized population-based approach, the proportion of non-dominated solutions in the population increases. Since EMO algorithms provide more emphasis to the non-dominated solutions, a large proportion of the old population gets emphasized, thereby not leaving much room for new solutions to be included in the population. This, in effect, reduces the selection pressure for better solutions in the population and results in poor convergence. Over and above this difficulty at algorithmic level, handling large-objectives is not only computationally expensive, it is also a challenge for proper decision-making, as visualizing a Pareto-optimal frontier which is more than three-dimensional, is extremely difficult. Amidst all these, a natural question arises, if it is even worth applying EMO methods, for large-objective problems. In [1], we had highlighted that there may exist large-objective problems, which have redundant objectives, that is, although the problem may have, say **M** objectives but the Pareto-optimal front involves a much lower-dimensional interaction. There, we addressed solving such problems by suggesting a principal component analysis (PCA) based NSGA-II procedure which progresses iteratively from within the search space towards Pareto-optimal region by adaptively finding most anti-correlated lower-dimensional interactions. While PCA yields a smaller dimensional linear subspace that best represents the full data according to a minimum square-error criterion, it may be ineffective in revealing the underlying dimensionality when the data points live on a non-linear manifold (manifolds are spaces that are locally linear but unlike Euclidean subspaces, they can be globally non-linear) or that the data structure is nongaussian. The strength of our earlier proposal of PCA-NSGA-II algorithm emerged from the fact that we could relate most important directions in the data set (in terms of variance) to the importance of objectives, given a multi-objective optimization problem. Now if the determination of important directions in data set is erroneous, the inferences drawn about importance of objectives and hence the determination of redundant objectives will be meaningless. Hence, it would be worthwhile to assess situations in which PCA is likely to extract erroneous directions. Such situations can be best examined under the question: "Does the data live in a low-dimensional subspace" or "Does the data live on a low-dimensional sub manifold", which we examine in the following section.

2 Difficulties with PCA

To highlight difficulties with standard PCA, let us begin with a concrete example of DTLZ5(2,5) [1]. Our earlier proposed PCA-NSGA-II , when tested for this problem, brought out f_3 and f_5 as critical objectives and the rest as redundant. Let us investigate these results in light of two facts. (i) Fact one relates to the property of DTLZ5(2,M) problems where the Pareto-optimal front corresponds to the last and any one of the rest, objective. In this context, declaration of f_3 and f_5 as critical is right. (ii) Fact two relates to the criteria of judging an objective set as critical. PCA-NSGA-II is expected to declare those objectives as critical which apart from being in conflict with each other, also account for variances larger than those declared redundant. From Figure 1, f_4 and f_5 can be seen to account for largest variance, amongst the set of five objectives. This