Multidimensional Data Representation Aimed at the Interpretation of Results from Multiobjective Optimization

Y. Collette[1,2], P. Siarry[1] and H.-I. Wong[2]

[1] Université Paris 12, LERISS, 61 Av. du Général de Gaulle, 94010 Créteil, France
[2] EdF Pôle industrie - RNE/PhR, 1 Av. du Général de Gaulle, 92141 Clamart, France

Abstract. In this study, we present some methods aimed at 2D representation of results from high dimension Multiobjective Optimization problems. First, we present the Principal Component Analysis (PCA) and the Multidimensional Data Scaling (MDS), their efficiencies and their drawbacks. Then, we present a new method called Multidimensional Concordance Discordance Mapping (MCDM). Lastly, we point out the need of methods that perform specific tasks like tradeoff surface conservation, topology conservation and so on.

1 Introduction

Multiobjective optimization is a growing domain of interest. Many publications cover this domain but, due to a lack of tools to perform dimension reduction, most of these works are limited to the study of biobjective or triobjective optimization problems. The need of a good multidimensional data representation method pops up dramatically. Indeed industrial applications are, most of the time, real multiobjective (number of objectives greater than 3) optimization problems. In addition, choosing a multiobjective solution is easier when solutions are represented on the same graph so as to be compared between each others. So, a multidimensional data representation method can be viewed both as a tool for data representation and a tool for decision aid.

2 Principal Component Analysis

The PCA method comes from statictics. It allows to measure distribution of data along virtual axes that minimize the dispersion of points. This method is quite old and is well described in statistics books like [Tenenhaus]. Its efficiency lies in its ability to measure the loss of information. Moreover, this method is reproductible. The main drawback comes from the sensibility of PCA to non-linear data distribution. Indeed, when considering this case, it's very difficult to find virtual axes that minimize the dispersion. So, when projecting data into 2D space, the loss of information may be very important and, the resulting set may be unworkable.

3 Multidimensional Data Scaling

The MDS philosophy is the following: "Conservation of the distances between points". The MDS proceed as follows: first, it chooses random points in a 2D space in same number as in the original space. Second, it moves points so as to minimize a "stress" function. Third, when the stress function reaches a local minimum, the method ends. This method overcomes the main drawback of PCA. The philosophy of MDS allows us to perform comparisons between points when the value of the stress function is very low. The drawbacks of MDS are the following ones: most of the time, it's not reproductible; if the size of the set is very high, the running time may be very long; lastly, when the value of the stress function is high, the resulting set may be unworkable. [Sammon] is a good introducing paper to MDS and [XGVIS] is a good visualization software that perform MDS.

4 Multidimensional Concordance Discordance Mapping

This method is inspired from the ELECTRE I method (Concordance, Discordance). Its aim is to overcome the non-reproductibility and time-consuming drawbacks of preceeding methods. MCDM proceeds as follows: first, we find the best point (composed with the minimum on each coordinate) and the worst point (composed with the maximum on each coordinate). Second, for each point in the original space, we calculate the distance between that point and the best point (called Dmin) and the distance between the point and the worst point (called Dmax). Lastly, we draw these points in a (Dmin, Dmax) space. The method allows us to extract good solutions by searching points that minimize the Dmin coordinate (to be closer to the ideal point) and maximize the Dmax coordinate (to be farther from the worst point).

5 Conclusion

Finally, we can say there's a lot of methods to perform dimension reduction. Some of these methods are complementary but there still exists a need in methods that are "Pareto conservative". Except for this lack, we can consider that the four methods described above can be studied further to improve their applicability on multiobjective problems. We believe that MCDM is particularly suitable for this task, because it makes easier the choice of a good solution among several candidates.

References

[Sammon] J.M. Sammon "A nonlinear mapping for data structure analysis.", IEEE Transactions on Computers volume C-18 number 5 pages 401-409, 1969

[Tenenhaus] M. Tenenhaus "Méthodes statistiques en gestion", DUNOD, 1996

[XGVIS] URL:http://www.research.att.com/~andreas/xgobi