Soft Computing 9 (2005) 13–20 © Springer-Verlag 2003
DOI 10.1007/s00500-003-0329-4

Faster convergence by means of fitness estimation

J. Branke, C. Schmidt

Abstract Evolutionary algorithms usually require a large number of objective function evaluations before converging to a good solution. However, many real-world applications allow for only very few objective function evaluations. To solve this predicament, one promising possibility seems to not evaluate every individual, but to just estimate the quality of some of the individuals. In this paper, we estimate an individual's fitness on the basis of previously observed objective function values of neighboring individuals. Two estimation methods, interpolation and regression, are tested and compared. The experiments show that by using fitness estimation, it is possible to either reach a better fitness level in the given time, or to reach a desired fitness level much faster (roughly half the number of evaluations) than if all individuals are evaluated.

Keywords Evolutionary algorithm, Fitness estimation, Interpolation, Regression, Search history

1 Introduction

For many real-world applications, the number of calls to the objective function is very limited, e.g. because an evaluation is very time consuming or expensive, or because the approach requires user interaction. Evolutionary algorithms (EAs) don't seem to be suited to such problems, since they usually require many evaluations before producing good results.

One possible remedy is to replace some of the costly objective function evaluations by fitness estimates, based on an approximate model of the fitness landscape. Since EAs repeatedly sample the search space at different points, such a model can be built and refined during the actual EA run.

In this paper, we examine the use of two standard local approximation schemes, namely interpolation and regression, which are based on previously evaluated neighboring individuals. In each generation, a fixed percentage of the population is evaluated with the accurate objective function, while the remaining individuals' fitnesses are estimated. The individuals to be evaluated accurately are determined based on their estimated fitness and uncertainty. As we will show, using this approximation scheme speeds up convergence, yielding the same solution quality with significantly fewer actual fitness evaluations.

The paper is structured as follows: First, we survey some related work in Sect. 2. Section 3 discusses a number of important design decisions when using fitness estimation. Details on interpolation and local regression are presented in Sect. 4. The estimation models are evaluated empirically in Sect. 5. The paper concludes with a summary and some ideas for future work.

2 Related work

The idea of approximating fitness values of some individuals based on information generated during the run, i.e. based on fitness values of individuals generated previously, has lately gained increasing attention. A recent survey on this topic can be found in Jin and Sendhoff (2002).

In the simplest case, the fitness of a new individual is derived from its parents' fitnesses. This "fitness inheritance" has first been suggested by Smith et al. (1995). Sastry et al. (2001) reexamined the approach and provide a theoretical analysis, demonstrating that fitness inheritance is indeed able to save a substantial portion of fitness evaluations. They also highlight the influence of the population size and show that for inappropriately large population sizes, the savings may be even higher.

Other approaches attempt to construct a more global model of the fitness landscape based on previous evaluations. Several such approaches can be found in the literature, mainly differing in the model that is used to approximate the landscape, and the selection of data points used to construct the model. Jin et al. (2000, 2002) developed a sophisticated approach using feed-forward neural networks to estimate fitness values, Ratle (1998, 1999) and Emmerich et al. (2002) use a Kriging approach, while El-Beltagy et al. (1999), El-Beltagy (2000) apply Gaussian Process Networks. Rasheed et al. (2002a, b) suggested to cluster the data and construct separate approximation models for the different clusters. In their experiments, quadratic regression performed somewhat better than neural network models.

Statistical approximation models have also been used to improve fitness estimates for optimization problems with stochastic fitness functions. Of course, an individual's fitness can be estimated from sampling the fitness function

Published online: 4 September 2003

J. Branke (✉), C. Schmidt
Institute AIFB, University of Karlsruhe,
76128 Karlsruhe, Germany
E-mail: {branke|csc}@aifb.uni-karlsruhe.de

several times. But as has been shown e.g. in Branke (1998), Branke et al. (2001), or Sano and Kita (2000, 2002), the estimate can be improved significantly by taking into account the fitness estimates of neighboring individuals, i.e. by using a more global approximation model.

Using just the parents to estimate a child's fitness neglects a lot of the information available. On the other hand, while global models take much more information into account, it is hardly possible to build a globally useful approximate model for complex problems with high dimensionality. Also, building such a model from a lot of data generally is a time-consuming process. In this paper, we suggest to build a local model through interpolation or regression, taking into account the intuitively most important information - the closest neighbors. The approach can work with relatively large sets of data points (we usually use the whole history of the search), since for building the actual estimation model, only a few data points are selected.

3 On using fitness estimates

When using an approximate model to estimate the fitness of some of the individuals, there are several design decisions that have to be made which are discussed in the following paragraphs. Some details of interpolation and regression will be discussed in Sect. 4.

Approximation model: There are certainly many possible approximation models, and several have already been used in combination with EAs (cf. Section 2). Simple local models, however, have a number of advantages: they are well-known and established techniques, relatively fast, and take into account the intuitively most important information, the closest neighbors. As we will show in this paper, estimation accuracy is not the only factor that should guide the selection of the approximation model. Different approximation models produce different "virtual" landscapes, some smooth, some rugged, which may pose different challenges to the EA.

Model-data: Every evaluation with the real objective function produces a data point that could potentially be taken into account by the approximation model. However, the larger the data set, the longer it usually takes to construct the model, which is why generally, it is tried to keep the data set small, e.g. by deleting old data points. We keep *all* previous evaluations in a history, and then just select the closest neighbors to build a specific estimation model. That way, all data is preserved and is potentially available for use, while the construction of the model is still fast since only the most relevant data points are actually used to construct the model. Note that finding the closest neighbors requires time linear in the number of samples in the history. But since we assume problems with computationally expensive fitness evaluations, it is save to assume that the number of samples in the history is limited, and that the search effort is negligible compared to the fitness evaluation. If that is not the case, one would have to limit the number of samples in the history e.g. by always deleting the least recently used samples.

Update-Frequency: Depending on the approximation model, it is a time-consuming process to update the model when new data becomes available. That is why in most publications, the model is only updated once in a while. For the case of interpolation or regression, any new evaluation is just added to the history, and is immediately available for use for future estimates. Thus, the approach allows to always use the most up-to-date data.

Elitism: Since the fitness of some individuals in the population is only estimated, it is likely that the best individual in the population is just overevaluated and not necessarily the truly best individual. Therefore, we decided to use the best accurately evaluated individual as an elite [cf. El-Beltagy (2000)].

How many individuals to evaluate: Jin et al. (2000) propose two alternatives: to accurately evaluate a certain percentage of the individuals per generation, or to evaluate all individuals but only in some generations. El-Beltagy (2000) has used a fixed percentage of the population. There are good reasons for both approaches, but evaluating only a portion of the individuals in each generation gives us an additional degree of freedom: we can actively choose which individuals to evaluate from a given set (see next paragraph). It may be argued that the estimation accuracy should increase when the population converges and the individuals become more and more similar. From that, one could conclude that the fraction of individuals evaluated should also increase over time. However, for interpolation and local regression, the accuracy automatically increases over time due to an increasing number of available data points and a focus of the search to an interesting region of the search space. Therefore we decided to evaluate a fixed fraction of the individuals in the population.

Which individuals to evaluate: Intuitively, good individuals are more important, since they will have a stronger influence on the future of the search than bad individuals. Another idea is to favor individuals in regions of the search space with few previous evaluations, because their fitness estimates are less reliable.

Jin et al. (2000) have proposed to evaluate the individuals with best estimated fitness, El-Beltagy (2000) suggested to additionally take into account an individual's distance to the data points used to create the current approximation model (which in a sense corresponds to the uncertainty of the estimate). Finally, Emmerich et al. (2002) suggest a linear combination of estimated fitness and estimated standard error to select the individuals to be evaluated.

In this paper, we examine three possibilities: evaluating the best individuals according to the estimate, evaluating the most uncertain, and evaluating those selected according to a combined measure. For the latter two, an uncertainty measure is required. Since such a measure does not exist for interpolation, we propose the following qualitative uncertainty measure u_j for an individual j:

$$u_j = \frac{1}{\sum_{i=1}^{k} 1/d_{ij}}$$