Global Characterization of the CEC 2005 Fitness Landscapes Using Fitness-Distance Analysis

Christian L. Müller and Ivo F. Sbalzarini

Institute of Theoretical Computer Science and Swiss Institute of Bioinformatics,
ETH Zurich, CH-8092 Zurich, Switzerland
christian.mueller@inf.ethz.ch, ivos@ethz.ch
http://www.mosaic.ethz.ch

Abstract. We interpret real-valued black-box optimization problems over continuous domains as black-box landscapes. The performance of a given optimization heuristic on a given problem largely depends on the characteristics of the corresponding landscape. Designing statistical measures that can be used to classify landscapes and quantify their topographical properties is hence of great importance. We transfer the concept of fitness-distance analysis from theoretical biology and discrete combinatorial optimization to continuous optimization and assess its potential to characterize black-box landscapes. Using the CEC 2005 benchmark functions, we empirically test the robustness and accuracy of the resulting landscape characterization and illustrate the limitations of fitness-distance analysis. This provides a first step toward a classification of real-valued black-box landscapes over continuous domains.

Keywords: Fitness landscape, landscape characterization, fitness-distance correlation, continuous black-box optimization.

1 Introduction

Real-valued optimization problems over continuous parameter spaces ("continuous optimization problems") are ubiquitous in science and engineering. They occur in many practical applications ranging from simple parameter identification in data–model fitting to intrinsic design-parameter optimization in complex technical systems. In a black-box optimization problem only zeroth-order information about the objective function is available. The objective function may be discontinuous or noisy, and analytic gradients or higher-order information may be unknown or inexistent. The diversity of real-world continuous black-box optimization problems hampers a clean classification of problem structure and complexity. Nevertheless, an interesting approach is provided by the landscape metaphor. Ever since Sewall Wright introduced the fitness landscape imagery to evolutionary biology [1] it has been a highly influential concept in many subfields of biology and, more recently, also in combinatorial optimization [2].

We advocate that the fitness landscape perspective also offers a way to establish a more refined analysis of continuous black-box optimization problems. Inspired by our shared visual experience of natural terrains and sceneries, we consider the continuous input variables a high-dimensional landscape domain. Neighborhood or nearness in this landscape domain is defined by a suitable distance metric. We interpret the

C. Di Chio et al. (Eds.): EvoApplications 2011, Part I, LNCS 6624, pp. 294–303, 2011.

scalar objective function value as a height or elevation over the landscape domain. The landscape metaphor encourages a characterization in terms of topographical features, such as valleys, ridges, mountain peaks, and plateaus. In order to underline our view of black-box optimization problems as high-dimensional, complex landscapes we use the term *black-box landscape*. It is conceivable that certain landscape topologies allow efficient optimization while others almost surely lead to failure of a given search heuristic. Despite the tremendous number of novel continuous black-box optimization heuristics published in the past two decades, limited attention has been paid to the question what global topology a certain problem instance has, how to quantify it, and how success or failure of a certain algorithm can be related to landscape topology. In this paper we attempt a first step toward filling this gap. We propose to characterize real-valued black-box landscapes solely based on zeroth-order information, i.e., within a statistical sampling framework. We transfer the well-known concepts of fitness-distance plots and fitness-distance correlation from evolutionary biology and combinatorial optimization into the continuous black-box optimization context.

This paper is structured as follows: We first present a number of conceivable landscape topologies and comment on their impact on the performance of continuous search heuristics. In Section 3 we consider the concepts of fitness-distance plots and fitness-distance correlation. After a short review of the topic we present a number of tools that are applicable to continuous black-box landscapes. We apply these techniques to the IEEE CEC 2005 benchmark functions in Section 4 in order to test their capacity to quantify certain landscape topologies. We conclude this work and suggest future studies in Section 5.

2 Landscape Topologies and Their Impact on Continuous Black-Box Search Heuristics

We sketch a number of conceivable landscape topologies in Fig. 1. The simplest topology is a convex structure (Fig. 1a). This landscape has only one minimum, the global one. If one knows in advance that both the landscape domain and the objective function are convex, there is a wealth of exact and efficient techniques for finding this global minimum. A globally convex single-funnel landscape topology (Fig. 1b) consists of a number of local minima that can be seen as high-frequency perturbations to an underlying convex structure. Functions with this topology, also known as “big valley structures” [2], have been analyzed theoretically by Hu and co-workers [3]. In the evolutionary optimization community, Hansen and Kern [4] pointed out that “if the local optima can be interpreted as perturbations of an underlying unimodal function”, the Evolution Strategy with Covariance Matrix Adaptation (CMA-ES) performs well. The well-known Rastrigin function and the Ackley function belong to this class of landscapes. This observation led Lunacek and Whitley to introduce the dispersion metric [5] as statistical measure that attempts detecting such landscape topologies and may thus serve as a predictor for success or failure of CMA-ES. Both convex and globally convex landscapes are also termed “single-funnel landscapes” [5,6]. Another archetypal landscape structure is the “double-funnel topology” (Fig. 1c). Whenever the funnel that contains the global minimum covers a much smaller part of the domain than the other funnels, this