On the Behaviour of the $(1, \lambda)$-\(\sigma\)SA-ES for a Constrained Linear Problem

Dirk V. Arnold

Faculty of Computer Science, Dalhousie University
Halifax, Nova Scotia, Canada B3H 4R2
dirk@cs.dal.ca

Abstract. This paper analyses the behaviour of the $(1, \lambda)$-\(\sigma\)SA-ES with deterministic two-point rule when applied to a linear problem with a single linear constraint. Equations that describe the single-step behaviour of the strategy are derived and then used to predict the strategy’s multi-step behaviour. The findings suggest that mutative self-adaptation will result in convergence of the $(1, \lambda)$-ES to non-stationary points if the angle between the gradient vector of the objective function and the normal vector of the constraint plane is small. Comparisons with the behaviour of evolution strategies that employ other step size adaptation mechanisms are drawn.

1 Introduction

Step size adaptation mechanisms and constraint handling techniques are important components of evolutionary algorithms (EAs) for constrained real valued optimisation. Most step size adaptation mechanisms have been devised with unconstrained optimisation in mind. Conversely, constraint handling techniques are often designed without much thought to their impact on step size adaptation. Schwefel [16] as early as the 1970s showed that a commonly employed step size adaptation mechanism may result in convergence to non-stationary points in an environment as simple as a linear problem with a single linear constraint.

An understanding of the interaction between step size adaptation mechanisms and constraint handling techniques is crucial for the design of EAs for constrained real valued optimisation. The Handbook of Evolutionary Computation [5, page B2.4:11f] lists a small number of studies that consider the behaviour of evolution strategies applied to simple constrained problems. Rechenberg [14] studies the performance of the $(1+1)$-ES for the axis-aligned corridor model. Schwefel [16] considers the performance of the $(1, \lambda)$-ES in the same environment. Beyer [6] analyses the performance of the $(1+1)$-ES for a constrained, discus-like function. All of those have in common that the constraint planes are oriented such that their normal vectors are perpendicular to the gradient vector of the objective function. In contrast, Schwefel’s work [16] suggests that convergence to

\[ \text{(1+1)-ES for the axis-aligned corridor model.} \]

\[ \text{Schwefel [16] considers the performance of the } (1, \lambda)\text{-ES in the same environment.} \]

\[ \text{Beyer [6] analyses the performance of the } (1+1)\text{-ES for a constrained, discus-like function.} \]

\[ \text{All of those have in common that the constraint planes are oriented such that their normal vectors are perpendicular to the gradient vector of the objective function. In contrast, Schwefel’s work [16] suggests that convergence to} \]

\[ \text{See [9] for an explanation of the } (\mu/\rho + \lambda) \text{ terminology.} \]

© Springer-Verlag Berlin Heidelberg 2012
non-stationary points may occur in situations where the angle between those vectors, which we refer to as the constraint angle, is small. Studying the behaviour of EAs applied to a linear problem with a linear constraint of general orientation is fundamental as owing to Taylor’s theorem, any smooth problem will appear increasingly linear as the step size of the strategy decreases. Arnold and Brauer \cite{3} derive analytical results for the (1 + 1)-ES with success probability based step size adaptation and provide a quantitative confirmation of Schwefel’s findings. More recent work \cite{2, 1} analyses the behaviour of the (1,\(\lambda\))-ES with cumulative step size adaptation for the constrained linear problem and compares two constraint handling techniques. It is found that convergence to non-stationary points in the face of small constraint angles is not unique to success probability based step size adaptation mechanisms.

The goal of this paper is to study the behaviour of the (1,\(\lambda\))-\(\sigma\)SA-ES, i.e., the (1,\(\lambda\))-ES that employs mutative self-adaptation \cite{16, 13} for step size control, when applied to a linear problem with a single linear constraint of general orientation. We assume that constraints are handled by resampling infeasible offspring candidate solutions. The work complements prior research that analyses the behaviour of mutative self-adaptation in unconstrained settings, including that by Hansen \cite{10} who considers unconstrained linear problems, Beyer \cite{7, 8} who considers spherically symmetric functions, and Meyer-Nieberg and Beyer \cite{12} and Arnold and MacLeod \cite{4} who consider ridge functions.

The remainder of this paper is organised as follows. Section 2 briefly describes the problem and the evolution strategy considered. Section 3 derives equations describing the single-step behaviour of the strategy. Section 4 considers multiple time steps and employs the balance criterion proposed by Lunacek and Whitley \cite{11} in order to predict whether the strategy converges to a non-stationary point of the objective function. Section 5 concludes with a brief discussion of the findings and contrasts them with corresponding results for other step size adaptation mechanisms.

2 Problem and Algorithm

As in \cite{3, 2, 1}, throughout this paper we consider the problem of maximising a linear function \(f : \mathbb{R}^n \rightarrow \mathbb{R}, n \geq 2\), with a single linear constraint. We assume that the gradient vector of the objective function forms an acute angle with the normal vector of the constraint plane. Without loss of generality, we choose a Euclidean coordinate system with its origin located on the constraint plane, and with its axes oriented such that the \(x_1\)-axis coincides with the gradient direction \(\nabla f\), and the \(x_2\)-axis lies in the two-dimensional plane spanned by the gradient vector and the normal vector of the constraint plane. The angle between those two vectors is referred to as the constraint angle and denoted by \(\theta\) as illustrated in Fig. 1. Constraint angles of interest are in the open interval \((0, \pi/2)\). The unit normal vector of the constraint plane expressed in the chosen

\textsuperscript{2} Strictly speaking, the task is one of amelioration rather than maximisation, as a finite maximum does not exist. We do not make that distinction here.