Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 4031 – 4047
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Monotonous random search on a torus: Integral upper bounds for the complexity

A. Tikhomirov[a,*], T. Stojunina[b], V. Nekrutkin[b]

[a]*Novgorod State University, Bol'shaya Sankt-Peterburgskaya 41, Veliky Novgorod 173003, Russia*
[b]*Mathematics Department, St. Petersburg University, Universitetsky av. 28, Petrodvorets, St. Petersburg 198504, Russia*

Received 8 January 2006; received in revised form 28 June 2006; accepted 6 April 2007
Available online 27 April 2007

Abstract

The paper consists of three parts. The first part is dedicated to a Markov monotonous random search on a general optimization space. Under certain restrictions, an upper bound for the complexity of search is presented in an integral form, suitable for further analysis. This estimate is applied to the case of a torus, where several specific results on the rate of convergence are obtained with the help of a supplementary optimization problem, discussed in Appendix.

MSC: 90C30; 60J05; 65C05

Keywords: Global optimization; Random search methods; Markov sequences

1. Introduction

Consider a metric space (X, ρ) and suppose that f is a certain objective function defined on X. Let f have a unique minimizer $x_0 = \mathrm{argmin}_{x \in X} f(x)$ and assume that our aim is to find x_0 with accuracy $\varepsilon > 0$. To estimate x_0, we use random search sequences of a special kind.

Set $B_x = \{y \in X, \text{ such that } f(y) \leqslant f(x)\}$ and consider a Markov chain $\{\xi_i, i \geqslant 0\}$ with transition probabilities

$$R_\ell(x, \cdot) = \delta_x(\cdot) P_\ell(x, X \backslash B_x) + P_\ell(x, \cdot \cap B_x), \quad \ell > 0, \tag{1}$$

where δ_x stands for the distribution concentrated at the point x. As usual, $P_\ell(x, \cdot)$ is a probability measure for any $x \in X$ while $P_\ell(\cdot, A)$ is a measurable function for any Borel set $A \subset X$. Obviously, $R_\ell(x, B_x) = 1$; this implies that $f(\xi_i) \leqslant f(\xi_{i-1})$ with probability 1 for all $i > 0$. Therefore, the sequence $\{\xi_i, i \geqslant 0\}$ is called *monotonous*.

It is useful to present a general algorithmic scheme for the simulation of the random sequence $\{\xi_i, i \geqslant 0\}$ with $\xi_0 = x \in X$.

* Corresponding author.
E-mail address: Alexey.Tikhomirov@novsu.ru (A. Tikhomirov).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2007.04.019

Algorithm 1. 1. Set $\xi_0 = x$ and the iteration number $\ell = 1$.
2. Obtain a point η by sampling from the distribution $P_\ell(\xi_{\ell-1}, \cdot)$.
3. If $f(\eta) \leqslant f(\xi_{\ell-1})$ then set $\xi_\ell = \eta$. Alternatively, set $\xi_\ell = \xi_{\ell-1}$.
4. Set $\ell = \ell + 1$ and go to Step 2.

In view of the structure of Algorithm 1, the distributions $P_\ell(x, \cdot)$ can be called *trial transition functions*.

Since we are required to estimate x_0 with accuracy ε, we are interested in the distribution of a random variable $v_\varepsilon = \min\{i \geqslant 0,\ \text{such that } \xi_i \in S_\varepsilon\}$, where S_r stands for the closed ball of radius r with center x_0. Yet it may happen that $\xi_j \in S_\varepsilon$ and $\xi_{j+1} \notin S_\varepsilon$ for some j. To avoid such inconvenience, we introduce a family of sets

$$M_r = \{x \in S_r,\ \text{such that } f(x) < f(y) \text{ for any } y \notin S_r\} \tag{2}$$

and use M_r instead of S_r. Since the Markov chain under consideration is monotonous, it remains in M_r once it has hit it. Thus we introduce a random variable $\tau_\varepsilon = \min\{i \geqslant 0,\ \text{such that } \xi_i \in M_\varepsilon\} \geqslant v_\varepsilon$. If we know the distribution of τ_ε, then we obtain a stopping rule for the random sequence $\{\xi_i, i \geqslant 0\}$. The function $\mathsf{E}_x \tau_\varepsilon$ can be used as a characteristic of *the complexity* of the random search.

Note that the Markov random sequences of Algorithm 1 have a rather simple structure. (It can be considered as too simple for practical needs.) However, it turns out that under certain assumptions on (X, ρ) and very mild restrictions on *f*, the *global* random search method under consideration can be reasonably fast, at least in some asymptotical sense.

More precisely, as it is proved in Nekrutkin and Tikhomirov (1993) (see also Nekrutkin and Tikhomirov (1989) for the particular case of a torus), there exist trial transition functions $P_\ell(x, \cdot)$ such that $\mathsf{E}_x \tau_\varepsilon \leqslant C(X, f, \rho, x) \ln^2 \varepsilon$ for any 'nondegenerate' function *f*. Moreover, these trial distributions do not depend on ℓ and are presented explicitly.

This result seems promising. Indeed, many methods of local optimization need $\mathrm{O}(|\ln \varepsilon|)$ steps to attain the ε-neighborhood of x_0 but require much stronger restrictions concerning the objective function.

If $X \subset \mathbb{R}^d$, then (see, e.g., Zhigljavsky, 1991) the so-called *pure random search* methods need on average $\mathrm{O}(\varepsilon^{-d})$ calculations of the objective function to hit S_ε for any reasonable ρ and any *f*. Note that these methods formally correspond to Algorithm 1 under the supposition that the trial distributions $P_\ell(x, \cdot)$ do not depend on ℓ, x, and ε. On the other hand, as proved in Nekrutkin and Tikhomirov (1993) (see also Tikhomirov and Nekrutkin, 2004 for a review of the whole area), the order of growth of $\mathsf{E}_x \tau_\varepsilon$ typically cannot be smaller than $|\ln \varepsilon|$ for any choice of trial distributions $P_\ell(x, \cdot)$. Therefore, the choice of trial distributions in Algorithm 1 is really important.

In this paper we investigate the Markov monotonous random search sequences of Algorithm 1 in a wider context. Section 2 is devoted to random search methods with a fixed initial point in a general optimization space (X, ρ). All definitions, conditions, and restrictions are collected in Section 2.1. The complexity of a symmetrical random search is investigated in Sections 2.2 and 2.3. The main result (see Theorem 2.1 and Remark 2.3) gives a general and useful upper bound for the complexity.

This estimate is used in Section 3 to analyze Markov monotonous random search on a multi-dimensional torus. The choice of a torus as an example of an *optimization space* (X, ρ) needs explanation.

The 'technical' explanation is very simple. In view of the discussion of Section 2.1, any optimization space under consideration must satisfy several conditions, and the main condition CX1 requires that the volume of a ball of any fixed radius does not depend on the center of this ball. Of course, any bounded subset *X* of $\mathbb{R}^d$ equipped with the Euclidean (or any other 'usual') metric does not satisfy this condition. On the contrary, there is no problem with the condition CX1 for a torus.

On the other hand, the feasible region $X = [0, 1]^d$ is standard for testing optimization methods for different classes of objective functions. Suppose that a unique minimizer x_0 of an objective function *f* belongs to $(0, 1)^d$. Then we can reduce *X* to $\mathbb{I}^d = (0, 1]^d$ and the difference between the Euclidean case and the case of the torus will only be in the choice of the metric ρ. If we equip $\mathbb{I}^d$ with the Euclidean metric, we obtain the standard situation; the choice of the metric (16) (or of any other equivalent metric) leads to the torus.

Still there are a lot of metrics that turn the set $\mathbb{I}^d$ into the *d*-dimensional torus. The 'uniform' metric (16) is the most convenient in view of the simple structure of a ball in this metric.

Returning to results of this paper, we solve three problems related to the *d*-dimensional torus. Firstly, in Section 3.1 we consider a homogeneous symmetrical Markov random search with a random initial point and get the upper bound