PROCEEDINGS OF THE YEREVAN STATE UNIVERSITY

Physical and Mathematical Sciences 2026, **60**(1), p. 23–33

Mathematics

FINITE-DIFFRENCE STOCHASTIC SCHEMES FOR MINIMIZING A STRONGLY QUASICONVEX NON-DIFFERENTIABLE FUNCTION ON $\mathbb{R}^n$. THE NURMINSKII METHOD

R. A. KHACHATRYAN * , Z. B. HOVHANNISYAN **

Chair of Numerical Analysis and Mathematical Modelling, YSU, Armenia

In this work, stochastic approximation methods based on finite differences are investigated for the problem of minimizing quasiconvex functions.

The main result of this work is the derivation of convergence rate estimates for stochastic finite-difference methods in the case of quasiconvex functions.

The obtained results significantly extend the applicability of stochastic finite-difference methods to nonsmooth quasiconvex optimization problems and provide a rigorous justification for their use in black-box settings, where the oracle returns only function values.

https://doi.org/10.46991/PYSUA.2026.60.1.023

MSC2020: Primary: 65k; Secondary 90C30.

Keywords: optimization, quasi convex function, subgradient, fnite-diference scheme, convergence rate.

Introduction. In this study, stochastic approximation techniques based on finite-difference estimators are analyzed for solving minimization problems involving quasiconvex objective functions. Classical approaches to convex optimization typically assume access to exact gradients or stochastic subgradients. In many practical applications, however, only function value information is available.

To address this setting, we consider stochastic quasi-gradient methods that employ randomized finite-difference approximations. In particular, batch two-point schemes are investigated, which produce unbiased estimates of the gradient of a suitably smoothed version of the objective function.

The main contribution of this work is the derivation of convergence rate bounds for stochastic finite-difference algorithms in the case of quasiconvex functions. It is shown that the expected optimality gap decreases at the rate $O\left(\frac{1}{\sqrt{k}}\right)$.

* E-mail: khrafik@ysu.am

** E-mail: zorhh@ysu.am

Finite-difference Stochastic Approximation. The study of quasiconvex functions has been an active area of research in mathematical optimization for several decades [1–4]. A substantial body of work is devoted to the analysis and solution of quasiconvex optimization problems. In particular, [4] was the first to propose an iterative method, the *Normalized Gradient Descent (NGD) method*, and to establish that it attains an ε-approximate solution within a finite number of iterations for differentiable quasiconvex functions.

In large-scale optimization problems, exact gradients or subgradients are often unavailable or expensive to compute. In such settings, *stochastic gradients* are frequently employed and have been widely used for solving both convex and nonconvex optimization problems [5].

However, constructing reliable stochastic gradient estimators is itself a nontrivial task. As a practical alternative, gradients are often approximated using *finite-difference schemes*, which enable efficient derivative-free optimization while retaining favorable convergence properties.

Finite-difference methods for the minimization of nonsmooth Lipschitz functions were initially proposed by A.M. Gupal in the article [6] and subsequently developed in his book [3].

The following notation are used throughout the paper:

- (x,y) denotes the scalar product of vectors $x, y \in \mathbb{R}^n$.
- $\mathbb{E}\Big[\xi\Big]$ denotes the mathematical expectation (expected value) of a random variable ξ.
- $\mathbb{E}\Big[\xi|x\Big]$ denotes the conditional expectation of ξ given x.
- $f'(x)$ denotes the gradient of f at the point x.
- $B_\delta(x)$ denotes the closed ball of radius δ centered at the point x.

Consider the general optimization problem:

$$f(x) \to \min,\ x \in R^n. \tag{1}$$

The sequence x^k is generated by the following recurrence relation:

$$x^{k+1} = x^k - \rho_k g^k, \tag{2}$$

where vector g^k is chosen according to one of the following finite-difference stochastic approximations:

$$g^k = n\frac{f(x_k+\alpha_k u_k) - f(x_k-\alpha_k u_k)}{2\alpha_k}u_k. \tag{3}$$

Here, $\{u_k\}_{k\geq 0}$ is a sequence of independent random vectors, where each u_k is either uniformly distributed on the unit sphere $\mathbb{S}^{n-1}$ or follows the standard Gaussian distribution $\mathcal{N}(0, I_n)$.

This scheme is referred to as the *two-point randomized finite-difference* method [7]. By employing the aforementioned two-point finite-difference scheme, the authors in [7] analyze the sequence $\{x_k\}$ generated by (2) for a strongly convex function f and establish quantitative estimates for its convergence rate.

We note that the definition of strong convexity was first introduced by B.T. Polyak in his paper [8].

Now consider the problem (1) in the case, where the objective function f is strongly quasiconvex.

Most likely, the definition of a strongly quasi-convex function was given by Jovanović in his works [9, 10].

Definition. *[10]. A function f is said to be strongly quasiconvex on the convex set Ω, if there exists $\mu > 0$ such that for all $x, y \in \Omega$ and all $\lambda \in [0,1]$ we have*

$$f(\lambda x + (1-\lambda)y) \leq \max\{f(x), f(y)\} - \mu\lambda(1-\lambda)\|x-y\|^2.$$

In [11], it is shown that a continuous and strongly quasiconvex function has a unique global minimizer over the entire space. We will denote it by x^*.

Let us provide an example of a function that is strongly quasi-convex, but not convex. Consider the function

$$f(x) = -x^2 - x, \qquad x \in [0,1].$$

First compute the derivatives:

$$f'(x) = -2x - 1, \qquad f''(x) = -2.$$

Since

$$f''(x) = -2 < 0 \quad \text{for all } x,$$

the function is concave and therefore not convex.

Next we show that the function is strongly quasi-convex on $[0,1]$.

Recall that a differentiable function f is called μ-strongly quasi-convex if there exists $\mu > 0$ such that for all x, y

$$f(y) \leq f(x) \Rightarrow f'(x)(y-x) \leq -\mu(y-x)^2.$$

Observe that

$$f'(x) = -2x - 1 < 0 \qquad \forall x \in [0,1],$$

hence f is strictly decreasing on $[0,1]$. Therefore

$$f(y) \leq f(x) \quad \Rightarrow \quad y \geq x.$$

Now compute

$$f'(x)(y-x) = (-2x-1)(y-x).$$

Since $x \in [0,1]$, we have

$$-2x - 1 \leq -1.$$

Hence

$$f'(x)(y-x) \leq -(y-x).$$

From $x, y \in [0,1]$ it follows that

$$0 \leq y - x \leq 1,$$

which implies

$$-(y-x) \leq -(y-x)^2.$$

Thus,

$$f'(x)(y-x) \le -(y-x)^2.$$

Therefore the strong quasi-convexity condition holds with $\mu = 1$, i.e.

$$f'(x)(y-x) \le -1 \cdot (y-x)^2.$$

Hence the function $f(x) = -x^2 - x$ is 1-strongly quasi-convex on $[0,1]$, while it is not convex, since $f''(x) < 0$.

Let f be a function that is Lipschitz continuous with constant L_f on the set $M + B_1(0)$, where M is a certain compact set. Define the function $f(x,\alpha)$ as follows:

$$f(x,\alpha) = E_{u\sim B_1(0)}\Big[f(x+\alpha u)\Big] = \frac{1}{Vol(B_1(0))} \int\limits_{\|u\|\le 1} f(x+\alpha u)\,du.$$

The following results holds:

1)

$$|f(x,\alpha) - f(x)| \le L_f \alpha \;\; \forall x \in M. \tag{4}$$

2) The function $f(x,\lambda)$ is differentiable (see [11], Lemma 2.6). By differentiating this function and applying Stokes' theorem, we obtain:

$$f'(x,\alpha) = \frac{n}{\alpha}\mathbb{E}_{u\sim S^{n-1}}\Big[f(x+\alpha u)u\Big].$$

Therefore

$$f'(x,\alpha) = -\frac{n}{\alpha}\mathbb{E}_{u\sim S^{n-1}}\Big[f(x-\alpha u)u\Big].$$

By adding these two equations, we obtain

$$\begin{gathered} f'(x,\alpha) = \frac{n}{2\alpha}\mathbb{E}_{u\sim S^{n-1}}\Big[\Big(f(x+\alpha u) - f(x-\alpha u)\Big)u\Big], \\ \|f'(x,\alpha)\| \le nL_f \;\; \forall x \in M. \end{gathered} \tag{5}$$

From this representation, it follows that the vector g^k is a stochastic gradient of the function $f(x,\alpha)$ at iteration k. Since, according to the first point, this function is a smooth approximation of f, the vector g^k is often referred to in the literature as a stochastic quasi-gradient.

Finite-difference Stochastic Schemes for Minimizing a Strongly Quasiconvex Non-differentiable Function on $\mathbb{R}^n$: Convergence Rate. As already mentioned in the introduction, in many practical problems the gradient of the objective function is not directly available, and only function values can be accessed. In such cases, a widely used approach in the literature is to approximate the subgradient by means of the *randomized two-point estimator* g^k. For the class of smooth and strongly convex function Yu. Nesterov [7] established the following convergence rate (in expected function value) theorem.

Theorem 1. *Let $f : \mathrm{R}^n \to R$ be a strongly convex function and with Lipschitz gradient. Consider the stochastic gradient-free method:*

$$x^{k+1} = x^k - \rho_k g^k,$$

where $u_k \sim \mathcal{N}(0, I_n)$, $\rho_k = 1/k$, $\alpha_k = 1/\sqrt{k}$ and g^k is given by (3). Then

$$\mathbb{E}[f(x^k)] - f(x^*) = O\left(\frac{\ln k}{k}\right).$$

In the proof of the next main theorem we will need the following important property of locally Lipschitz and strongly quasiconvex functions.

Proposition. *[12]. Let $f : \mathbb{R}^n \to \mathbb{R}$ be strongly quasiconvex function with module $\mu > 0$ and locally Lipschitz-continuous. Then, for all $x \in \mathbb{R}^n$, one has*

$$\frac{\mu}{2}\|x - x^*\|^2 \leq (v, x - x^*),\ v \in \partial f(x),$$

where $\partial f(x)$ is Clarke subdifferential.

The Clarke subdifferential of a Lipschitz function f at the point x is defined as

$$\partial f(x) = \{v \in \mathbb{R}^n : f^o(x,h) \geq (v,h) \ \ \forall h \in \mathbb{R}^n\},$$

where

$$f^o(x,h) = \limsup_{\lambda \to 0,\, y \to x} \frac{f(y+\lambda h) - f(y)}{\lambda}.$$

We consider the following iterative minimization procedure for a locally Lipschitz function f, as proposed by Nurminskii in [13]:

$$x^{k+1} = \begin{cases} x^k - \rho_k g^k, & f(x^k) \leq f(x^0) + c, \\ x^0, & f(x^k) > f(x^0) + c, \end{cases} \tag{6}$$

where x^0 is the initial approximation, c is a positive constant, and the vector g^k is iteratively updated following the finite-difference scheme by formula (3).

Denote $M = \{x \in R^n : f(x) \leq f(x^0) + c\}$. In [1] it is proven that the level sets of a strongly quasiconvex and continuous function are compact. Therefore, M is a convex compact set.

The following convergence theorem holds:

Theorem 2. *Let the following assumptions hold:*

1) *$f(x)$ is strongly quasiconvex with modulus μ and locally Lipchitz continuous on R^n.*

2) *We minimize f over the entire space using scheme (6), choosing the parameters*

$$\rho_k = \frac{1}{k}, \quad \alpha_k = \frac{1}{\sqrt{k}}.$$

Then

$$\mathbb{E}[\|x^k - x^*\|^2 = O\left(\frac{1}{\sqrt{k}}\right).$$

The proof of the Theorem requires the following two lemmas.

Lemma 1. *[14]. Let $(\mathcal{F}_k)$, $k \geq 0$, be a filtration, and let*

$$\{\mathbb{X}_k\},\ \{A_k\},\ \{B_k\}$$

be sequences of nonnegative, $\mathcal{F}_k$-adapted random variables satisfying the following conditions:

1) $\mathbb{E}\big[\mathbb{X}_{k+1} \mid \mathcal{F}_k\big] \le \mathbb{X}_k - A_k + B_k,\ k \ge 0;.$

2) $\sum_{k=1}^{\infty} B_k < \infty$ *(almost surely);*

3) $\mathbb{X}_k \ge 0,\ k \ge 0.$

Then $\mathbb{X}_k$ *converges almost surely to a finite limit, and* $\sum_{k=1}^{\infty} A_k < \infty$.

Remark. *This statement remains valid if the condition requiring* $\mathbb{X}_k$ *to be nonnegative is replaced by the assumption that it is bounded from below. Indeed, suppose a lower bound is known*

$$\mathbb{X}_k \ge m \quad \forall k.$$

Let us introduce

$$\mathbb{Z}_k := \mathbb{X}_k - m.$$

Then, if there was already an inequality for the original variable, substituting gives the same type of inequality for $\mathbb{Z}_k$. *Consequently,*

$$\mathbb{Z}_k \to \mathbb{Z}_\infty \quad \textit{almost surely,}$$

and therefore

$$\mathbb{X}_k = \mathbb{Z}_k + m \to \mathbb{Z}_\infty + m \quad \textit{almost surely.}$$

Lemma 2. *The sequence* $\{x^k\}$ *leaves the set*

$$M = \{x \in \mathbb{R}^n : f(x) \le f(x^0) + c\}$$

only finitely many times with probability one and that the sequence $f(x^k)$ *converges almost surely.*

Proof. It is known [3] (Lemma 2.3, p. 36) that the gradient of the function $f(x,\alpha)$ is Lipschitz on the set M, with a Lipschitz constant $L_\alpha = D/\alpha$, where D is some constant. Therefore,

$$f(x^{k+1},\alpha_k) \le f(x^k,\alpha_k) - \rho_k(f'(x^k,\alpha_k),g^k) + \frac{L_{\alpha_k}}{2}\rho_k^2\|g^k\|^2.$$

Let us take the conditional mathematical expectation of both sides; taking into account (5), we obtain

$$\mathbb{E}\big[f(x^{k+1},\alpha_k) \mid x^k\big] \le f(x^k,\alpha_k) - \rho_k\|f'(x^k,\alpha_k)\|^2 + \frac{Dn^2L_f^2\rho_k^2}{2\alpha_k}.$$

Let us bring this inequality into the form of the Robbins lemma. We have

$$\begin{aligned}\mathbb{E}\big[f(x^{k+1},\alpha_{k+1}) \mid x^k\big] &= \mathbb{E}\big[f(x^{k+1},\alpha_k) \mid x^k\big] + \mathbb{E}\big[f(x^{k+1},\alpha_{k+1}) - f(x^{k+1},\alpha_k) \mid x^k\big] \\ &\le f(x^{k+1},\alpha_k) + L_f|\alpha_{k+1}-\alpha_k|.\end{aligned}$$

Thus, in this particular case, the Robbins–Siegmund inequality takes the following form:

$$\mathbb{E}\big[\mathbb{X}_{k+1} \mid \mathcal{F}_k\big] \le \mathbb{X}_k - A_k + C_1\frac{1}{k\sqrt{k}},$$

where
$$A_k = \rho_k \|f'(x^k, \alpha_k)\|^2, \quad \mathbb{X}_k \equiv f(x^k, \alpha_k),$$
and C_1 is some constant.

Science the variable $\mathbb{X}_k$, is bounded from below on the convex compact set M, one can apply the Robbins–Siegmund Lemma with respect to $\mathbb{X}_k$.

Now we consider the case when $x^{k+1} = x^0$. If $f(x^k) > f(x^0) + c$, then $x^{k+1} = x^0$ and
$$\mathbb{X}_{k+1} = f(x^0, \alpha_{k+1}) \le f(x^0) + L_f \alpha_{k+1}.$$

On the other hand,
$$\mathbb{X}_k = f(x^k, \alpha_k) \ge f(x^k) - L_f \alpha_k > f(x^0) + c - L_f \alpha_k.$$

Thus
$$\mathbb{X}_{k+1} - \mathbb{X}_k \le -c + L_f(\alpha_k + \alpha_{k+1}).$$

Let the sequence $\{\mathbb{X}_k\}$ be defined by $\mathbb{X}_k = f(x^k, \alpha_k)$ with $\alpha_k = k^{-1/2}$. Suppose the following bound holds in the no-reset case:
$$\mathbb{E}[\mathbb{X}_{k+1} \mid \mathcal{F}_k] \le \mathbb{X}_k - \rho_k \|f'(x^k, \alpha_k)\|^2 + \frac{C}{k^{3/2}},$$
and in the reset case:
$$\mathbb{X}_{k+1} \le \mathbb{X}_k - c + L_f(\alpha_k + \alpha_{k+1}).$$

Define the Robbins–Siegmund terms as follows:

If no reset occurs:
$$A_k = \rho_k \|f'(x^k, \alpha_k)\|^2, \quad B_k = \frac{C}{k^{3/2}}.$$

If reset occurs:
$$A_k = \max\{0, c - L_f(\alpha_k + \alpha_{k+1})\},$$
$$B_k = \max\{0, L_f(\alpha_k + \alpha_{k+1}) - c\}.$$

Then
$$\mathbb{E}[\mathbb{X}_{k+1} \mid \mathcal{F}_k] \le \mathbb{X}_k - A_k + B_k$$
with $A_k \ge 0$, $B_k \ge 0$, and $\sum_{k=1}^{\infty} B_k < \infty$, where $\mathcal{F}_k = \{x^1, x^2, \ldots, x^k\}$.

If the number of returns were infinite, then there exists a subsequence such that
$$\mathbb{X}_{k_j} = f(x^{k_j}, \alpha_{k_j}) > f(x^0) + c - L_f \alpha_{k_j}$$
and
$$\mathbb{X}_{k_j+1} = f(x^0, \alpha_{k_j+1}) \le f(x^0) + L_f \alpha_{k_j+1}.$$

Since $\mathbb{X}_k \to \mathbb{X}^\infty$, it follows that
$$\mathbb{X}_{k_j} \to \mathbb{X}^\infty \quad \text{and} \quad \mathbb{X}_{k_j+1} \to \mathbb{X}^\infty.$$

However,
$$f(x^0) + c \le \mathbb{X}^\infty \le f(x^0),$$

which is a contradiction. Therefore, the sequence leaves the set M only finitely many times.

It follows from the Robbins–Siegmund lemma that the sequence $f(x^k, \alpha_k)$ converges almost surely. Since

$$|f(x^k, \alpha_k) - f(x^k)| \leq L_f \alpha_k \to 0,$$

it follows from inequality (4). Hence $f(x^k)$ converges almost surely.

Now, let us consider the case of strongly quasi-convex functions. □

Proof of Theorem 2. We have

$$f'(x, \alpha) = \mathbb{E}_{u \sim B_1(0))}[v(x + \alpha u)],$$

where $v : M \to \mathbb{R}^n$ is any measurable function such that

$$v(x) \in \partial f(x) \quad \text{for all } x \in M,$$

with $\partial f(x)$ denoting the Clarke subdifferential of f.

Indeed, since f is locally Lipschitz, by Rademacher's theorem it is differentiable a.e.; denote the set of differentiability by D, with $|D^c| = 0$. Moreover, $\|f'(x)\| \leq L_f$ a.e.

For any $h \in \mathbb{R}^n$,

$$\frac{f(x+th, \alpha) - f(x, \alpha)}{t} = \mathbb{E}_{u \sim B_1(0)} \left[\frac{f(x+th+\alpha u) - f(x+\alpha u)}{t} \right].$$

For a.e. u such that $x + \alpha u \in D$, we have

$$\lim_{t \to 0} \frac{f(x+th+\alpha u) - f(x+\alpha u)}{t} = (f'(x+\alpha u), h).$$

Moreover,

$$\left| \frac{f(x+th+\alpha u) - f(x+\alpha u)}{t} \right| \leq L_f \|h\|,$$

so by dominated convergence,

$$\lim_{t \to 0} \frac{f(x+th, \alpha) - f(x, \alpha)}{t} = \mathbb{E}_{u \sim B_1(0)} \left[(f'(x+\alpha u), h) \right].$$

Thus

$$f'(x, \alpha) = \mathbb{E}_{u \sim B_1(0)}[f'(x + \alpha u)].$$

Since $f'(x) \in \partial f(x)$ for a.e. x, and $\partial f(x)$ is nonempty, convex, compact-valued, there exists a measurable selection $v(x) \in \partial f(x)$ such that

$$v(x) = f'(x) \quad \text{a.e.}$$

Hence

$$f'(x + \alpha u, \alpha) = v(x + \alpha u) \quad \text{a.e. in } u,$$

which implies

$$\mathbb{E}_{u \sim B_1(0)}[f'(x + \alpha u)] = \mathbb{E}_{u \sim B_1(0)}[v(x + \alpha u)].$$

Therefore,

$$f'(x, \alpha) = \mathbb{E}_{u \sim B_1(0)}[v(x + \alpha u)].$$

By Proposition, we also have

$$\frac{\mu}{2}\|x-x^*\|^2 \le (v, x-x^*), \quad v\in\partial f(x) \quad \forall x\in\mathbb{R}^n.$$

Then

$$(f'(x,\alpha), x-x^*) \ge \frac{\mu}{2}\mathbb{E}_{u\sim B_1(0)}\left[\|x+\alpha u-x^*\|^2\right] - \alpha\mathbb{E}_{u\sim B_1(0)}\left[(v(x+\alpha u),u)\right] \tag{7}$$

$$\ge \frac{\mu}{2}\|x-x^*\|^2 + \frac{\mu}{2}\alpha^2 - \alpha L_f,$$

where $L_f \ge \|v\|$, $v\in\partial f(x)$, $x\in M+B_1(0)$.

Since, according to Lemma 2, the sequence $\{x^k\}$ leaves the set M only a finite number of times, starting from some index k_0, we can estimate the difference $\|x^k-x^*\|^2$ as follows:

$$\|x^{k+1}-x^*\|^2 = \|x^k-x^*\|^2 - 2\rho_k(g^k, x^k-x^*) + \rho_k^2\|g^k\|^2,\ k\ge k_0.$$

From this, we obtain

$$\mathbb{E}\left[\|x^{k+1}-x^*\|^2\right] \le \mathbb{E}\left[\|x^k-x^*\|^2\right] - 2\rho_k\left(\frac{\mu}{2}\|x^k-x^*\|^2 + \frac{\mu}{2}\alpha_k^2 - \alpha_k L_f\right) + \rho_k^2 E\left[\|g^k\|^2\right].$$

Substituting ρ_k, α_k, m_k, there exists a constant $R>0$ such that

$$\mathbb{E}\left[\|x^{k+1}-x^*\|^2\right] \le \left(1-\frac{1}{k}\right)\mathbb{E}\left[\|x^k-x^*\|^2\right] + \frac{R}{k\sqrt{k}}.$$

Define $a_k = \mathbb{E}\left[\|x^k-x^*\|^2\right]$. Then

$$a_{k+1} \le \left(1-\frac{1}{k}\right)a_k + \frac{R}{k^{3/2}}.$$

Multiplying both sides by k and defining $b_k=(k-1)a_k$, we get

$$b_{k+1} \le b_k + \frac{R}{\sqrt{k}}.$$

Summing from $k=1$ to T gives

$$b_{T+1} \le b_1 + R\sum_{j=1}^{T}\frac{1}{\sqrt{j}} \le 2R\sqrt{T} + b_1.$$

Therefore,

$$a_{T+1} \le \frac{2R}{\sqrt{T}} = O\Big(\frac{1}{\sqrt{T}}\Big).$$

Hence, we obtain

$$\mathbb{E}\left[\|x^k-x^*\|^2\right] = O\Big(\frac{1}{\sqrt{k}}\Big).$$

□

Conclusion. Stochastic finite-difference methods were studied for minimizing quasiconvex functions when only function values are available. Using randomized two-point estimators, unbiased approximations of the gradient of a smoothed objective were constructed. It was shown that the expected optimality gap decreases at the rate $O\left(\frac{1}{\sqrt{k}}\right)$.

Received 16.03.2026
Reviewed 08.04.2026
Accepted 16.04.2026

REFERENCES

1. Lara F. On Strongly Quasiconvex Functions: Existence Results and Proximal Point Algorithm. *JOTA* **192** (2022), 891–911.
https://doi.org/10.1007/s10957-021-01996-8
2. Lara F., Marcavillaca R.T., Yuong P.T. Characterizations, Dynamical Systems and Gradient Methods for Strongly Quasiconvex Functions. (2024).
https://doi.org/10.48550/arXiv.2410.03534
3. Mikhalevich V.S., Gupal A.M., Norkin V.I. *Methods of Nonconvex Optimization*. Moscow, Nauka (1983).
4. Nesterov Yu.E. Minimization Methods for Non-smooth Convex and Quasiconvex Functions. *Matekon* **29** (1984), 519–531.
5. Shamir O., Zhang T. Stochastic Gradient Descent for Non-smooth Optimization. (2012).
https://doi.org/10.48550/arXiv.1212.1824
6. Gupal A.M. *Algorithms for Finding the Extremum of Nondifferentiable Functions with Constraint.* Kyiv, Institute of Cybernetics, Preprint (1976).
7. Nesterov Yu.E., Spokoiny V. Random Gradient-Free Minimization of Convex Functions. *Foundations of Computational Mathematics* **17** (2017), 527–566.
https://doi.org/10.1007/s10208-015-9296-2
8. Polyak B.T. Existence Theorems and Convergence of Minimizing Sequences in Extremum Problems with Restrictions. *Soviet Math. Dokl.* **7** (1966), 72–75.
9. Jovanovic M.V. A Note on Strongly Convex and Quasiconvex Functions. *Math. Notes* **60** (1996), 584–585.
https://doi.org/10.1007/BF02309176
10. Jovanovic M.V. Strongly Quasiconvex Quadratic Functions. *Publications de l'institut Mathematique Nouvelle Serie* **53** (1993), 153–156.
11. Hazan E. *Introduction to Online Convex Optimization*. Essential Knowledge, Boston-Delt (2016).
12. Grad S.M., Lara F., Marcavillaca R.T. *Strongly Quasiconvex Functions, What We Know (So Far).* (2024).
https://doi.org/10.48550/arXiv.2410.23055
13. Nurminskii E.A. *Numerical Methods for Solving Deterministic and Stochastic Minimax Problems*. Kiev, Naukova Dumka (1979).
14. Robbins H., Siegmund D. A Convergence Rate Theorem for Non Negative Almost Supermartingales and Some Applications. *Optimizing Methods in Statistics* **8** (1971), 233–257.
https://doi.org/10.1016/B978-0-12-604550-5.50015-8

Ռ. Ա. ԽԱՉԱՏՐՅԱՆ, Զ. Բ. ՀՈՎՀԱՆՆԻՍՅԱՆ

ՎԵՐՋԱՎՈՐ ՏԱՐԲԵՐԱԿԱՆ ՍՏՈԽԱՍՏԻԿ ՍԽԵՄԱՆԵՐ՝ ՈՒԺԵՂ ՔՎԱԶԻՈՒՌՈՒՑԻԿ ՈՉ ԴԻՖԵՐԵՆՑԵԼԻ ՖՈՒՆԿՑԻԱՅԻ ՄԻՆԻՄԻԶԱՑԻԱՅԻ ՀԱՄԱՐ $\mathbb{R}^n$ ՏԱՐԱԾՈՒԹՅՈՒՆՈՒՄ։ ՆՈՒՐՄԻՆՍԿՈՒ ՄԵԹՈԴԸ

Աշխատանքում ուսումնասիրվում են ստոխաստիկ մոտարկման մեթոդներ՝ հիմնված վերջավոր տարբերությունների վրա՝ ուժեղ քվազիուռուցիկ ֆունկցիաների մինիմիզացման խնդրի համար։

Աշխատանքի հիմնական արդյունքը հանդիսանում է վերջավոր տարբերությունների ստոխաստիկ մեթոդների զուգամիտության արագության գնահատականների ստացումը քվազիուռուցիկ ֆունկցիաների դեպքում։

Ստացված արդյունքները էականորեն ընդլայնում են վերջավոր տարբերությունների ստոխաստիկ մեթոդների կիրառման ոլորտը ոչ ողորկ քվազիուռուցիկ օպտիմիզացիայի խնդիրներում՝ ապահովելով ալգորիթմի խիստ հիմնավորումն այն դեպքերում, երբ օրակուլին հասանելի է միայն սև արկղի ինֆորմացիան։

Р. А. ХАЧАТРЯН, З. Б. ОГАНИСЯН

СТОХАСТИЧЕСКИЕ СХЕМЫ КОНЕЧНЫХ РАЗНОСТЕЙ ДЛЯ МИНИМИЗАЦИИ СИЛЬНО КВАЗИВЫПУКЛОЙ НЕДИФФЕРЕНЦИРУЕМОЙ ФУНКЦИИ НА $\mathbb{R}^n$: МЕТОД НУРМИНСКОГО

В данной работе исследуются методы стохастического приближения на основе конечных разностей для задачи минимизации сильно квазивыпуклых функций.

Основным результатом работы является вывод оценок скорости сходимости для стохастических методов конечных разностей в случае квазивыпуклых функций.

Полученные результаты существенно расширяют область применения стохастических методов конечных разностей к негладким задачам квазивыпуклой оптимизации, обеспечивают строгое обоснование алгоритма в случаях, когда оракулу доступна только информация черного ящика.