Liet. matem. rink., **44**, *spec. nr.*, 2004, 603–608

Vienalaikio trikdžio stochastinės aproksimacijos konvergavimas Lipšico funkcijų klasėje

Vaida BARTKUTĖ, Leonidas SAKALAUSKAS (MII)
el. paštas: vaidaba@one.lt, sakal@ktl.mii.lt

1. Įvadas

Optimizavimo uždavinių, kai tikslo funkcija yra nediferencijuojama, sprendimas yra aktuali teorinė ir praktinė problema [1, 3]. Efektyvius skaitmeninius nediferencijuojamo optimizavimo algoritmus leidžia sukurti SPSA metodas, nes stochastiniam gradientui įvertinti kiekvienoj iteracijoj pakanka žinoti tik vieną arba dvi tikslo funkcijos reikšmes. Žinomi SPSA algoritmai remiasi tikslo funkcijos suglodinimu naudojant Bernulio modelį [4] arba suglodinimo operatorių, tolygiai pasiskirsčiusį daugiamačiame hiperkube [1]. Šiame darbe įvedamas SPSA algoritmas, kuris remiasi Gauso suglodinimo operatoriumi, bei įrodomas šio algoritmo konvergavimas Lipšico funkcijų klasėje.

2. Optimizavimo metodo konvergavimo tyrimas

Nagrinėsime optimizavimo uždavinį

$$f(x) \to \min, \tag{1}$$

čia $f\colon \Re^n \to \Re$ funkcija, tenkinanti Lipšico sąlygą su konstanta K, $\min_{x\in\Re^n} f(x) - \infty$, o $\partial f(x)$ – jos subdiferencialas. Iš Lipšico sąlygos išplaukia: $\sup_{\substack{g\in\partial f(x)\\ x\in\Re^n}} |g| \leqslant K$ [1]. Pažymėkime stacionarių taškų ir stacionarių reikšmių aibes: $X^* = \{x \mid 0 \in \partial f(x)\}$ ir $F^* = \{z \mid z = f(x), x \in X^*\}$. Funkcijos suglodinimui naudosime Gauso operatorių: $\overline{f}(x,\sigma) = Ef(x+\sigma\xi)$, $\xi \sim N(0, I_n)$, čia $\sigma \geqslant 0$ suglodinimo parametras. Užrašysime suglodintą funkciją integraliniu pavidalu:

$$\begin{aligned}\overline{f}(x,\sigma) &= (2\pi)^{-\frac{n}{2}} \int_{\Re^n} f(x+\sigma y)\cdot p(y)\,\mathrm{d}y \\ &= (2\pi)^{-\frac{n}{2}} \cdot \sigma^{-n} \int_{\Re^n} f(y)\cdot p\Big(\frac{y-x}{\sigma}\Big)\,\mathrm{d}y, \quad \text{čia } p(y) = \exp\Big(-\frac{\|y\|^2}{2}\Big).\end{aligned}$$

Pastebėsime, kad tokiu būdu suglodintos funkcijos yra be galo diferencijuojamos įprastine prasme. Iš Rademacherio teoremos [1] seka, kad apibendrinto gradiento daugiareikšmiškumo taškų aibės Borelio matas lygus nuliui. Kadangi $f(x)$ apibendrintas gradientas yra aprėžtas, o nagrinėjamas matas yra absoliučiai tolydus, galime įvesti suglodintos funkcijos gradientą

$$\overline{g}(x,\sigma)=\frac{\mathrm{d}\overline{f}(x,\sigma)}{\mathrm{d}x}=E\partial f(x+\sigma\xi), \tag{2}$$

kur daugiareikšmiškumo taškuose imama bet kuri apibendrinto gradiento reikšmė. Nesunku įsitikinti, kad taip pat

$$\overline{g}(x,\sigma)=(2\pi)^{-\frac{n}{2}}\cdot\sigma^{-n}\int_{\Re^n}f(y)\cdot\frac{\partial p\left(\frac{y-x}{\sigma}\right)}{\partial x}\,\mathrm{d}y=\frac{E(f(x+\sigma\xi)-f(x))\cdot\xi}{\sigma}. \tag{3}$$

Iš (3) seka, kad funkcijos stochastinį gradientą galima įvertinti tokiu būdu:

$$g(x,\sigma,\xi)=\frac{(f(x+\sigma\xi)-f(x))\cdot\xi}{\sigma}. \tag{4}$$

Aišku, kad $Eg(x,\sigma,\xi)=\overline{g}(x,\sigma)$. Iš Lipšico sąlygos išplaukia, kad egzistuoja ir yra tolygiai aprėžti visi įverčio (4) baigtinės eilės momentai, kai $\sigma>0$, pvz., $E\|g(x,\sigma,\xi)\|^2\leqslant C_2=K^2\cdot n\cdot(n+2)$.

1 LEMA. *Jei $f(x)$ tenkina Lipšico sąlygą su konstanta K, tai*

$$\left|\overline{f}(x,\sigma_1)-\overline{f}(x,\sigma_2)\right|\leqslant C_1\cdot K\cdot|\sigma_1-\sigma_2|,$$

$$\left\|\overline{g}(x_1,\sigma)-\overline{g}(x_2,\sigma)\right\|\leqslant\frac{C_1\cdot K\cdot\|x_1-x_2\|}{\sigma},$$

čia $C_1=E\|\xi\|$.

Įrodymas. Iš Lipšico sąlygos ir (4) formulės gauname:

$$\begin{aligned}
\left|\overline{f}(x,\sigma_1)-\overline{f}(x,\sigma_2)\right|&=\left|E\big(f(x+\sigma_1\xi)-f(x+\sigma_2\xi)\big)\right|\\
&\leqslant E\left|f(x+\sigma_1\xi)-f(x+\sigma_2\xi)\right|\leqslant K\cdot|\sigma_1-\sigma_2|\cdot E\|\xi\|\leqslant C_1\cdot K\cdot|\sigma_1-\sigma_2|,\\
\left\|\overline{g}(x_1,\sigma)-\overline{g}(x_2,\sigma)\right\|&=\left\|\frac{E(f(x_1+\sigma\xi)-f(x_1)-f(x_2+\sigma\xi)+f(x_2))\cdot\xi}{\sigma}\right\|\\
&=\left\|\frac{E(f(x_1+\sigma\xi)-f(x_2+\sigma\xi))\cdot\xi}{\sigma}\right\|\\
&\leqslant\frac{E\|(f(x_1+\sigma\xi)-f(x_2+\sigma\xi))\cdot\xi\|}{\sigma}\leqslant\frac{K\cdot\|x_1-x_2\|\cdot E\|\xi\|}{\sigma}.
\end{aligned}$$

2 LEMA. *Tegul $f(x)$ tenkina Lipšico sąlygą su konstanta K, be to, bet kuriam $\varepsilon>0$ galima rasti tokį $\delta>0$, kad visiems y, $\|y-x\|\leqslant 2\delta$, kur funkcija $f(x)$ diferencijuojama įprastine prasme, tolygiai x atžvilgiu galioja $\min_{z\in\partial f(x)}\|\frac{\partial f(y)}{\partial y}-z\|\leqslant\frac{\varepsilon}{2}$. Tegul dar $\sigma_i\to 0$. Tuomet bet kuriam $\varepsilon>0$ galima rasti tokį k, kad visiems y, $\|y-x\|\leqslant\delta$, kur funkcija $f(x)$ diferencijuojama įprastine prasme, ir visiems $i\geqslant k$ tolygiai x atžvilgiu galioja $\min_{z\in\partial f(x)}\|\overline{g}(y,\sigma_i)-z\|\leqslant\varepsilon$.*

Įrodymas. Žymėsime $I(A)$ aibės A indikatorių. Tegul $z_1 \in \partial f(x)$ toks, kad

$$\left\| E\left(\left(g(y+\sigma_i\xi)-z_1\right) \cdot I\left(\|\xi\| \leqslant \frac{\delta}{\sigma_i} \right) \right) \right\| = \min_{z \in \partial f(x)} \left\| E\left(\left(g(y+\sigma_i\xi)-z\right) \cdot I\left(\|\xi\| \leqslant \frac{\delta}{\sigma_i} \right) \right) \right\|,$$

o k yra toks, kad $\sigma_i \leqslant \frac{\delta}{2} \cdot \sqrt{\frac{\varepsilon}{K \cdot n}}$ visiems $i \geqslant k$. Tuomet turime, kad

$$E\left(\|g(y+\sigma_i\xi) - z_1\| \cdot I\left(\|\xi\| > \frac{\delta}{\sigma_i} \right) \right) \leqslant 2K \cdot P\left(\|\xi\| > \frac{\delta}{\sigma_i} \right) \leqslant \frac{2 \cdot K \cdot n \cdot \sigma_i^2}{\delta^2} \leqslant \frac{\varepsilon}{2}.$$

Tegul $H\colon \mathfrak{R}^n \to \partial f(x)$ yra selektorius, tai yra tokia mati funkcija, kad $\|g(y) - H(y)\| = \min_{z \in \partial f(x)} \|g(y) - z\|$ (žr. teoremą apie selektorius [2]), čia $g(y) = \frac{\partial f(y)}{\partial y}$. Tuomet, kai $\|y - x\| \leqslant \delta$, seka, kad

$$\begin{aligned}
\min_{z \in \partial f(x)} \|\overline{g}(y, \sigma_i) - z\| &\leqslant \left\| E\left(g(y+\sigma_i\xi) - z_1\right) \right\| \\
&= \left\| E\left(\left(g(y+\sigma_i\xi)-z_1\right) \cdot I\left(\|\xi\| \leqslant \frac{\delta}{\sigma_i} \right) \right) + E\left(\left(g(y+\sigma_i\xi)-z_1\right) \cdot I\left(\|\xi\| > \frac{\delta}{\sigma_i} \right) \right) \right\| \\
&\leqslant \min_{z \in \partial f(x)} \left\| E\left(\left(g(y+\sigma_i\xi)-z\right) \cdot I\left(\|\xi\| \leqslant \frac{\delta}{\sigma_i} \right) \right) \right\| \\
&\quad + \left\| E\left(\left(g(y+\sigma_i\xi)-z_1\right) \cdot I\left(\|\xi\| > \frac{\delta}{\sigma_i} \right) \right) \right\| \leqslant \varepsilon,
\end{aligned}$$

nes

$$\begin{aligned}
\frac{\varepsilon}{2} &\geqslant E\left(\left(\min_{z \in \partial f(x)} \|(g(y+\sigma_i\xi) - z)\| \right) \cdot I\left(\|\xi\| \leqslant \frac{\delta}{\sigma_i} \right) \right) \\
&= E\left\| \left(g(y+\sigma_i\xi) - H(y+\sigma_i\xi)\right) \cdot I\left(\|\xi\| \leqslant \frac{\delta}{\sigma_i} \right) \right\| \\
&\geqslant \left\| E\left(g(y+\sigma_i\xi) \cdot I\left(\|\xi\| \leqslant \frac{\delta}{\sigma_i} \right) \right) - E\left(H(y+\sigma_i\xi) \cdot I\left(\|\xi\| \leqslant \frac{\delta}{\sigma_i} \right) \right) \right\| \\
&\geqslant \min_{z \in \partial f(x)} \left(\left\| E\left(\left(g(y+\sigma_i\xi) - z\right) \cdot I\left(\|\xi\| \leqslant \frac{\delta}{\sigma_i} \right) \right) \right\| \right),
\end{aligned}$$

kadangi iš subdiferencialo iškilumo seka $E\left(H(y+\sigma_i\xi) \,\middle|\, \|\xi\| \leqslant \frac{\delta}{\sigma_i} \right) \in \partial f(x)$.

Įveskime seką

$$x^{k+1} = x^k - \rho_k \cdot g^k, \quad k = 1, 2, \ldots, \tag{5}$$

čia $g^k = g(x^k, \sigma_k, \xi_{k+1})$ stochastinio gradiento, apskaičiuoto pagal (4), reikšmė taške x^k, $\xi_1, \xi_2, \ldots$ yra nepriklausomos ξ kopijos, ρ_k – skaliarinis daugiklis, o σ_k – glodinimo parametro reikšmė iteracijoje k, x^0 – pradinis taškas. Seka (5) konverguoja į (1) sprendinį beveik tikrai (b. t.), kai iteracijų skaičius neaprėžtai didėja, jei yra išpildytos tam tikros sąlygos.

TEOREMA. *Jei funkcija* $f: \Re^n \to \Re$

A) *tenkina Lipšico sąlygą su konstanta* K,

B) *yra aprėžta iš apačios,*

C) $\underline{\lim}_{k\to\infty} \inf_{g\in\partial f(x^k)} |g| > 0$ *bet kokiai sekai* $\{x^k\}$, *jei* $\underline{\lim}_{k\to\infty} \|x^k\| = \infty$,

D) *bet kuriam* $\varepsilon > 0$ *galima rasti tokį* $\delta > 0$, *kad visiems* y, $\|y - x\| \leqslant 2\delta$, *kur funkcija* $f(x)$ *diferencijuojama įprastine prasme, tolygiai* x *atžvilgiu galioja* $\min_{z\in\partial f(x)} \left\| \frac{\partial f(y)}{\partial y} - z \right\| \leqslant \frac{\varepsilon}{2}$,

E) *funkcijos stacionarių reikšmių aibė* F^* *neturi vidinių taškų,*

F) $\{\rho_k\}$, $\{\sigma_k\}$ *yra neneigiamų skaičių sekos, tokios, kad*
$\sum_{k=1}^{\infty} \rho_k = \infty$, $\sum_{k=1}^{\infty} \rho_k^2 < \infty$, $\sigma_k \to 0$, $|\sigma_k - \sigma_{k-1}|/\rho_k \to 0$, $\frac{\rho_k}{\sigma_k} \to 0$,

tuomet visi sekos $\{x^k\}$, *apibrėžiamos formule* (5), *ribiniai taškai b.t. priklauso* X^*.

Žymėkime $\{\Theta_m\}_{m=0}^{\infty}$ σ-algebrų srautą, sugeneruotą sekos $\{x^m\}_{m=0}^{\infty}$.

3 LEMA. *Jei* $\{\varphi_k\}$ *yra tokia atsitiktinių dydžių seka, suderinta su* $\{\Theta_m\}_{m=0}^{\infty}$, *kad* $\sum_{k=0}^{\infty} E\varphi_k^2 < \infty$, *tai* $\sum_{k=0}^{\infty} \varphi_k - E(\varphi_k|\Theta_{k-1}.)$ *b.t. konverguoja.*

Lemos įrodymą galima rasti ([6], 1 lema, 2 priedo 4 skyrius).

Teoremos įrodymas. Iš 1 lemos turime:

$$\overline{f}(x^{k+1}, \sigma_{k+1}) \leqslant \overline{f}(x^{k+1}, \sigma_k) + C_1 \cdot K \cdot |\sigma_{k+1} - \sigma_k|. \tag{6}$$

Iš Lagranžo formulės seka:

$$\begin{aligned}
\overline{f}(x^{k+1}, \sigma_k) &= \overline{f}(x^k, \sigma_k) + \Big(\overline{g}\big(x^k + \tau(x^{k+1} - x^k), \sigma_k\big) - \overline{g}(x^k, \sigma_k)\Big)^T \cdot (x^{k+1} - x^k) \\
&\quad + \overline{g}(x^k, \sigma_k)^T \cdot (x^{k+1} - x^k) \\
&\leqslant \overline{f}(x^k, \sigma_k) + \overline{g}(x^k, \sigma_k)^T \cdot (x^{k+1} - x^k) + \frac{C_1 \cdot K}{\sigma_k} \cdot \|x^{k+1} - x^k\|^2 \\
&= \overline{f}(x^k, \sigma_k) - \rho_k g(x^k, \sigma_k, \xi_{k+1})^T \overline{g}(x^k, \sigma_k) + \frac{C_1 \cdot K \cdot \rho_k^2}{\sigma_k} \left\| g(x^k, \sigma_k, \xi_{k+1}) \right\|^2 \\
&= \overline{f}(x^k, \sigma_k) + \frac{C_1 \cdot K}{\sigma_k} \cdot \rho_k^2 \|g^k\|^2 - \rho_k \|\overline{g}^k\|^2 + \rho_k \overline{g}_k^T (\overline{g}^k - g^k) \\
&\leqslant \overline{f}(x^k, \sigma_k) + \frac{C_1 \cdot K \cdot C_2}{\sigma_k} \cdot \rho_k^2 - \rho_k \|\overline{g}^k\|^2 + \frac{C_1 \cdot K}{\sigma_k} \cdot \rho_k^2 \cdot \big(\|g^k\|^2 - E\|g^k\|^2\big) \\
&\quad + \rho_k (\overline{g}^k)^T (\overline{g}^k - g^k),
\end{aligned} \tag{7}$$

$0 \leqslant \tau \leqslant 1$ ir trumpumo dėlei pažymėta $\overline{g}^k = \overline{g}(x^k, \sigma_k)$, $g^k = g(x^k, \sigma_k, \xi_{k+1})$. Tarkime priešingai. Tada galima rasti tokius skaičius $\overline{s}$ ir δ, kad aibių $\{x \,|\, \|x^s - x\| \leqslant 2\delta\}$ seka neturi sankirtos su X^*, kai $s \geqslant \overline{s}$. Iš 2 lemos seka, kad galima rasti tokį $\varepsilon > 0$, kad

$\|\overline{g}^s\| \geqslant \varepsilon > 0$ tolygiai x atžvilgiu, jei $s \geqslant \bar{s}$, ir $\bar{s}$ yra pakankamai didelis. Tegul $\bar{s}$ yra toks, kad yra teisinga nelygybė, kai $i \geqslant \bar{s}$: $\frac{C_1 \cdot K \cdot |\sigma_{i+1} - \sigma_i|}{\rho_i} + \frac{\rho_i \cdot C_2 \cdot C_1 \cdot K}{\sigma_i} \leqslant \frac{\varepsilon^2}{2}$. Tada iš (6) ir (7) gauname:

$$\begin{aligned}
\bar{f}(x^{k+1}, \sigma_{k+1}) &\leqslant \bar{f}(x^s, \sigma_s) - \frac{\varepsilon^2}{2} \sum_{i=s}^{k} \rho_i \\
&\quad - \sum_{i=s}^{k} \rho_i \left(\frac{\varepsilon^2}{2} - \frac{C_1 \cdot K \cdot |\sigma_{i+1} - \sigma_i|}{\rho_i} - \frac{\rho_i C_1 \cdot K \cdot C_2}{\sigma_i} \right) \\
&\quad + \sum_{i=s}^{k} \rho_i^2 \cdot \frac{C_1 \cdot K}{\sigma_i} \cdot \left(\|g^i\|^2 - E\left(\|g^i\|^2 | \Theta_i\right) \right) + \sum_{i=s}^{k} \rho_i (\overline{g}^i)^T (\overline{g}^i - g^i) \\
&\leqslant \bar{f}(x^s, \sigma_s) - \frac{\varepsilon^2}{2} \sum_{i=s}^{k} \rho_i + \sum_{i=s}^{k} \rho_i (\overline{g}^i)^T (\overline{g}^i - g^i) \\
&\quad + \sum_{i=s}^{k} \rho_i^2 \cdot \frac{C_1 \cdot K}{\sigma_i} \cdot \left(\|g^i\|^2 - E\left(\|g^i\|^2 | \Theta_i\right) \right).
\end{aligned} \tag{8}$$

Kadangi $\sum_{i=0}^{k} \rho_i^2 (E((g^i)^T | \Theta_i) \cdot \overline{g}^i)^2 \leqslant \sum_{i=0}^{k} \rho_i^2 (E(\|g^i\|^2 | \Theta_i))^2 \leqslant C_2^2 \cdot \sum_{i=0}^{k} \rho_i^2$ ir $\sum_{k=1}^{\infty} \rho_k^2 < \infty$, iš 3 lemos išplaukia, kad $\sum_{i=0}^{\infty} \rho_i (\overline{g}^i)^T \cdot (\overline{g}^i - g^i)$ b.t. konverguoja. Panašiai įrodome, kad $\sum_{i=0}^{\infty} \rho_i^2 \cdot \frac{C_1 \cdot K}{\sigma_i} \cdot (\|g^i\|^2 - E(\|g^i\|^2 | \Theta_i .))$ b.t. konverguoja, nes egzistuoja tolygiai aprėžti visi įverčio (4) baigtinės eilės momentai. Kadangi $\sum_{i=s}^{\infty} \rho_i = \infty$, o $\bar{f}(x, \sigma)$ yra aprėžta iš apačios, gauname prieštaravimą. Taigi, turi egzistuoti begalinis posekis, konverguojantis į X^*. Dabar tarkime, kad sekoje (5) egzistuoja, arba aprėžtas posekis, konverguojantis į $x' \notin X^*$, arba posekis, konverguojantis į begalybę.

Pirma įrodysime, kad, jei sekoje $\{x^k\}_{k=0}^{\infty}$ yra aprėžtas posekis, konverguojantis į tašką x', kur $\inf_{g \in \partial f(x')} \|g\| > 0$, tai egzistuoja toks δ_0, kad esant bet kokiems $\delta \in (0, \delta_0]$, galima rasti tokias indeksų sekas $\{l_s\}_{s=0}^{\infty}$, $\{k_s\}_{s=0}^{\infty}$, $l_s < k_s$, kad $\|x^i - x'\| \leqslant 2\delta$ dėl visų $i \in [l_s, k_s - 1]$, ir

$$\overline{\lim_{s \to \infty}} f(x^{k_s}) < \lim_{s \to \infty} f(x^{l_s}). \tag{9}$$

Iš tikrųjų, iš prielaidų apie posekių, konverguojančių į $x' \notin X^*$ ir į $x \in X^*$, egzistavimą, išplaukia, kad galime rasti tokias indeksų sekas $l_s, k_s, l_s < k_s$, kad pakankamai mažam δ turime, kad $\|x^{l_s} - x'\| < \delta$, $\|x^{k_s} - x'\| > 2\delta$ ir $\|x^i - x'\| \leqslant 2\delta$, kai $i \in [l_s, k_s - 1]$. Dabar įrodysime, kad iš tokių sekų egzistavimo išplaukia (9). Turime, kad $x^{k_s} = x^{l_s} - \sum_{i=l_s}^{k_s-1} \rho_i g^i = x^{l_s} - \sum_{i=l_s}^{k_s-1} \rho_i \overline{g}^i - \sum_{i=l_s}^{k_s-1} \rho_i (g^i - \overline{g}^i)$. Iš Lemos 3 išplaukia, kad $\sum_{i=0}^{\infty} \rho_i \cdot (g^i - \overline{g}^i)$ b. t. konverguoja. Reiškia $\| \sum_{i=l_s}^{k_s-1} \rho_i (g^i - \overline{g}^i) \| < \frac{\delta}{2}$, kai $s \geqslant \bar{s}$, jei $\bar{s}$ yra pakankamai didelis. Todėl $\delta \leqslant \|x^{k_s} - x^{l_s}\| \leqslant \| \sum_{i=l_s}^{k_s-1} \rho_i \overline{g}^i \| +$

$\| \sum_{i=l_s}^{k_s-1} \rho_i (g^i - \overline{g}^i) \| \leqslant \| \sum_{i=l_s}^{k_s-1} \rho_i \overline{g}^i \| + \frac{\delta}{2}$. Tad $\| \sum_{i=l_s}^{k_s-1} \rho_i \overline{g}^i \| \geqslant \frac{\delta}{2}$ ir $\sum_{i=l_s}^{k_s-1} \rho_i \geqslant \frac{\delta}{2 \cdot \sqrt{C_2}}$, nes $\|\overline{g}^i\|^2 < C_2$. Tuomet iš (8) gauname įvertį: $\overline{f}(x^{k_s}, \sigma_{k_s}) \leqslant \overline{f}(x^{l_s}, \sigma_{l_s}) - \frac{\varepsilon^2 \cdot \delta}{8 \cdot \sqrt{C_2}}$. Iš tolygaus konvergavimo x atžvilgiu: $\overline{f}(x, \sigma) \to f(x)$, kai $\sigma \to 0$, seka (9). Pastebėsime, kad bet kuriems skaičiams f' ir f'', tokiems, kad $\overline{\lim}_{s \to \infty} f(x^{k_s}) < f' < f'' < \lim_{s \to \infty} f(x^{l_s})$, seka $f(x^k)$ kerta intervalą (f', f'') be galo daug kartų. Tuomet galima rasti tokias dvi sekas $\{x^{r_s}\}$, čia $\|x^{r_s} - x'\| \leqslant 2\delta$, ir $\{x^{p_s}\}$, kurioms

$$f(x^{r_s}) \leqslant f', \quad f(x^{r_s+1}) > f', \tag{10}$$

$$f(x^{p_s}) > f'', \quad f(x^i) > f', \quad r_s < i < p_s. \tag{11}$$

Seką $\{x^{r_s}\}$ galima laikyti konverguojančia; priešingu atveju vietoj šios sekos galima paimti jos konverguojantį posekį. Iš sekos $\{x^{r_s}\}$ konvergavimo, funkcijos f tolydumo, (10) ir $\rho_i \to 0$ turime, kad $\lim_{s \to \infty} f(x^{r_s}) = f'$. Tuomet iš (10) išplaukia, kad

$$\overline{\lim_{s \to \infty}} f(x^{p_s}) \geqslant \lim_{s \to \infty} f(x^{r_s}). \tag{12}$$

Kadangi F^* neturi vidinių taškų, tai galime f' parinkti tokį, kad $f' \notin F^*$. Kadangi pagal Lipšico sąlygą galime parinkti $\delta \leqslant \frac{f''-f'}{2 \cdot K}$, $\delta > 0$, tai dabar galime išvesti (9), kai seka $\{r_s\}$ atitinka $\{l_s\}$, o $\{p_s\}$ atitinka $\{k_s\}$. Tačiau pastaruoju atveju (12) prieštarauja (9).

Jei sekoje $\{x^k\}_{k=0}^{\infty}$ yra posekis, konverguojantis į begalybę, panašiai įrodome, kad egzistuoja toks δ_0, kad, kai $\delta \in (0, \delta_0]$, galima rasti tokias indeksų sekas $\{l_s\}_{s=0}^{\infty}$, $\{k_s\}_{s=0}^{\infty}$, kad $\inf_{x \in X^*} \|x^i - x\| \geqslant \delta$, dėl visų $i \in [l_s, k_s]$ ir $\overline{\lim}_{s \to \infty} f(x^{k_s}) < \lim_{s \to \infty} f(x^{l_s})$, toliau analogiškai aukščiau išnagrinėtam atvejui nustatome atitinkamų indeksų r_s, p_s egzistavimą, vedantį į prieštaravimą.

Straipsnio autoriai nuoširdžiai dėkoja recenzentui už pateiktas pastabas, padėjusias žymiai pagerinti straipsnį.

Literatūra

1. V.S. Michalevič, A.M. Gupal, V.I. Norkin, *Metody nevypukloj optimizacii*, Nauka, M. (1987).
2. K. Kuratowski, *Topology*, II, Academic Press, N.Y. (1968).
3. L. Sakalauskas, Apie centruojančių klaidžiojimų konvergavimą, *Liet. matem. rink.*, **34**(1), 120–130 (1994).
4. J.C. Spall, Multivariate stochastic approximation using a simultaneous perturbation gradient approximation, *IEEE Transactions on Automatic Control*, **37**, 332–341 (1992).
5. S. Sternberg, *Lekcii po diferencialnoj geometrii*, Nauka, M. (1970).
6. M.T. Vazan, Stochastic approximation, in: *Transactions in Mathematics and Mathematical Physics*, Cambridge University Press, Cambridge (1969).

SUMMARY

V. Bartkutė, L. Sakalauskas. Simultaneous perturbation stochastic aproximation for Lipshitz functions

The convergence conditions of Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm with Gaussian smoothing operator are established in the class of Lipshitz functions.

Keywords: stochastic approximation, Lipshitz function, differential.