ISSN 1648-8776
JAUNŲJŲ MOKSLININKŲ DARBAI. NR. 2 (40). 2013

REIKŠMINGŲ IMČIŲ METODAS DVIEJŲ ETAPŲ STOCHASTINIAM TIESINIAM UŽDAVINIUI

Valerijonas Dumskis, Leonidas Sakalauskas
Šiaulių universitetas

Įvadas

Stochastiniame programavime, kai atsitiktinių parametrų tikimybiniai pasiskirstymai yra žinomi, paprastai naudojama aproksimacija imtimis. Reikia taip sumodeliuoti visą atsitiktinio parametro erdvę, kuri būtų kiek įmanoma arčiau tikrosios, imant pakankamą imties dydį, kad tikimybinė tikslo funkcijos reikšmė kiekvienoje iteracijoje gerėtų. Tradiciniuose imtimis grįstuose metoduose tai pasiekiama modelio imitavimu su duotu imties dydžiu, po to apskaičiuojant tikimybinės tikslo funkcijos reikšmę (t. y. tikėtiną tikslo funkcijos reikšmę). Du tokie metodai yra: L pavidalo (*L shaped*) dekompozicijos metodas (Shapiro, 2000) ir nuoseklios paieškos Monte Karlo serijų įverčiais metodas (Sakalauskas, 2002, 2004). Esminis šio metodo bruožas – sumažinti scenarijų skaičių, kuris reikalingas gauti tikimybinės tikslo funkcijos statistinių įverčių priimtinus pasikliovimo intervalus.

Dydžio pasikliovimo intervalo ilgis priklauso nuo dispersijos, taigi, norint sumažinti pasikliovimo intervalo ilgį, reikia mažinti dispersiją. Stochastiniame programavime, kai aproksimuojama imtimis, dispersijos sumažinimo priemonė ir kartu scenarijų skaičiaus sumažinimo – reikšmingų imčių metodas. Reikšmingų imčių metodas remiasi tuo, kad tam tikros atsitiktinio kintamojo reikšmės turi didesnę įtaką parametrui, kurį norime įvertinti (Bucklew, 2004; Changhe, Druzdzel, 2006; Kurtz, Song, 2013). Jei šios „reikšmingos“ reikšmės yra dažniau parenkamos imtyje, įverčio dispersija gali būti sumažinta. Vadinasi, pagal reikšmingų imčių metodologiją reikia parinkti tokį pasiskirstymą, kuris „paskatina“ reikšmingas reikšmes. Realizuojant reikšmingų imčių metodą modeliavime, pagrindinė problema yra toks pasiskirstymo, kuris susijęs su duotuoju, parinkimas, kad kintamųjų reikšmės patektų į reikšmingas sritis.

Tyrimo tikslas – patobulinti stochastinio programavimo metodą.

Stochastinio programavimo patobulinimas, naudojant reikšmingų imčių metodą

Tarkime, kad reikia minimizuoti tam tikrą tikimybinę tikslo funkciją $F(z)$ (tikėtinos išlaidos, tikėtinas sutrikimo laikas ir pan.):

$$F(x) \equiv Ef(x,\omega) \to \min_{z \in D \subset \Re^n}, \qquad (1)$$

čia: $\omega \in \Omega$ yra elementarusis įvykis iš tikimybinės erdvės (Ω, Σ, P_x), funkcija $f : \Re^n \times \Omega \to \Re$ tenkina tam tikrus integruotumo, diferencijuotumo ir iškilumo reikalavimus, $D \subset \Re^n$ yra leistinų sprendinių aibė ir P_x yra tikimybinis matas, kuris yra absoliučiai tolydus ir apskritai priklauso nuo x, t. y. jis gali būti apibrėžtas tankio funkcija $p : \Re^n \times \Omega \to \Re_+$, E yra matematinės vilties simbolis.

Esant tokioms prielaidoms tikslo funkcijai, nagrinėjamas uždavinys užrašomas kartotiniu integralu:

$$F(x) \equiv \int_\Omega f(x,s) p(x,s) ds. \qquad (2)$$

Įveskime atsitiktinį vektorių, jį apibrėždami tankio funkcija $\varphi : \Re^m \times \Omega \to \Re_+$, kuri priklauso nuo determinuotų parametrų $y \in \Re^m$ vektoriaus, kurio parinkimą aptarsime žemiau. Dabar (2) tikslo funkcija tampa vidurkiu naujo atsitiktinio vektoriaus atžvilgiu:

$$F(x) \equiv \int_\Omega f(x,s) \cdot q(x,y,s) \cdot \varphi(y,s) ds, \qquad (3)$$

čia $q(x,y,s) = \dfrac{p(x,s)}{\varphi(y,s)}$. Atsitiktinės funkcijos po (3) integralu antrasis momentas yra:

$$D(x,y) = \int_\Omega \left(f(x,s) \cdot q(x,y,s)\right)^2 \cdot \varphi(y,s) ds =$$

$$= \int_\Omega \frac{\left(f(x,s) \cdot p(x,s)\right)^2}{\varphi(y,s)} ds. \qquad (4)$$

Lengva įsitikinti, kad atsitiktinės funkcijos po (3) integralu dispersija $D(x,y) - F(x)^2$ mažėja, jei antrasis momentas (4) mažėja. Vadinasi, keičiant (4)-oje kintamąjį y, galima sumažinti dispersiją, taigi idėja – optimizuoti y parinkimą.

Tuo tikslu sprendžiamas dviejų lygių stochastinis optimizavimo uždavinys:

$$F(x) \equiv \int_\Omega f(x,s) \cdot q(x,y,s) \cdot \varphi(y,s) ds \to$$

$$\to \min_{x \in D \subset \Re^n}, \qquad (5)$$

$$D(x,y) \equiv \int_\Omega \left(f(x,s) \cdot q(x,y,s)\right)^2 \cdot \varphi(y,s) ds \to$$

$$\to \min_y, \qquad (6)$$

čia: pirmasis uždavinys yra lyderio, o antrasis – pasekėjo uždavinys (Christiansen ir kt., 2001).

Nagrinėkime, pavyzdžiui, dviejų etapų stochastinį tiesinį uždavinį:

$$F(x) = c \cdot x + E\{\min_y [q \cdot z \mid W \cdot z + T \cdot x \le h,$$

$$z \in R_+^m]\} \to \min, \qquad (7)$$

čia: matricos W, T, A ir vektoriai c, q, h, b yra atitinkamų matavimų, h vektorius yra atsitiktinis daugiamatis normalusis $N(\mu,\Sigma)$, μ yra vidurkių vektorius, Σ yra kovariacijos matrica. Vektoriaus $N(y,\Sigma)$ tankio funkciją pažymėkime $p(s,y)$. Tegu vidurkių vektorius y tankio funkcijoje $p(s,y)$ yra keičiamas tam, kad sumažintume dispersiją. Pažymėkime

$$f(x,h)=\min_{\substack{W\cdot z+T\cdot x\le h\\ z\in R_+^m}} q\cdot z. \qquad (8)$$

Po nesudėtingų pertvarkymų turime:

$$F(x)\equiv\int_\Omega f(x,h)\cdot q(y,h)\cdot p(h,y)ds, \qquad (9)$$

čia:

$$q(y,h)=\frac{p(h,\mu)}{p(h,y)}=$$

$$=\exp\left(-(\mu-y)^T\cdot\Sigma^{-1}\cdot(\mu+y-2h)\right). \qquad (10)$$

Atitinkamai:

$$D(x,y)=\int_\Omega\frac{\left(f(x,h)\cdot p(h,\mu)\right)^2\cdot}{p(h,y)}dh. \qquad (11)$$

Lengva įsitikinti, kad (4) dispersijos gradientas parametrų vektoriaus y atžvilgiu yra:

$$\frac{\partial D(x,y)}{\partial y}=$$

$$=\int_\Omega(h-y)\cdot\Sigma^{-1}\cdot d(x,y,h)\cdot p(h,y)dh, \qquad (12)$$

$$d(x,y,h)=\left(f(x,h)\cdot p(h,\mu)\right)^2\cdot\frac{dp(h,y)}{dy}. \qquad (13)$$

Pastebėsime, kad (7) tikslo funkcijos gradientas gali būti išreikštas per tam tikros vektorinės funkcijos matematinę viltį. Iš tikrųjų, pagal tikslo funkcijos dualumą ji gali būti išreikšta taip:

$$F(x)=c\cdot x+E\left\{\max_u[(h-T\cdot x)\cdot u\,|u\cdot W^T+q\ge 0,\right.$$
$$\left. u\in R_+^s]\right\}.$$

Pagal matematinės vilties diferencijavimo taisykles ir maksimumo savybes (Rubinstein, Melamed, 1998; Sakalauskas, 2002, 2004) galime įsitikinti, kad tikslo funkcijos gradientas yra lygus matematinei vilčiai:

$$\nabla_x F(x)=Eg(x,h), \qquad (14)$$

čia: $g(x,h)=c-T\cdot u^*$, (15)

o u^* yra dualaus tiesinio uždavinio:

$$(h-T\cdot x)^T\cdot u^*=$$

$$=\max_u[(h-T\cdot x)^T\cdot u\,|\,u\cdot W^T+q\ge 0,$$
$$u\in R_+^s]$$ sprendinys

Monte Karlo reikšmingų imčių metodas stochastiniam programavimui

Apibrėžkime Monte Karlo įverčius ir iteracinę procedūrą (7) stochastinio programavimo uždavinio sprendimui reikšmingų imčių metodu. Tarkime, kad bet kuriam leistinam sprendiniui $z\in D\subset R^n$ galima sugeneruoti tam tikro dydžio N Monte Karlo imtį:

$$H=(h^1,h^2,...,h^N), \qquad (16)$$

čia h^j, $1\le j\le N$ yra nepriklausomi vienodai pasiskirstę atsitiktiniai vektoriai, kurių tankis yra $p(h,\cdot y)$. Įverčiai, naudojant šią imtį, yra apskaičiuojami pagal (8), (13) ir (15):

$$\widetilde{F}(x,y)=\frac{1}{N}\sum_{j=1}^N\eta^j, \qquad (17)$$

$$\widetilde{G}(x,y)=\frac{1}{N}\sum_{j=1}^N g^j, \qquad (18)$$

$$\widetilde{D}(x,y)=\frac{1}{N}\sum_{j=1}^N(h^j-y)\cdot\Sigma^{-1}\cdot d^j, \qquad (19)$$

čia pažymėta: $\eta^j=f(x,h^j)$, $g^j=g(x,y,h^j)$, $d^j=d(x,y,h^j)$, $1\le j\le N$.

Tarkime, pradinė aproksimacija $x^0\in D\subset\Re^n$, $y^0=\mu$, pradinis imties dydis N^0 yra duoti ir atsitiktinė seka $\left\{x^t,y^t,N^t\right\}_{t=0}^\infty$ yra apibrėžta taip:

$$x^{t+1}=x^t-\rho\cdot\widetilde{G}_\varepsilon(x^t,y^t), \qquad (20)$$

$$y^{t+1}=y^t-\alpha\cdot\widetilde{D}(x^t,y^t), \qquad (21)$$

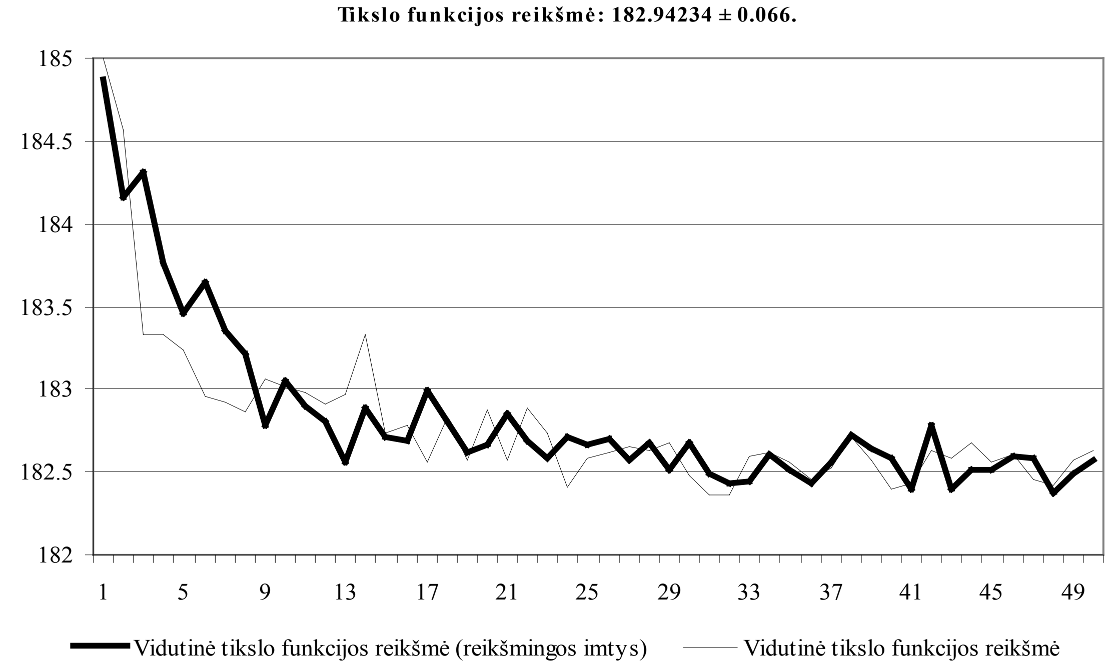
$$N^{t+1}=\frac{1}{b^t}\cdot N^t, \qquad (22)$$

čia: $\widetilde{G}_\varepsilon(x^t,y^t)$ yra stochastinio gradiento ε-projekcija į leistiną aibę (žr. apibrėžimą – Sakalauskas, Zilinskas), $\rho>0,\alpha>0,b^t>0$ yra tam tikri metodo parametrai.

Galima įrodyti, kad (20)–(22) procedūra konverguoja beveik tikrai (b. t.) į (5), (6) uždavinio sprendinį, atitinkamai (determinuotai ar stochastiškai) parinkus metodo parametrų reikšmes (žr. Sakalauskas, 2002, 2004). Pastebėsime, kad Monte Karlo imties dydžio pagal (22) parinkimas leidžia išspręsti uždavinį, atliekant priimtiną skaičiavimų kiekį ir, pasiremiant apibrėžtų Monte Karlo įverčių asimptotiniu Gauso konvergavimu, sukonstruoti statistines algoritmo stabdymo taisykles (Sakalauskas, 2002; Sakalauskas, Zilinskas, 2008). Iš tikrųjų, iteracinis procesas gali būti stabdomas, patikrinus statistinę hipotezę apie tikslo funkcijos gradiento lygybę nuliui, kai tikslo funkcijos įverčio pasikliovimo intervalas ilgis tampa mažesniu už reikiamą tikslumą.

Kompiuterinis modeliavimas

Suformuluotas reikšmingų imčių metodas buvo ištirtas kompiuteriniame modeliavime, sprendžiant dviejų etapų stochastinį tiesinį optimizavimo uždavinį, kuris

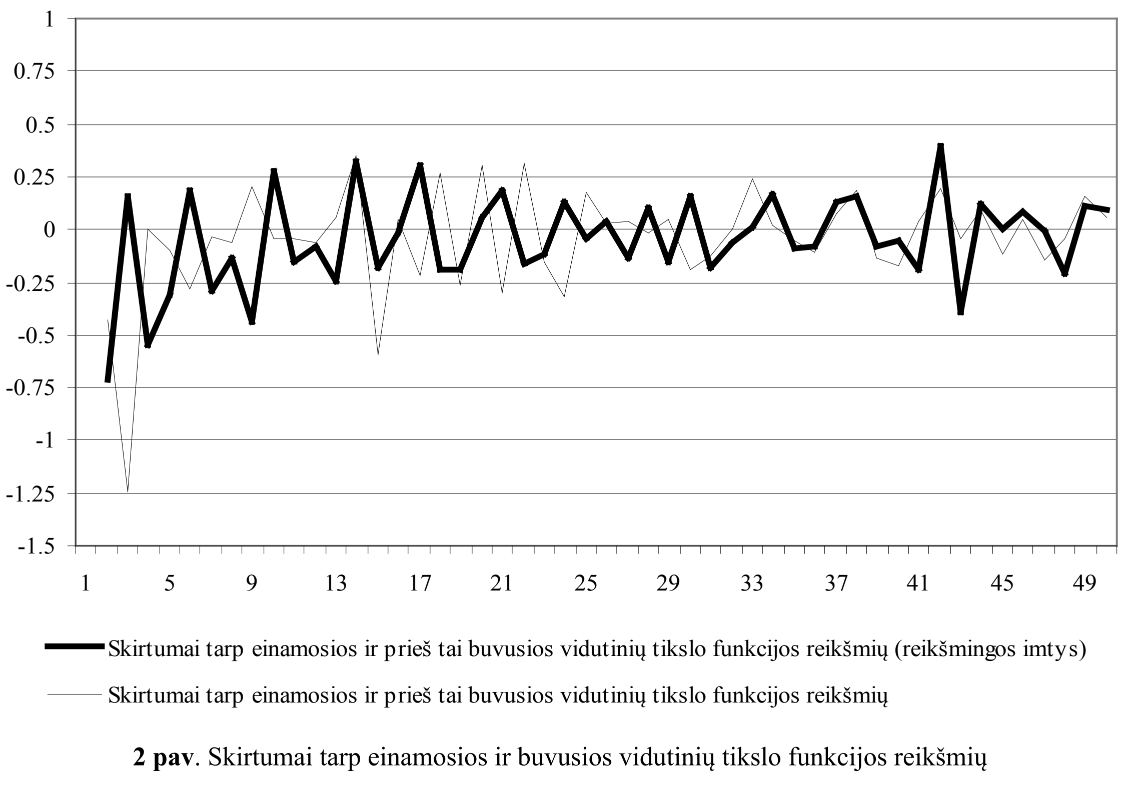


1 pav. Vidutinės tikslo funkcijos reikšmės

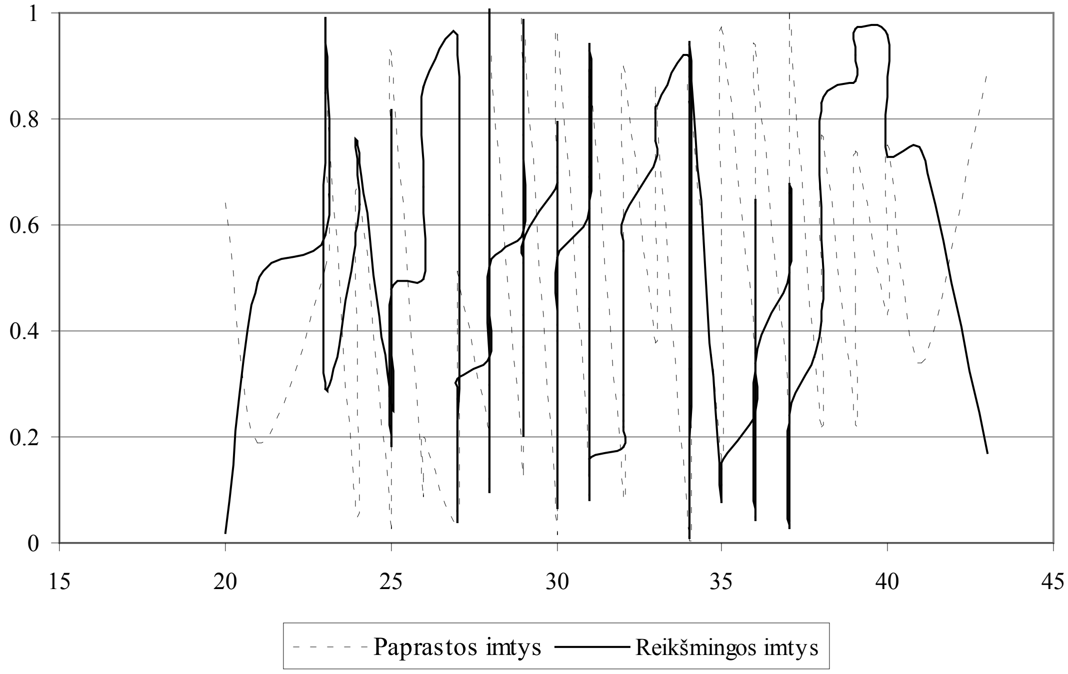
yra paimtas iš duomenų bazės http://www.math.bme.hu/~deak/twostage/ l1/20x20.1 (taip pat žr. Sakalauskas, Zilinskas, 2008).

Uždavinio matavimai yra tokie: pirmame etape yra 10 eilučių ir 20 kintamųjų, antrame etape – 20 eilučių ir 30 kintamųjų. Tikslo funkcijos įvertis, kuris yra pateiktas duomenų bazėje, yra 182,94234 ± 0,066.

Pavyzdys aprašytu metodu buvo spręstas 50 kartų su tokiomis parametrų reikšmėmis: $\rho = 0.0005, \alpha = 0.1$. b^t buvo parinktas pagal (Sakalauskas, Zilinskas, 2008) aprašytą būdą. Iteracijų skaičius buvo fiksuojamas, kai stabdymo sąlygos buvo pirmą kartą patenkinamos (statistinė hipotezė apie tikslo funkcijos gradiento lygybę nuliui nėra atmetama ir tikslo funkcijos įverčio pasikliovimo intervalo ilgis neviršija priimtinos reikšmės $\varepsilon = 2$). 1 pav. pavaizduotos vidutinės tikslo funkcijos reikšmės, taikant reikšmingų imčių metodą, 2 pav. – skirtumai tarp einamosios ir prieš tai buvusios vidutinės tikslo funkcijos reikšmių, 3 pav. – iteracijų skaičiaus, kai patenkintos algoritmo stabdymo sąlygos, dažnis.



2 pav. Skirtumai tarp einamosios ir buvusios vidutinių tikslo funkcijos reikšmių



3 pav. Iteracijų skaičiaus dažnis

Išvada

Tyrimo rezultatai leidžia teigti, kad reikšmingų imčių metodas leidžia sumažinti iteracijų skaičių, kuris reikalingas, kad būtų patenkintos stabdymo sąlygos, taip pat sumažinti Monte Karlo imties dydį kiekvienoje iteracijoje.

Literatūra

1. Bucklew J. A., 2004, *Introduction to rare event simulation*. Springer-Verlag, New York.
2. Changhe Yuan C., Druzdzel M. J., 2006, Importance sampling algorithms for Bayesian networks: Principles and performance. *Mathematical and Computer Modelling*. Vol. 43. No 9–10. P. 1189–1207.
3. Christiansen S., Patriksson M., Wynter L., 2001, Stochastic bilevel programming in structural optimization. *Structural and Multidisciplinary Optimization*. Vol. 21. No 5. P. 361–371
4. Kurtz N., Song J., 2013, Cross-entropy-based adaptive imp[ortance sampling using Gaussian mixture. *Structural Safety*. Vol. 42. P. 35–44.
5. Rubinstein R. Y., Melamed B., 1998, Modern Simulation and Modeling. *Wiley Series in Probability and Statistics: Applied Probability and Statistics*. J. Wiley&Sons, N. Y.
6. Sakalauskas L., 2002, Nonlinear Stochastic Programming by Monte-Carlo estimators. *European Journal on Operational Research* . Vol. 137. P. 558–573.
7. Sakalauskas L., 2004, Nonlinear stochastic optimization by Monte-Carlo estimators. *Informatica*. Vol. 15 (2). P. 271–282.
8. Sakalauskas L., Žilinskas K., 2008, Epsilon-projection method for two-stage SLP. *Lietuvos matematikos rinkinys*. T. 48/49, spec. Nr. P. 320–326.
9. Shapiro A., 2000, Stochastic Programming by Monte-Carlo simulation methods. *Stochastic Programming E-Print Series*.

IMPORTANCE SAMPLING METHOD FOR TWO-STAGE LINEAR STOCHASTIC PROBLEM

Valerijonas Dumskis, Leonidas Sakalauskas

Summary

In this paper, a two-stage linear stochastic problem is investigated, and solved by applying the method of approximation by sampling. When we approximate by samples, importance sampling is used, i.e., we choose samples for which values of variables lie in important regions. To solve the problem, we proposed a procedure in which the variance of sample and the ε - projection of stochastic gradient is used. This procedure converges almost certainly to the true solution of the problem. In this procedure, the size of the sample is regulated. The conditions for termination of the algorithm are also stated. The iterative process can be stopped by testing the statistical hypothesis of the equality of the gradient of the objective function to zero, when the confidence interval of the estimator of the objective function becomes of the proper accuracy. We then use importance sampling conditions for the stopping algorithm which are satisfied at the iterations whose number is less than using simple samples. The proposed method was investigated by solving the two-stage linear stochastic problem. Thus, the results obtained enable us to conclude that importance sampling enables us to decrease the number of iterations needed to achieve the termination conditions as well as to decrease the Monte-Carlo sample size at each iteration.

REIKŠMINGŲ IMČIŲ METODAS DVIEJŲ ETAPŲ STOCHASTINIAM TIESINIAM UŽDAVINIUI

Valerijonas Dumskis, Leonidas Sakalauskas

Santrauka

Nagrinėjamas dviejų etapų tiesinis stochastinis programavimo uždavinys, kurį sprendžiant naudojamas aproksimacijos imčių metodas. Aproksimuojant imtimis parenkamos reikšmingos imtys, t. y. imtys, prie kurių kintamųjų reikšmės patenka į reikšmingas sritis. Pateikėme uždavinio sprendimo procedūrą, kurioje naudojama imties dispersija ir stochastinio gradiento ε - projekcija. Ši procedūra konverguoja beveik tikrai į uždavinio sprendinį. Šioje procedūroje yra reguliuojamas imties dydis. Taip pat yra suformuluotos algoritmo stabdymo taisyklės.

Iteracinis procesas gali būti stabdomas, patikrinus hipotezę apie tikslo funkcijos gradientų lygybę nuliui, kai tikslo funkcijos įverčio pasikliovimo intervalo ilgis tampa mažesniu už reikiamą tikslumą. Kai taikome reikšmingų imčių metodą, algoritmo stabdymo sąlygos yra patenkintos iteracijoje, kurios numeris yra mažesnis, nei taikant paprastas imtis.

Pasiūlytas metodas buvo ištirtas, sprendžiant dviejų etapų tiesinį stochastinį uždavinį. Taigi, gauti rezultatai leidžia teigti, kad reikšmingų imčių metodas sumažina iteracijų skaičių, prie kurio yra patenkintos algoritmo stabdymo sąlygos ir taip pat sumažina Monte Karlo imties dydį kiekvienoje iteracijoje.

Prasminiai žodžiai: dviejų etapų tiesinis stochastinis programavimo uždavinys, Monte Karlo metodas, reikšmingos imtys, konvergavimas.

Įteikta 2013-06-17