VILNIAUS UNIVERSITETAS

Gražina Gimbutienė

STATISTINIAIS IR LIPŠICO TIKSLO FUNKCIJOS MODELIAIS PAGRĮSTI NEIŠKILOS GLOBALIOSIOS OPTIMIZACIJOS ALGORITMAI

Daktaro disertacijos santrauka
Fiziniai mokslai, informatika (09P)

Vilnius, 2017

Disertacija rengta 2013–2017 metais Vilniaus universitete.

Mokslinis vadovas:

prof. habil. dr. Antanas Žilinskas (Vilniaus universitetas, fiziniai mokslai, informatika - 09P).

Disertacija ginama viešame Gynimo tarybos posėdyje:

Pirmininkas
prof. habil. dr. Gintautas Dzemyda (Vilniaus universitetas, fiziniai mokslai, informatika - 09P).
Nariai:
prof. dr. Kęstutis Dučinskas (Klaipėdos universitetas, fiziniai mokslai, informatika - 09P),
prof. dr. Audronė Jakaitienė (Vilniaus universitetas, fiziniai mokslai, informatika - 09P),
dr. Dmitri E. Kvasov (Kalabrijos universitetas, Italija, fiziniai mokslai, informatika - 09P),
dr. Remigijus Paulavičius (Vilniaus universitetas, fiziniai mokslai, informatika - 09P).

Disertacija bus ginama viešame Gynimo tarybos posėdyje 2017 m. rugsėjo mėn. 15 d. 12 val. Vilniaus universiteto Matematikos ir informatikos instituto 203 auditorijoje.
Adresas: Akademijos g. 4, LT–08663 Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2017 m. rugpjūčio mėn. 14 d.
Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu: *www.vu.lt/lt/naujienos/ivykiu-kalendorius*.

VILNIUS UNIVERSITY

Gražina Gimbutienė

ALGORITHMS FOR NON-CONVEX GLOBAL OPTIMIZATION BASED ON THE STATISTICAL AND LIPSCHITZ OBJECTIVE FUNCTION MODELS

Summary of doctoral dissertation
Physical sciences, informatics (09P)

Vilnius, 2017

The dissertation work was carried out at Vilnius University from 2013 to 2017.

Scientific supervisor:

Prof. Dr. Habil. Antanas Žilinskas (Vilnius University, physical sciences, informatics - 09P).

The defense council:

Chairman
Prof. Dr. Habil. Gintautas Dzemyda (Vilnius University, Physical Sciences, Informatics - 09P).
Members:
Prof. Dr. Kęstutis Dučinskas (Klaipėda University, Physical Sciences, Informatics - 09P),
Prof. Dr. Audronė Jakaitienė (Vilnius University, Physical Sciences, Informatics - 09P),
Dr. Dmitri E. Kvasov (University of Calabria, Italy, Physical Sciences, Informatics - 09P),
Dr. Remigijus Paulavičius (Vilnius University, Physical Sciences, Informatics - 09P).

The dissertation will be defended at the public meeting of the council in the auditorium 203 of the Vilnius University Institute of Mathematics and Informatics on the 15th of September, 2017 at 12:00.
Address: Akademijos str. 4, LT–08663 Vilnius, Lithuania.

The summary of the dissertation was distributed on the 14th of August, 2017.
The dissertation is available at the library of Vilnius University.

Turinys

1. Įvadas

1.1. Tyrimų sritis

Sprendžiant taikomuosius uždavinius dažnai iškyla poreikis surasti kintamųjų kombinaciją, atitinkančią geriausią kiekybiškai išreikšto kriterijaus reikšmę, pvz., mažiausią produkto kainą. Atvejis, kai ieškomas visoje kriterijaus apibrėžimo srityje geriausias sprendinys, žinomas kaip globalaus optimizavimo uždavinys, šiame darbe formuluojamas taip:

$$\min_{\mathbf{x}\in A} f(\mathbf{x}), \tag{1}$$

kur $f(\mathbf{x})$ yra neiškila tikslo funkcija, o $A = \{\mathbf{x} \in \mathbb{R}^d : a_i \leq x_i \leq b_i, i = 1, \ldots, d\}$ yra apibrėžimo sritis. (24) uždavinys dažniausiai sprendžiamas skaitiniais algoritmais, nes analitinis sprendimas žinomas tik išskirtiniais atvejais.

Dažnas atvejis, kai tinkamos kintamųjų kombinacijos pasirinkimą nulemia ne vienas, o keletas tarpusavyje prieštaringų kriterijų $f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})$, $k \geq 2$, $\mathbf{x} \in A$, , pvz., produkto kaina ir kokybė. Tokia situacija formalizuojama daugiakriterinio optimizavimo uždaviniu:

$$\min_{\mathbf{x}\in A}(f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})). \tag{2}$$

(25) uždavinio sprendiniai sudaro aibę, tenkinančią Pareto optimalumo apibrėžimą, t. y. aibėje nėra dviejų tokių sprendinių, kad vienas iš jų būtų ne prastesnis už kitą pagal visus kriterijus ir griežtai geresnis bent pagal vieną. Pareto aibės aproksimavimas yra sudėtingas uždavinys, dažniausiai sprendžiamas skaitiniais metodais.

Šiame darbe nagrinėjami statistiniais ir Lipšico tikslo funkcijos modeliais pagrįsti optimizacijos algoritmai, skirti (24) ir (25) uždaviniams spręsti gana tipiškoje *juodosios dėžės* situacijoje. Tokiu atveju tikslo funkcijos reikšmės gali būti apskaičiuojamos apibrėžimo srities taškuose, tačiau funkcijos analitinė išraiška nėra žinoma.

Statistinis tikslo funkcijos modelis naudingas, kai jos reikšmės skaičiavimas ilgai trunka arba brangiai kainuoja. Atlikus keletą tikslo funkcijos įvertinimų, jos statistinis modelis gali būti traktuojamas kaip reikšmės prognozė bei su ja susijęs patikimumo matas neištirtuose apibrėžimo srities taškuose. Remdamasis modeliu, optimizavimo algoritmas įvertina potencialą aptikti naujame taške geresnę tikslo funkcijos reikšmę už jau žinomas.

Kita vertus, Lipšico tikslo funkcijos modeliu pagrįsti optimizavimo algoritmai remiasi gana realistiška prielaida, kad tikslo funkcijos kitimo greitis yra aprėžtas. Šis rėžis, vadinamoji Lipšico konstanta, dažniausiai yra nežinomas, ir tai yra didžiausia šios klasės metodų problema. Pagrindinis prielaidos privalumas — algoritmų teorinio tyrimo paprastumas.

1.2. Darbo aktualumas

Įvairūs statistiniai tikslo funkcijos modeliai, tokie kaip Gausinės stochastinės funkcijos arba apibrėžimo srities dekompozicijai pritaikyti statistiniai modeliai, buvo naudoti globaliosios optimizacijos algoritmuose. Siekiant pasirinkti tinkamą modelį, reikia atsižvelgti į turimą informaciją apie tikslo funkciją bei su modeliu susijusį skaičiavimų sudėtingumą. Dėl to praverstų sistemingas tyrimas, parodantis, kokią įtaką įvairių savybių tikslo funkcijų optimizavimo rezultatams turi skirtingi statistiniai modeliai.

Dėl prielaidos apie aukštą tikslo funkcijos įvertinimų kainą statistiniais tikslo funkcijos modeliais pagrįstuose optimizacijos algoritmuose nauji funkcijos įvertinimo taškai parenkami atidžiai planuojant, t. y. skiriant daug skaičiavimo resursų. Pastebėta, kad globalios paieškos algoritmams prireikia daug funkcijos įvertinimų, kol globaliojo minimumo artinys tampa norimo tikslumo, ypač tais atvejais, kai tikslo funkcija sudėtinga ir turi daug lokaliųjų minimumų. Aukšta kaina skatina didinti algoritmų efektyvumą, mažinant tikslo funkcijos įvertinimų skaičių, reikalingą norimam rezultatui pasiekti, pavyzdžiui, kombinuojant globalią ir lokalią paieškos strategijas.

Idealiu atveju optimizavimo algoritmas turėtų būti efektyvus ir teoriškai pagrįstas. Neseniai pasirodęs statistiniu modeliu pagrįstas globalaus optimizavimo algoritmas su simpleksiniais apibrėžimo srities posričiais veikia iteratyviai išrinkdamas simpleksinius posričius dalijimui, remiantis tam tikru euristiškai apibrėžtu kriterijumi. Intuityvu, kad šis kriterijus atitinka tikimybę simplekso viduje aptikti geresnę funkcijos reikšmę. Matematinė ryšio su pagerinimo tikimybe išraiška suteiktų teorinį pagrindimą vienam iš naujausių nagrinėjamos klasės algoritmų.

Pastaraisiais metais sukurta keletas sėkmingų vieno kriterijaus (vienmačių ir daugiamačių) Lipšico optimizavimo metodų su stačiakampiais apibrėžimo srities posričiais, kuriuose hiperstačiakampiai dalijami į tris lygias dalis. Tai vadinamieji adaptyvūs įstrižaininiai algoritmai. Vėliau šis dalijimo būdas buvo pritaikytas ir dviejų kriterijų dvimačiam optimizavimui. Iš teorinio taško įdomu panagrinėti šio dalijimo būdo optimalumą, pradedant nuo santykinai paprasto vienmačio dvikriterinio uždavinio, ir tokiu būdu šį dalijimo būdą teoriškai pagrįsti. Nors optimalūs algoritmai nėra plačiai taikomi, jie gali būti naudojami palyginimui ir tam tikros jų savybės gali būti būdingos plačiau taikomiems algoritmams.

1.3. Darbo tikslai ir uždaviniai

Šiuo darbu siekiama efektyvios ir teoriškai pagrįstos optimalių sprendimų paieškos. Teoriškai ir eksperimentiškai nagrinėjami statistiniais modeliais pagrįsti vieno kriterijaus optimizavimo algoritmai bei jiems pasiūlomos euristinio pobūdžio modifikacijos. Taip pat atliekamas teorinis optimalumo tyrimas vieno kintamojo dvikriterinėje Lipšico optimizacijoje.

Darbo tikslai:

1. Ištirti naudojamo statistinio modelio įtaką globaliosios optimizacijos algoritmų veikimui.
2. Padidinti statistinio globaliosios optimizacijos algoritmo, naudojančio stačiakampę apibrėžimo srities dekompoziciją, globaliojo minimumo aproksimavimo efektyvumą, matuojamą tikslo funkcijos įvertinimų skaičiumi.
3. Teoriškai pagrįsti simpleksinio globaliosios optimizacijos algoritmo apibrėžimą, remiantis simpleksinio statistinio modelio savybėmis.
4. Ištirti blogiausio atvejo optimalumą vieno kintamojo dvikriterinėje Lipšico optimizacijoje.

Siekiant išvardytų tikslų, suformuluoti šie uždaviniai:

1. Pasiūlyti eksperimentinę metodiką ir suformuluoti rekomendacijas dėl statistinio modelio pasirinkimo realizuojant klasikinius globaliojo optimizavimo algoritmus.
2. Pasiūlyti būdų sukombinuoti neseniai sukurtą statistinį globaliosios paieškos algoritmą su lokaliąja paieška, kad būtų padidintas globalaus minimumo aproksimavimo efektyvumas, matuojamas funkcijos įvertinimų skaičiumi.

3. Ištirti asimptotines simpleksinio statistinio modelio savybes ir panaudoti jas siekiant nustatyti teorinį ryšį tarp dviejų simplekso išrinkimo kriterijų simpleksinėje globaliojoje optimizacijoje: euristiškai apibrėžto ir efektyviai apskaičiuojamo kriterijaus bei teoriškai pagrįsto kriterijaus, susijusio su pagerinimo tikimybe.

4. Ištirti viename žingsnyje blogiausiu atveju optimalų intervalo dalijimą į tris dalis vieno kintamojo dvikriterinėje Lipšico optimizacijoje bei realizuoti atitinkamą optimalų algoritmą.

1.4. Darbo naujumas ir rezultatai

Pasiūlyta eksperimentinė metodika, skirta sistemingai įvertinti pasirinktojo statistinio modelio įtaką skirtingų savybių tikslo funkcijų optimizavimui. Remiantis eksperimentų su vienmačiais ir dvimačiais statistiniais modeliais iš Gausinių stochastinių funkcijų klasės rezultatais, suformuluotos statistinio modelio pasirinkimo rekomendacijos.

Pasiūlytos dvi euristinės neseniai sukurto statistinio globaliosios optimizacijos algoritmo modifikacijos, skirtos pasiekti norimo tikslumo globaliojo minimumo artinį panaudojus mažesnį skaičių tikslo funkcijos įvertinimų. Pradinis globaliosios paieškos algoritmas, paremtas stačiakampe apibrėžimo srities dekompozicija, modifikacijose kombinuojamas su skirtingomis lokaliosios paieškos strategijomis. Pasiūlytosios modifikacijos veikia geriau už pirminį bei kitus į palyginimą įtrauktus šiuolaikinius algoritmus.

Pateiktas teorinis pagrindimas euristiškai apibrėžtam simplekso išrinkimo kriterijui, naudojamam neseniai sukurtame globaliosios optimizacijos algoritme. Teorinis pagrindimas gautas, nustačius ryšį tarp minėto euristinio kriterijaus ir pagerinimo tikimybę atitinkančios išraiškos tuo atveju, kai tikslo funkcija modeliuojama simplekse stacionariu izotropiniu Gausiniu atsitiktiniu lauku.

Ištirtas intervalo dalijimo į tris dalis optimalumas viename žingsnyje blogiausiu atveju Lipšico rėžio Pareto frontui paklaidos atžvilgiu. Parodyta, kad intervalo dalijimas į tris lygias dalis tenkina optimalumo apibrėžimą. Realizuotas atitinkamas optimalus algoritmas ir eksperimentiškai pademonstruotas jo veikimas.

1.5. Ginamieji teiginiai

1. *P-algoritmas*, realizuotas naudojant stacionarią izotropinę Gausinę stochastinę funkciją su eksponentine koreliacija, veikia geriausiai optimizuojant įvairias vienmates ir dvimates tikslo funkcijas.

2. *Maksimalaus tikėtino pagerinimo algoritmas*, realizuotas naudojant stacionarią izotropinę Gausinę stochastinę funkciją su koreliacija, užtikrinančia mažą svyravimų mastą, tinka santykinai paprastoms tikslo funkcijoms optimizuoti.

3. Lokalios ir globalios paieškos strategijų derinimas pasiūlytuose hibridiniuose algoritmuose *GB* ir *Cluster* efektyvesnis sunaudotų tikslo funkcijos įvertinimų skaičiaus atžvilgiu nei pirminis optimizavimo algoritmas *Rect-1*, kurį jie išplečia, kai tikslo funkcijos sudėtingos ir turi daug lokaliųjų minimumų.

4. Euristiškai apibrėžtas simplekso išrinkimo kriterijus yra asimptotiškai ekvivalentus pagerinimo tikimybę atitinkančiai išraiškai.

5. Intervalo dalijimas į tris lygias dalis vieno kintamojo dvikriterinėje Lipšico optimizacijoje yra viename žingsnyje optimalus blogiausiu atveju.

2. Statistiniais ir Lipšico tikslo funkcijos modeliais pagrįsta globalioji optimizacija

Pradžioje aptarsime kertines tolimesniame tekste vartojamas sąvokas. Globaliosios optimizacijos uždavinio formuluotėje (24) $f(\mathbf{x})$ vadinama *tikslo funkcija*, o A — *apibrėžimo sritimi*. *Globalusis minimumas* žymimas $f^* = \min_{\mathbf{x}\in A} f(\mathbf{x})$, o *globaliojo minimumo taškų aibė* — $X^* = \{\mathbf{x}^* : f(\mathbf{x}^*) = f^*\}$. *Globaliosios optimizacijos algoritmas* yra procedūra, generuojanti seką taškų $\mathbf{x}_i, i = 1, \ldots, n$, tokią, kad globaliojo minimumo f^* artinys (*geriausia žinoma reikšmė*) $y_{on} = \min_{i=1,\ldots,n} f(\mathbf{x}_i)$ artėja į f^*, kai $n \to \infty$. Papildomai algoritmas sugeneruoja ir tam tikro *globalaus minimumo taško* $\mathbf{x}^* \in X^*$ artinį. Taškai $(\mathbf{x}_i, y_i), y_i = f(\mathbf{x}_i), i = 1, \ldots, n$, vadinami *tikslo funkcijos įvertinimais*, o $\mathbf{x}_i$, $i = 1, \ldots, n$, — *tikslo funkcijos įvertinimo taškais.*

2.1. Statistiniu modeliu pagrįsta globalioji optimizacija

Kai tikslo funkcijos reikšmių apskaičiavimas yra brangus, verta investuoti į tinkamą šio proceso planavimą. Tokiu atveju praverčia statistinio tikslo funkcijos $f(\mathbf{x})$ modelio $\xi(\mathbf{x}), \mathbf{x} \in A$, naudojimas. Statistinis modelis gali būti interpretuojamas kaip funkcijos aproksimacija, gauta naudojant jau turimus jos įvertinimus $(\mathbf{x}_i, y_i), y_i = \xi(\mathbf{x}_i) = f(\mathbf{x}_i)$, $i = 1, \ldots, n$, ir apimanti statistiniais terminais išreikštą sąlyginę reikšmių prognozę $m(\mathbf{x}|\xi(\mathbf{x}_i) = y_i, i = 1, \ldots, n)$ dar neištirtuose taškuose $\mathbf{x} \in A, \mathbf{x} \neq \mathbf{x}_i, i = 1, \ldots, n$, bei prognozės tikslumo matą $s^2(\mathbf{x}|\xi(\mathbf{x}_i) = y_i, i = 1, \ldots, n)$.

Klasikiniai globaliosios optimizacijos algoritmai, naudojantys statistinį modelį, priskiria kiekybinį įvertį kiekvienam taškui $\mathbf{x} \in A$, parodantį tinkamumą jame skaičiuoti funkcijos reikšmę. Renkantis naują tašką tikslo funkcijai apskaičiuoti, srityje A maksimizuojama šio įverčio reikšmė. Šio tipo algoritmų tyrimai rutuliojosi dviem kryptimis. Pirmoji apima optimalaus Bajeso algoritmo [14] modifikacijas, tarp kurių labiausiai išpopuliarėjo *Maksimalaus tikėtino pagerinimo algoritmas*:

$$\mathbf{x}_{n+1} = \arg\max_{\mathbf{x} \in A} \mathbb{E}(\max(\xi(\mathbf{x}) - y_{on}, 0)|\xi(\mathbf{x}_i) = y_i, i = 1, \ldots, n); \tag{3}$$

čia $\mathbb{E}(\cdot|\cdot)$ žymi sąlyginį vidurkį. Antrąją kryptį sudaro racionalių sprendimų teorijos kontekste išvesto *P-algoritmo* modifikacijos. *P-algoritmas* formuluojamas taip:

$$\mathbf{x}_{n+1} = \arg\max_{\mathbf{x} \in A} \mathbb{P}(\xi(\mathbf{x}) \leq y_{on} - \epsilon_n|\xi(\mathbf{x}_i) = y_i, i = 1, \ldots, n); \tag{4}$$

čia y_{on} yra geriausia surasta reikšmė, o $\epsilon_n > 0$ - parametrų seka.

Bendrosios (3) ir (4) algoritmų išraiškos konkretizuojamos, pasirinkus statistinį modelį $\xi(\mathbf{x})$. Klasikiniai statistiniai modeliai, naudoti globaliosios optimizacijos algoritmuose, yra Gausinės stochastinės funkcijos. Pirmasis literatūroje žinomas šios klasės modelis — vienmatis Wiener procesas [12]. Dėl skaičiavimų sudėtingumo, susijusio su įvairių stochastinių funkcijų naudojimu daugiamačiu atveju, buvo pasiūlyti apibendrintieji statistiniai modeliai [23], leidžiantys modeliuoti $f(\mathbf{x})$ kaip atsitiktinių dydžių šeimą, bendresnę nei stochastinė funkcija. Siekiant tolimesnio skaičiavimų supaprastinimo, remiantis pastarojo tipo modeliais pasiūlyti simpleksinėms ir stačiakampėms apibrėžimo srities dekompozicijoms pritaikyti statistiniai modeliai [24]. Iš esmės tai kiekviename apibrėžimo srities posrityje apibrėžti nepriklausomi statistiniai modeliai, atsižvelgiant į funkcijos reikšmes posričio viršūnėse. Perspektyvu toliau tyrinėti būtent šiuos modelius ir vystyti jiems pritaikytus algoritmus, nes jie palankūs tiek skaičiavimų supaprastinimo, tiek teorinio algoritmų tyrimo atžvilgiu.

2.2. Lipšico modeliu pagrįsta globalioji optimizacija

Sakoma, kad funkcija $f(\mathbf{x})$ tenkina Lipšico sąlygą, jei

$$|f(\mathbf{x}_1) - f(\mathbf{x}_2)| \leq L\|\mathbf{x}_1 - \mathbf{x}_2\|, \forall \mathbf{x}_1, \mathbf{x}_2 \in A; \tag{5}$$

čia L yra Lipšico konstanta, o $\| \cdot \|$ žymi Euklidinę arba kitokią normą. (5) prielaida reiškia, kad funkcijos kitimo greitis yra aprėžtas. Ji supaprastina teorinį algoritmų tyrimą, be to, yra teisinga daugeliui taikomųjų uždavinių. Didžiausia tokio funkcijos modeliavimo problema — nežinoma Lipšico konstanta, kuri algoritmuose sprendžiama naudojant iš anksto fiksuotą konstantos reikšmę, jos adaptyvų įvertį arba leistinų konstantų aibę.

Vieno kintamojo vieno kriterijaus Lipšico optimizacija yra teoriškai nuodugniai ištirta. Ypač daug dėmesio skirta algoritmams, optimaliems blogiausiu atveju. Įdomu, kad žinomas *geriausias įmanomas* algoritmas [7], naudojamas kaip atskaitos taškas algoritmų efektyvumui matuoti. Sukurta nemažai algoritmų, tarp kurių populiariausias *Pijavskij-Shubert* algoritmas [16, 22], taip pat ir tokių, kurių efektyvumas artimas geriausiam įmanomam. Būtų įdomu apibendrinti teorinės analizės elementus didesnio kriterijų skaičiaus atvejui.

Daugiamatis atvejis teoriškai tyrinėtas mažiau, nes daugiamatė analizė gerokai sudėtingesnė. Atitinkami algoritmai arba apibendrina *Pijavskij-Shubert* algoritmą, arba tam tikru būdu interpretuoja šakų ir rėžių algoritminę schemą, naudodami tiek stačiakampius, tiek simpleksinius apibrėžimo srities posričius. Vienas sėkmingiausių daugiamačių Lipšico metodų yra *DIRECT* algoritmas [11], kurio vėlesni patobulinimai gali būti sėkmingai integruojami ir statistiniais modeliais grįstuose globaliosios optimizacijos algoritmuose.

3. Statistinio modelio pasirinkimas

Siekiant ištirti, kaip statistinio modelio pasirinkimas veikia globaliosios optimizacijos rezultatus, pasiūlyta eksperimentinė metodika ir atliktas atitinkamas sistemingas tyrimas su keletu vienmačių ir dvimačių Gausinių stochastinių funkcijų (žr. 1 ir 2 lenteles) bei dviem klasikiniais globaliosios optimizacijos algoritmais (3) ir (4). Algoritmai realizuoti naudojant kiekvieną iš pateiktų modelių, taigi gauta $2 \times (3 + 6) = 18$ algoritmų versijų. Sugeneruota po 1000 kiekvieno modelio realizacijų ir jos naudotos kaip tikslo funkcijos. Optimizavimo eksperimentai atlikti su visomis įmanomomis algoritmų ir tikslo funkcijų kombinacijomis, vienmačiu atveju sustojant po $N = 35$, o dvimačiu — po $N = 50$ funkcijos įvertinimų.

Gautus rezultatus apibendrina 1 paveikslas. Kiekviena laužtė atitinka algoritmo realizaciją su pasirinktu statistiniu modeliu, nurodytu legendoje. X ašyje pažymėti modeliai, pagal kuriuos generuotos tikslo funkcijos. Kiekvienas laužtės taškas rodo, koks procentas iš 1000 optimizacijos eksperimentų pasirinkto ir tikrojo modelio kombinacijai baigėsi nesėkme, t. y. globalusis minimumas nebuvo surastas. Kuo kreivė žemesnė, tuo algoritmo versijos rezultatai geresni.

Kaip nurodyta 1 ir 2 lentelėse, modelių *Wiener*, *Exponential* ir *Spheric* realizacijoms būdingas didelis svyravimų mastas, dėl to jie laikomi sudėtingais, o likusieji — paprastesniais. Remiantis gautais rezultatais, 3 lentelėje pateikta suformuluotų algoritmo ir modelio pasirinkimo rekomendacijų santrauka, atsižvelgiant į numatomą tikslo funkcijos sudėtingumą. Rezultatai rodo, kad *P-algoritmo*, realizuoto naudojant *Exponential* modelį, veikimas geriausias optimizuojant įvairias vienmates ir dvimates tikslo funkcijas. *Maksimalaus tikėtino pagerinimo algoritmas*, realizuotas naudojant mažą svyravimų mastą užtikrinančius modelius, tinka santykinai paprastoms tikslo funkcijoms optimizuoti.

4. Statistiniu modeliu grindžiama globalioji optimizacija su stačiakampiais posričiais

Statistiniais modeliais grindžiama globalioji optimizacija pirmiausia skirta brangioms tikslo funkcijoms optimizuoti, dėl to verta ieškoti būdų padidinti globaliojo minimumo aproksimavimo efektyvumą, išreiškiamą tikslo funkcijos įvertinimų skaičiumi. Tęsiant

1 lentelė. Vienmačiai atsitiktiniai procesai.

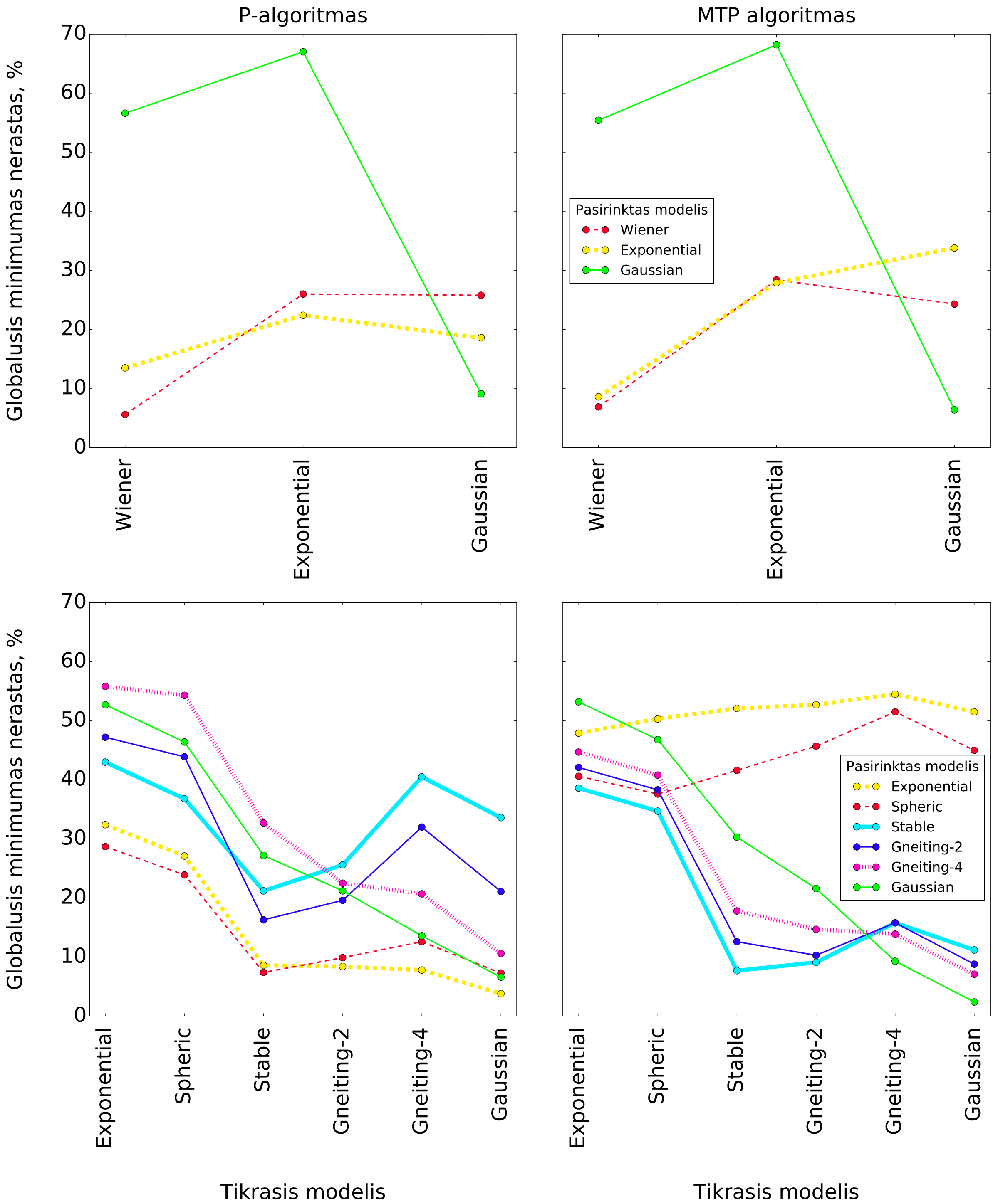
Procesas		
Wiener	Aprašymas:	Gausinis procesas su vidurkiu 0 ir kovariacine funkcija $cov(\xi(x_1), \xi(x_2)) = \sigma^2 \min(x_1, x_2), x_1, x_2 \geq 0$; Gausiniai prieaugiai $\xi(x+\delta) - \xi(x) \sim N(0, \sigma^2\delta)$ nepriklausomi nesikertantiems intervalams.
	Parametrai:	$\sigma > 0$
	Svyravimų mastas:	Didelis.
Exponential	Aprašymas:	Stacionarus Gausinis procesas su eksponentine koreliacija $\rho(\tau) = \exp\left(-\frac{\tau}{c}\right), \tau \geq 0$.
	Parametrai:	μ, $\sigma > 0$, $c > 0$
	Svyravimų mastas:	Didelis.
Gaussian	Aprašymas:	Stacionarus Gausinis procesas su Gausine koreliacija $\rho(\tau) = \exp\left(-\left(\frac{\tau}{c}\right)^2\right), \tau \geq 0$.
	Parametrai:	μ, $\sigma > 0$, $c > 0$
	Svyravimų mastas:	Labai mažas.

2 lentelė. Dvimačiai stacionarūs izotropiniai Gausiniai atsitiktiniai laukai su parametrais μ, $\sigma > 0$, $c > 0$.

Laukas	**Koreliacinė funkcija** $\rho(\tau), \tau \geq 0$	**Svyravimų mastas**
Exponential	$\exp\left(-\frac{\tau}{c}\right)$	Didelis
Spheric	$1 - 1.5\frac{\tau}{c} + 0.5\left(\frac{\tau}{c}\right)^3 \mathbb{1}_{[0,1]}\left(\frac{\tau}{c}\right)$	Didelis
Stable	$\exp\left(-\left(\frac{\tau}{c}\right)^\alpha\right)$, $\alpha = 1.8$	Mažas
Gneiting-2	$(1 + \beta\frac{\tau}{c})(1 - \frac{\tau}{c})^\beta \mathbb{1}_{[0,1]}(\frac{\tau}{c})$, $\beta = \gamma + 2\kappa + 0.5$, $\gamma = 1.5$, $\kappa = 1$	Mažas
Gneiting-4	$\left(1 + \beta\frac{\tau}{c} + \frac{\beta^2-1}{3}(\frac{\tau}{c})^2\right)(1 - \frac{\tau}{c})^\beta \mathbb{1}_{[0,1]}(\frac{\tau}{c})$, $\beta = \gamma + 2\kappa + 0.5$, $\gamma = 1.5$, $\kappa = 2$	Mažas
Gaussian	$\exp\left(-\left(\frac{\tau}{c}\right)^2\right)$	Labai mažas

3 lentelė. Modelio pasirinkimo rekomendacijų santrauka. Sutrumpinimai: *P* — *P-algoritmas*; *MTP* — *Maksimalaus tikėtino pagerinimo algoritmas*.

Dimensija	**Tikslo funkcijos sudėtingumas**	**Rekomendacijos**
1	Nežinomas	*P/MTP + Wiener/Exponential*
	Didelis	*P/MTP + Wiener/Exponential*
	Mažas	*P/MTP + Gaussian*
2	Nežinomas	*P + Exponential/Spheric* *MTP + Stable*
	Didelis	*P+ Exponential/Spheric*
	Mažas	*P/MTP + Gaussian* *MTP + Gneiting-2/Gneiting-4*



1 pav. Procentas atvejų, kai globalusis minimumas nerastas. **Viršuje:** Rezultatai su 1-mačiais modeliais, $N = 35$. **Apačioje:** Rezultatai su 2-mačiais modeliais, $N = 50$.

seriją darbų, nagrinėjančių statistinės globaliosios optimizacijos algoritmų konvergavimą [3–6], neseniai sukurtas statistinis globaliosios optimizacijos algoritmas [2], paremtas stačiakampe leistinosios srities dekompozicija (jam priskiriame pavadinimą *Rect*). Pastebėta, kad algoritmas skiria per daug dėmesio lokaliųjų minimumų aplinkos tyrinėjimui ir dėl to sulėtėja globaliojo minimumo paieška. Disertacijoje pasiūlyti du šią problemą sprendžiantys algoritmai, kombinuojantys globalios paieškos algoritmą *Rect* su lokaliąja paieška.

Algoritmas *Rect* dekomponuoja apibrėžimo sritį į hiperstačiakampių aibę, kurią kiekviename žingsnyje detalizuoja. Tai daroma apskaičiuojant statistinio kriterijaus reikšmę kiekvienam hiperstačiakampiui, išrenkant vieną, kuriam ši reikšmė didžiausia, bei pakeičiant jį dviem naujais, gautais padalijus pradinį hiperstačiakampį į dvi lygias dalis. Disertacijoje nagrinėti keli išrinkimo kriterijaus skaičiavimo variantai, tačiau pagrindu modifikacijoms pasirinkta algoritmo versija *Rect-1*, naudojanti šį kriterijų:

$$\hat{\eta}(R) = \frac{V(R)}{(L(\mathbf{c}) - y_{on} + g(v))^{d/2}}; \tag{6}$$

čia $\mathbf{c}$ — hiperstačiakampio centras, $V(R)$ — hiperstačiakampio tūris, $L(\mathbf{c})$ funkcijos reikšmių hiperstačiakampio viršūnėse vidurkis, y_{on} — geriausia žinoma funkcijos reikšmė, v — mažiausio hiperstačiakampio dekompozicijoje tūris, $g(v)$ — mažėjanti funkcija.

4.1. Modifikacija I: algoritmas *GB*

Pirmos pasiūlytos algoritminės modifikacijos *GB* pseudokodas pateiktas 1 algoritme. Pagrindinė idėja panaši į Lipšico optimizacijai pasiūlytąsias darbuose [15, 17]. Jos esmė — dviejų algoritmo fazių apibrėžimas: standartinės ir globalios, tarp kurių persijungiama atsižvelgiant į konkrečioje fazėje pasiektą progresą. Fazės pradžioje geriausia žinoma reikšmė y_{on} išsaugoma kintamajame s_{best}, o vėlesnis progresas įvertinamas jos atžvilgiu pagal sąlygą:

$$y_{on} \le s_{best} - 0.01|s_{best}|. \tag{7}$$

Hiperstačiakampė apibrėžimo srities dekompozicija žymima raide D, o jos poaibis, iš kurio renkami hiperstačiakampiai dalijimui, žymimas $\bar{D}$. Globalioje fazėje aibei $\bar{D}$ priklauso dalis didžiausių hiperstačiakampių, taip ribojant mažesniųjų dalijimą, o standartinėje fazėje nagrinėjami ir mažesnieji hiperstačiakampiai. Hiperstačiakampių išrinkimo ir dalijimo būdas sutampa su naudojamu *Rect-1* algoritme, tačiau kai patenkinta (7) sąlyga, padalijama 2^d geriausiam taškui artimiausių hiperstačiakampių.

1 algoritmas. *GB* algoritmo pseudokodas.

Inicializacija.
$phase \leftarrow STANDARD$.
while nepatenkinta sustojimo sąlyga **do**
    Išrinkti ir padalinti hiperstačiakampį(-ius).
    if $phase = STANDARD$ **then** ▷ $STANDARD$ fazės pradžia.
        if (7) sąlyga patenkinta **then**
            Užfiksuoti fazės pradžios rekordinę reikšmę: $s_{best} \leftarrow y_{on}$.
        else if $v < v_{small}$ **then**
            $phase \leftarrow GLOBAL$.
            Užfiksuoti fazės pradžios rekordinę reikšmę: $s_{best} \leftarrow y_{on}$.
            $\bar{D} \leftarrow \{R \in D : V(R) \geq v_{large}\}$.
            Inicializuoti $GLOBAL$ fazės iteracijų skaitliuką: $i_{glob} \leftarrow 0$.
        end if
    else ▷ $GLOBAL$ fazės pradžia.
        if (7) sąlyga patenkinta **then**
            $phase \leftarrow STANDARD$.
            Užfiksuoti fazės pradžios rekordinę reikšmę: $s_{best} \leftarrow y_{on}$.
            $\bar{D} \leftarrow \{R \in D : V(R) \geq v_{best}\}$.
        else
            $i_{glob} \leftarrow i_{glob} + 1$.
            if $i_{glob} \bmod i_{period} = 0$ **then** ▷ Atlikti vieną $SECURITY$ iteraciją.
                $\bar{D} \leftarrow \{R \in D : V(R) \geq v_{best}\}$.
                Išrinkti ir padalinti hiperstačiakampį(-ius).
                if (7) sąlyga patenkinta **then**
                    $phase \leftarrow STANDARD$.
                    Užfiksuoti fazės pradžios rekordinę reikšmę: $s_{best} \leftarrow y_{on}$.
                    $\bar{D} \leftarrow \{R \in D : V(R) \geq v_{best}\}$.
                else
                    $\bar{D} \leftarrow \{R \in D : V(R) \geq v_{large}\}$.
                end if
            end if
        end if
    end if
end while

4.2. Modifikacija II: algoritmas *Cluster*

Antros algoritminės modifikacijos *Cluster* pseudokodas pateiktas 2 algoritme. Algoritmu *Rect-1* atliekama globalioji paieška, kol sugeneruojamas ganėtinai mažas hiperstačiakampis. Tuomet hierarchiniu klasterizavimo algoritmu sudaromi hiperstačiakampių klasteriai, iš kurių lokaliosios paieškos algoritmu apdorojamas tas, kuriame aptikta geriausia tikslo funkcijos reikšmė. Ištirtas regionas pašalinimas iš tolimesnės globaliosios paieškos. Procesas kartojamas, kol patenkinama sustojimo sąlyga. Jei sustojimo sąlyga patenkinama anksčiau nei sugeneruojamas ganėtinai mažas hiperstačiakampis, algoritmo *Cluster* veikimas identiškas algoritmo *Rect-1* veikimui.

2 algoritmas. *Cluster* algoritmo pseudokodas.

Apskaičiuoti $f(\mathbf{x})$ reikšmes hiperstačiakampio $A = [0,1]^d$, $n \leftarrow 2^d$, $D \leftarrow \{A\}$ viršūnėse.
while nepatenkinta sustojimo sąlyga **do**
  Nustatyti geriausią hiperstačiakampį: $R^* \leftarrow \max_{R \in D} \hat{\eta}(R)$.
  Pakeisti $R^* \in D$ dviem naujais, gautais padalijus R^*.
  Padidinti n naujų funkcijos įvertinimų skaičiumi.
  if $v < v_{small}(d)$ **then**
    Klasterizuoti dekompozicijai D priklausančius hiperstačiakampius.
    Įvykdyti lokalią paiešką klasteryje, kuriam priklauso geriausia tikslo funkcijos reikšmė.
    Pašalinti dalį hiperstačiakampių iš tolimesnės paieškos.
  end if
end while

4.3. Eksperimentai

Pasiūlytieji algoritmai eksperimentiškai palyginti su pradiniu algoritmu *Rect-1* bei dviem kitais šiuolaikiniais optimizavimo algoritmais: *DIRECT* [11] ir *DIRECTl* [8, 9], siekiant nustatyti jų tinkamumą sudėtingiems globaliosios optimizacijos uždaviniams spręsti. Taikyta eksperimentinė metodika, artima darbams [15, 17]. Naudotos 6 GKLS testinių funkcijų generatoriaus [1, 10] klasės iš [17], turinčios po 100 funkcijų. Algoritmai buvo stabdomi sugeneravus tašką $\mathbf{x}_i = (x_{i1}, \ldots, x_{id}), i = 1, \ldots, n$, tenkinantį sąlygą

$$|x_{ij} - x_j^*| \le \sqrt[d]{\Delta}|b_j - a_j|, j = 1, \ldots, d; \tag{8}$$

4 lentelė. Maksimalus funkcijos įvertinimų skaičius GKLS klasėms: $\max_{i=1,\ldots,100} n_i$.

Klasė	**Rect-1**	**GB**	**Cluster**	**DIRECT**	**DIRECTl**
1	2121	**449**	463	1159	2318
2	9833	1401	**1130**	3201	3414
3	12595	5626	**5582**	12507	13309
4	46024	8355	**7589**	$> 10^6(4)$	29233
5	131341	41210	**26394**	$> 10^6(4)$	118744
6	416556	47597	**44272**	$> 10^6(4)$	287857

5 lentelė. Vidutinis funkcijos įvertinimų skaičius GKLS klasėms: $\max_{i=1,\ldots,100} n_i$.

Klasė	**Rect-1**	**GB**	**Cluster**	**DIRECT**	**DIRECTl**
1	342.70	218.08	248.71	**198.89**	292.79
2	1485.15	**607.75**	662.33	1063.78	1267.07
3	3958.98	2759.15	3200.74	**1117.70**	1785.73
4	8027.53	**3765.39**	4364.15	>42322.65	4858.93
5	19181.31	13176.30	**10660.82**	>47282.89	18983.55
6	39642.24	18097.49	**16953.56**	>95708.25	68754.02

čia $\mathbf{x}^* = (x_1^*, \ldots, x_d^*)$ — globalaus minimumo taškas, $A = \prod_{i=1}^{d}[a_i, b_i]$ — apibrėžimo sritis, o Δ — tikslumo parametras. Leista atlikti ne daugiau kaip 10^6 tikslo funkcijos įvertinimų.

4 ir 5 lentelėse kiekvienai funkcijų klasei pateiktas funkcijos įvertinimų skaičius, kurio prireikė blogiausiu atveju ir vidutiniškai, kol buvo patenkinta sustojimo sąlyga (8). Rezultatai rodo, kad modifikacijos *GB* ir *Cluster* eikvoja mažiau tikslo funkcijos įvertinimų nei pradinis algoritmas *Rect-1*, sprendžiant sudėtingus optimizavimo uždavinius. Šiai uždavinių klasei *Cluster* algoritmo veikimas blogiausiu atveju yra geriausias.

5. Globaliojo optimizavimo su simpleksiniais posričiais asimptotikos tyrimas

Plėtojant simpleksinio statistinio optimizavimo teoriją, siekiama užtikrinti neseniai sukurto globaliosios optimizacijos algoritmo [5] teorinį pagrįstumą. Minėtas algoritmas naudoja euristinį paprastai apskaičiuojamą simplekso išrinkimo kriterijų:

$$\eta_i = \frac{V(S_i)}{\frac{1}{d+1}\sum_{j=1}^{d+1} z_j - y_{on}}; \tag{9}$$

čia $V(S_i)$ — simplekso S_i tūris, $\mathbf{z} = (z_1, z_2, \ldots, z_{d+1})$ — tikslo funkcijos reikšmės simplekso viršūnėse, o $y_{on} < \min_{j=1,\ldots,n} y_j$ yra siekiama reikšmė $(n+1)$-ame algoritmo žingsnyje. Siekta šį kriterijų teoriškai pagrįsti, susiejant jį su tikimybe, kad simplekso viduje bus rasta geresnė tikslo funkcijos reikšmė už jau žinomas.

Uždavinys formuluotas kaip dviejų kriterijų ekvivalentumo įrodymas tuo atveju, kai nagrinėjamas simpleksas taisyklingas ir jo kraštinės ilgis $\delta \to 0$. Tarkime, kad tikslo funkcijos modelis yra stacionarus izotropinis Gausinis atsitiktinis laukas $\xi(\mathbf{x})$, $\mathbf{x} \in A \subset \mathbb{R}^d$, su vidurkiu μ, dispersija σ^2 ir koreliacine funkcija $\rho(\tau) = \exp(-c\tau^2)$. Be to, lauko realizacijos reikšmės žinomos simplekso viršūnėse $\mathbf{x}_i$, $i = 1, \ldots, d+1$: $\xi(\mathbf{x}_i) = z_i$. Pažymėkime simplekso svorio centrą $\mathbf{a}$. Tuomet pagerinimo tikimybę atitinka šis kriterijus:

$$\tilde{p}(\delta) = \frac{s(\delta)}{m(\delta, \mathbf{z}) - y_{on}}; \tag{10}$$

čia $m(\delta, \mathbf{z})$ ir $s(\delta)$ žymi, atitinkamai, atsitiktinio lauko sąlyginį vidurkį ir dispersiją taške $\mathbf{a}$. Taisyklingajam simpleksui apskaičiuotą (9) kriterijų pažymėkime $\eta(\delta)$. Siekiama parodyti, kad $\eta(\delta)$ ir $\tilde{p}(\delta)$ kriterijai ekvivalentūs, kai $\delta \to 0$.

5.1 teorema. *Galioja šis sąryšis:*

$$\tilde{p}(\delta) = \frac{\sigma c \delta^2}{(d+1)(\tilde{z} - y_{on})}\sqrt{\frac{d}{2}} + o(\delta^2). \tag{11}$$

5.2 išvada. *Kai $d = 2$, galioja šis asimptotinis ekvivalentumo sąryšis:*

$$\tilde{p}(\delta) \sim \frac{4\sqrt{3}}{9}\sigma c \eta(\delta). \tag{12}$$

Įrodant 5.1 teoremą, išvestos šios asimptotinės atsitiktinio lauko charakteristikos:

$$m(\mathbf{a}|\mathbf{a}_i,\xi(\mathbf{a}_i),i=1,\ldots,d+1) = \frac{1}{d+1}\sum_{i=1}^{d+1}\xi(\mathbf{a}_i) + o(\delta), \tag{13}$$

$$\frac{s^2(\mathbf{a}|\mathbf{a}_i,\xi(\mathbf{a}_i),i=1,\ldots,d+1)}{\sigma^2} = \frac{c^2 d}{2(d+1)^2}\delta^4 + o(\delta^4),\ \text{kai}\ \delta\to 0. \tag{14}$$

5.2 išvada išreiškia ieškomą ekvivalentumo sąryšį dvimačiu atveju ir pagrindžia euristinio kriterijaus (9) apibrėžimą. Be to, aukštesnės dimensijos atveju 5.1 teorema pateikia supaprastintą pagerinimo tikimybės atitikmenį, kurį galima naudoti naujuose algoritmuose.

6. Optimalus algoritmas Lipšico modeliu grindžiamai vienmatei dvikriterinei optimizacijai

Pastaraisiais metais sukurta keletas sėkmingų vieno kriterijaus Lipšico optimizavimo metodų su stačiakampiais leistinosios srities posričiais, kuriuose hiperstačiakampiai dalijami į tris lygias dalis [13, 18–21]. Siekiant teorinio tokio dalijimo būdo pagrindimo, jis nagrinėjamas optimalumo požiūriu, pradedant nuo santykinai paprasto vienmačio dvikriterinio uždavinio:

$$\begin{aligned}&|f_k(u)-f_k(t)| \le L_k|u-t|, k=1,2, \\ &\forall u,t\in[x_j,x_{j+1}], L=(L_1,L_2)^T, L_k>0, k=1,2.\end{aligned} \tag{15}$$

Toliau bus naudojami šie žymėjimai:

$$f_1(t)=y_t, f_2(t)=z_t, f_1(x_i)=y_i, f_2(x_i)=z_i, i=j,j+1, \tag{16}$$

$$\delta y = |y_j - y_{j+1}|, \delta z = |z_j - z_{j+1}|, \tag{17}$$

$$C = \frac{1}{2}\sqrt{L_1^2+L_2^2}, \tag{18}$$

$$\Psi = \{f(\cdot) : \text{funkcija } f(\cdot) \text{ tenkina (15) sąlygą ir } f(x_i)=(y_i,z_i), i=j,j+1\}. \tag{19}$$

Intervalui $[r_1,r_2], r_1\le r_2$, tokiam, kad $[r_1,r_2]\subseteq[x_j,x_{j+1}]$, kai žinomos reikšmės $f(x_j), f(x_{j+1})$,

apatinio Lipšico rėžio Pareto frontui paklaida lygi

$$\bar{\Delta}((r_1, r_2, x_j, x_{j+1}), (f(x_j), f(x_{j+1}))) = C \times \begin{cases} r_2 - r_1, & \text{jei } r_2 - r_1 \le \beta, \\ \beta, & \text{priešingu atveju,} \end{cases} \tag{20}$$

$$\text{čia } \beta = \beta((x_j, x_{j+1}), (f(x_j), f(x_{j+1}))) = (x_{j+1} - x_j) - \min\left(\frac{\delta y}{L_1}, \frac{\delta z}{L_2}\right). \tag{21}$$

Nagrinėjamas intervalo $[r_1, r_2]$, $x_j \le r_1 < r_2 \le x_{j+1}$, dalijimas į tris dalis taškais $\tilde{a}, \tilde{b} \in (r_1, r_2), \tilde{a} < \tilde{b}$. Viename žingsnyje blogiausiu atveju optimalių dalijimo taškų pasirinkimo uždavinys formuluojamas taip:

$$(\tilde{a}, \tilde{b}) = \arg\min_{a,b} \bar{\Delta}^*((a, b, r_1, r_2, x_j, x_{j+1}), (f(x_j), f(x_{j+1}))); \tag{22}$$

$$\text{čia } \bar{\Delta}^*(\cdot) = \max_{f(\cdot)\in\Psi} \max \{ \bar{\Delta}((r_1, a, x_j, b), (f(x_j), f(b))), \bar{\Delta}((a, b, x_j, x_{j+1}), (f(x_j), f(x_{j+1}))), \bar{\Delta}((b, r_2, a, x_{j+1}), (f(a), f(x_{j+1}))) \}. \tag{23}$$

Šią formuluotę reikia suprasti taip, kad esant blogiausioms įmanomoms kriterijų reikšmėms dalijimo taškuose, $\tilde{a}$ ir $\tilde{b}$ minimizuoja didžiausią iš trims subintervalams apskaičiuotų paklaidų.

6.1 teorema. *Tegu $x_j \le r_1 < r_2 \le x_{j+1}$ ir $\beta = \beta((x_j, x_{j+1}), (f(x_j), f(x_{j+1})))$ apibrėžta (21) formule. (22) uždavinio sprendinys yra: 1) $\tilde{a} = r_1 + \frac{1}{3}(r_2 - r_1), \tilde{b} = r_1 + \frac{2}{3}(r_2 - r_1)$, jei $\frac{1}{3}(r_2 - r_1) < \beta$, ir 2) bet kokie $\tilde{a} \in (r_1, r_2), \tilde{b} \in (r_1, r_2)$, priešingu atveju. Tuomet $\bar{\Delta}^*(\cdot)$ lygu $\frac{C}{3}(r_2 - r_1)$ pirmuoju atveju ir $C\beta$ — antruoju.*

6.1 teorema teigia, kad intervalo dalijimas į tris lygias dalis yra optimalus. Tai pagrindžia analogiškos dalijimo schemos naudojimą literatūroje žinomuose Lipšico optimizacijos algoritmuose. Disertacijoje pateikta teorinė analizė įgalina realizuoti optimalų vieno kintamojo dvikriterinės optimizacijos algoritmą, kuris pašalina dominuojamus intervalus iš tolimesnės paieškos ir efektyviai paskirsto tikslo funkcijų įvertinimo taškus apibrėžimo srityje. Pateiktos teorijos apibendrinimas aukštesnei dimensijai gali būti naudingas tolimesniems adaptyviųjų diagonalinių algoritmų tyrimams.

7. Rezultatai ir išvados

Disertacijoje suformuluotos statistinio modelio pasirinkimo iš Gausinių stochastinių funkcijų klasės rekomendacijos, remiantis pasiūlyta eksperimentine metodika. Pasiūlyta rinktis stochastinę funkciją atsižvelgiant į numatomą tikslo funkcijos sudėtingumą.

Globaliosios optimizacijos algoritmas, paremtas stačiakampe apibrėžimo srities dekompozicija, buvo realizuotas keletu būdų ir eksperimentiškai palygintas su kitais šiuolaikiniais algoritmais. Pasiūlytos dvi euristinės šio algoritmo modifikacijos, derinančios globalią ir lokalią paieškos strategijas. Pirmoji modifikacija, algoritmas *GB*, pakaitomis vykdo išreikštinai apibrėžtas globalią ir lokalią paieškos fazes, atsižvelgdamas į dabartinėje fazėje pasiektą pagerėjimą. Antroji modifikacija, algoritmas *Cluster*, naudoja klasterizavimo procedūrą gerai ištirtiems regionams identifikuoti.

Įrodytas asimptotinis dviejų statistinių simplekso išrinkimo kriterijų, naudojamų globaliosios optimizacijos algoritmuose, ekvivalentumas.

Apibrėžtas viename žingsnyje blogiausiu atveju optimalaus intervalo dalijimo į tris dalis vieno kintamojo dvikriterinėje Lipšico optimizacijoje uždavinys. Optimalumo sąlyga formuluota naudojant Lipšico rėžio Pareto frontui paklaidos sąvoką. Įrodyta, kad optimalumo sąlygą tenkina intervalo dalijimas į tris lygias dalis. Realizuotas atitinkamas optimalus algoritmas.

Iš atliktų tyrimų galima daryti tokias išvadas:

1. Eksperimentų rezultatai rodo, kad *P-algoritmas*, realizuotas naudojant stacionarią izotropinę Gausinę stochastinę funkciją su eksponentine koreliacija, veikia geriausiai optimizuojant įvairias vienmates ir dvimates tikslo funkcijas. Dėl to šią algoritmo versiją rekomenduojama naudoti, kai tikslo funkcijos sudėtingumas nėra iš anksto žinomas. Paprastos tikslo funkcijos gali būti sėkmingai optimizuojamos *Maksimalaus tikėtino pagerinimo algoritmu*, realizuotu naudojant stacionarią izotropinę Gausinę stochastinę funkciją su koreliacija, užtikrinančia mažą svyravimų mastą.

2. Pasiūlytos dvi euristinės globalios paieškos spartinimo strategijos, skatinančios globaliosios paieškos algoritmą *Rect-1* neeikvoti tikslo funkcijos įvertinimų neperspektyvių lokaliųjų minimumų aplinkos tyrinėjimui. Pasiūlytosioms modifikacijoms prireikia mažiau tikslo funkcijos įvertinimų nei pradiniam algoritmui, optimizuojant sudėtingas daugiaekstremalias tikslo funkcijas. Tokiems uždaviniams blogiausiu atveju geriausias *Cluster* algoritmo veikimas.

3. Du simplekso išrinkimo kriterijai, naudojami simpleksiniais statistiniais modeliais besiremiančiuose globaliosios optimizacijos algoritmuose, susieti asimptotiniu ekvivalentumo sąryšiu, mažėjant simplekso dydžiui. Pirmojo kriterijaus išraiška paprasta, tačiau euristinė. Antrasis kriterijus yra pagerinimo tikimybės analogas, tačiau jo išraiška sudėtinga. Natūralu, kad paprastesnė išraiška pranašesnė, tačiau ji teoriškai nepagrįsta. Parodytas ekvivalentumo sąryšis suteikia trūkstamą teorinį paprastosios išraiškos pagrindimą dvimačiu atveju. Be to, gauta supaprastina antrojo kriterijaus išraiška aukštesnės dimensijos atveju, kuri galėtų būti naudojama naujuose optimizacijos algoritmuose.

4. Intervalo dalijimas į tris lygias dalis vieno kintamojo dvikriterinėje Lipšico optimizacijoje yra viename žingsnyje blogiausiu atveju optimalus. Optimalumo sąlyga formuluojama naudojant Lipšico rėžio Pareto frontui apibrėžimo srities intervale paklaidos sąvoką. Šis optimalus dalijimo būdas pagrindžia analogišką dalijimo būdą dviejų kintamųjų dvikriterinėje Lipšico optimizacijoje. Pateiktas tyrimas leidžia pašalinti dominuojamus regionus iš tolimesnės paieškos ir dėl to efektyviai paskirstyti funkcijų įvertinimo taškus apibrėžimo srityje. Pateiktos analizės apibendrinimas aukštesnei dimensijai galėtų būti naudingas tiriant adaptyviuosius diagonalinius algoritmus.

8. Doktorantės publikacijos disertacijos tema

Publikacijos periodiniuose leidiniuose

1. A. Žilinskas, G. Gimbutienė. On one-step worst-case optimal trisection in univariate bi-objective Lipschitz optimization. *Communications in Nonlinear Science and Numerical Simulation*, 35:123 – 136, 2016. ISSN 1007-5704. doi: http://dx.doi.org/10.1016/j.cnsns.2015.11.002.

2. G. Gimbutienė, A. Žilinskas. A two-phase global optimization algorithm for black-box functions. *Baltic Journal of Modern Computing*, 3(3):214, 2015. ISSN 2255-8950.

Publikacijos recenzuojamuose konferencijų leidiniuose

3. A. Žilinskas, G. Gimbutienė. Statistical models for global optimization: how to choose an appropriate one? In A. M. A. C. Rocha, M. Fernanda P. Costa, and E. M. G. P. Fernandes, editors, *Proceedings of the XIII global optimization workshop GOW'16*, University of Minho, Braga, Portugal, 2016. ISBN 978-989-20-6764-3.

4. G. Gimbutienė, A. Žilinskas. Clustering-based statistical global optimization. In *AIP Conference Proceedings*, volume 1776, 2016. doi: http://dx.doi.org/10.1063/1.4965342. URL `http://scitation.aip.org/content/aip/proceeding/aipcp/10.1063/1.4965342`.

5. A. Žilinskas, G. Gimbutienė. On an asymptotic property of a simplicial statistical model of global optimization. In A. Migdalas and A. Karakitsiou, editors, *Optimization, Control, and Applications in the Information Age: In Honor of Panos M. Pardalos's 60th Birthday*, volume 130 of *Springer Proceedings in Mathematics and Statistics*, pages 383–391. Springer International Publishing, Cham, 2015. ISBN 978-3-319-18567-5. doi: http://dx.doi.org/10.1007/978-3-319-18567-5_20.

Pranešimai konferencijose

1. G. Gimbutienė, A. Žilinskas. Investigation of the effect of the assumed statistical model in global optimization. Poster presentation at the *8-th International Workshop "Data Analysis Methods for Software Systems"*, Druskininkai, 2016.
2. A. Žilinskas, G. Gimbutienė. Statistical models for global optimization: how to choose an appropriate one? Presentation at the XIII-th Global Optimization Workshop GOW'16, Braga, Portugal, 2016.
3. G. Gimbutienė, A. Žilinskas. Clustering-based statistical global optimization. Presentation at the 2-nd International Conference and Summer School *Numerical Computations: Theory and Algorithms*, Pizzo Calabro, Italy, 2016.
4. A. Žilinskas, G. Gimbutienė. On one-step worst-case optimal trisection in univariate bi-objective Lipschitz optimization. Poster presentation at the *7-th International Workshop "Data Analysis Methods for Software Systems"*, Druskininkai, 2015.
5. G. Gimbutienė, A. Žilinskas. Dviejų etapų globalaus optimizavimo algoritmas juodos dėžės funkcijoms. Presentation at *Kompiuterininkų dienos - 2015*, Panevėžys, 2015.
6. A. Žilinskas, G. Gimbutienė. On an asymptotic property of a simplicial statistical model of global optimization. Poster presentation at the *6-th International Workshop "Data Analysis Methods for Software Systems"*, Druskininkai, 2014.

9. Trumpai apie autorę

G. Gimbutienė Vilniaus universitete įgijo informatikos bakalauro (2011 m.) ir magistro (2013 m.) laipsnį. Baigus magistrantūros studijas už akademinius pasiekimus jai suteiktas Magna Cum Laude diplomas. 2013–2017 m. — Vilniaus universiteto Matematikos ir informatikos instituto doktorantė. 2010–2011 m. m. pagal Erasmus mainų programą pusmetį mokėsi Aalborgo universitete (Danija), 2012 m. vasarą stažavosi Europos branduolinių mokslinių tyrimų organizacijoje (CERN) (Šveicarija). 2011–2012 m. dirbo "Barclays Technology Centre Limited Lithuanian Branch", Vilniuje, programinės įrangos priežiūros, o 2012–2013 m. — "Nortal, UAB", Vilniuje, programinės įrangos kūrimo srityje.

10. Summary

Research Context

A broad range of applications depends on selecting a collection of certain decision variables corresponding to the best value of some quantifiable objective, e. g. the price of production items. This thesis considers the problem of finding an alternative that is the best globally, formulated as the global optimization problem:

$$\min_{\mathbf{x}\in A} f(\mathbf{x}), \tag{24}$$

where $f(\mathbf{x})$ is a non-convex objective function defined over the feasible region $A = \{\mathbf{x} \in \mathbb{R}^d : a_i \leq x_i \leq b_i, i = 1, \ldots, d\}$. Usually the problem (24) is approached by numerical algorithms, as the analytical solutions are known only in exceptional cases.

There are situations when a single objective does not completely define the choice of the decision vector as several conflicting goals are involved. The problem statement is generalized to a set of objectives $f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})$, $k \geq 2$, $\mathbf{x} \in A$:

$$\min_{\mathbf{x}\in A}(f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})). \tag{25}$$

A solution to (25) is represented by a set of vectors, complying with the concept of Pareto optimality. It means that resulting vectors of objectives cannot be improved in terms of one objective without sacrificing some other objective. Approximating the Pareto set is a difficult problem, generally approached by numerical algorithms.

This thesis considers algorithms that use statistical and Lipschitz objective function models to solve the problems (24) and (25) in a rather common engineering situation where the objective function is available as a *black-box*. A statistical objective function model is useful when the objective function is expensive to evaluate. After some function evaluations have already been obtained, the statistical model can be viewed as a predictor of function values with an associated uncertainty measure at locations yet unexplored. Based on the model, an optimization algorithm assesses the suitability of a new point $\mathbf{x} \in A$ for becoming the next trial point. On the other hand, Lipschitz optimization methods rely on a key assumption that the rate of change of the objective function is bounded. Functions satisfying this assumption are said to comply with the Lipschitz model. Although the theoretical investigation of algorithms in this class is relatively simple, the main problem of the model is the generally unknown Lipschitz constant, i. e. the rate of change bound.

Relevance of the Study

The choice of a specific statistical model in global optimization algorithms should be based on the a priori available information about the problem, as well as computational complexity considerations. Therefore a systematic investigation of the impact that different models have on the optimization results of objectives with various characteristics would be useful.

Optimization methods using statistical objective function models primarily target expensive problems, therefore the number of trials has to be reduced as much as possible. The approximation of the global minimum by a recent global optimization algorithm in this class is decelerated by the presence of other local minimizers. Ways of preventing the algorithm from excessively exploring their vicinity are needed to save resources.

A simplicial statistical model was used in a recent global optimization algorithm with an established convergence rate. The algorithm operates by selecting a simplex to be partitioned at each iteration, based on a relatively simple heuristic criterion. A theoretical justification of this criterion could be established by mathematically linking it to the probability of finding a better function value inside the considered simplex.

Recently single-objective Lipschitz optimization algorithms based on adaptive diagonal partitions proved successful. These methods apply a trisection procedure to divide hyper-rectangles. Moreover, the same trisection procedure was used in bivariate bi-objective Lipschitz optimization. A theoretically interesting task is to demonstrate the optimality of the trisection procedure in question, starting from a relatively simple case of the univariate bi-objective Lipschitz optimization. Although, in general, the applicability of optimal algorithms is narrow, they can be used as benchmarks for comparison; moreover, some of their properties might be shared by other practically applicable algorithms.

Objectives of the Thesis

The present thesis aims at efficient and theoretically justified search for optimal solutions. Analysis of theoretical properties, as well as heuristic extensions in single-objective global optimization using statistical models are provided. Furthermore, the worst-case optimality in the univariate bi-objective Lipschitz optimization is investigated.

The objectives of the study are:

1. Investigate the effect of the assumed statistical model on the performance of global optimization algorithms.
2. Increase the global minimum approximation efficiency in terms of the number of function evaluations of a recent global optimization algorithm, relying on a hyper-rectangular decomposition-adjusted statistical objective function model.
3. Theoretically support the definition of a simplicial global optimization algorithm, using properties of a simplicial statistical model.
4. Investigate the worst-case optimality in univariate bi-objective Lipschitz optimization.

Scientific Novelty and Results

An experimental methodology was proposed aimed at a systematic assessment of the impact that the selected statistical objective function model has on the optimization of objective functions with various characteristics. Guidelines for the model selection were formulated based on the experimental results with a number of 1- and 2-dimensional Gaussian stochastic functions considered as statistical models.

Two heuristic extensions for a recent statistical global search algorithm relying on a hyper-rectangular decomposition of the feasible region were proposed, attempting to reduce the number of trials required solving difficult optimization problems. The suggested modifications combine global and local search techniques and are experimentally shown to perform better than the original and other contemporary algorithms under consideration.

Theoretical support for using a heuristic simplex selection criterion in a recent simplicial global optimization algorithm was provided. This was achieved by linking the criterion to the probability of improvement-related expression, when a stationary isotropic Gaussian random field, defined over a simplex, is used as a statistical objective function model.

The one-step worst-case optimal trisection of an interval in univariate bi-objective Lipschitz optimization was investigated with respect to the tolerance of the Lipschitz bound for the Pareto front. It was theoretically shown that trisection of an interval into three equal parts satisfies the considered definition of optimality. A corresponding optimal bi-objective optimization algorithm was implemented and its performance was demonstrated.

Statements to Be Defended

1. The *P-algorithm*, constructed assuming a stationary isotropic Gaussian stochastic function with an exponential correlation, performs the best for a variety of univariate and bivariate objective functions.
2. The *Maximum expected improvement algorithm*, constructed assuming a stationary isotropic Gaussian stochastic function with a correlation structure ensuring low short-range variability, is appropriate to use for optimizing relatively simple objective functions.
3. Balancing the local and global search strategies in proposed hybrid algorithms *GB* and *Cluster* consumes fewer function evaluations compared to the original global optimization algorithm *Rect-1* that they extend, when difficult multi-modal global optimization problems are optimized.
4. A heuristically defined simplex selection criterion is asymptotically equivalent to a probability of improvement-related expression.
5. The trisection of an interval into three equal parts in the univariate bi-objective Lipschitz optimization is one-step worst-case optimal.

Publications by the Author

1. A. Žilinskas and G. Gimbutienė. On one-step worst-case optimal trisection in univariate bi-objective Lipschitz optimization. *Communications in Nonlinear*

Science and Numerical Simulation, 35:123 – 136, 2016. ISSN 1007-5704. doi: http://dx.doi.org/10.1016/j.cnsns.2015.11.002.

2. G. Gimbutienė and A. Žilinskas. A two-phase global optimization algorithm for black-box functions. *Baltic Journal of Modern Computing*, 3(3):214, 2015. ISSN 2255-8950.

3. A. Žilinskas and G. Gimbutienė. Statistical models for global optimization: how to choose an appropriate one? In A. M. A. C. Rocha, M. Fernanda P. Costa, and E. M. G. P. Fernandes, editors, *Proceedings of the XIII global optimization workshop GOW'16*, University of Minho, Braga, Portugal, 2016. ISBN 978-989-20-6764-3.

4. G. Gimbutienė and A. Žilinskas. Clustering-based statistical global optimization. In *AIP Conference Proceedings*, volume 1776, 2016. doi: http://dx.doi.org/10.1063/1.4965342. URL `http://scitation.aip.org/content/aip/proceeding/aipcp/10.1063/1.4965342`.

5. A. Žilinskas and G. Gimbutienė. On an asymptotic property of a simplicial statistical model of global optimization. In A. Migdalas and A. Karakitsiou, editors, *Optimization, Control, and Applications in the Information Age: In Honor of Panos M. Pardalos's 60th Birthday*, volume 130 of *Springer Proceedings in Mathematics and Statistics*, pages 383–391. Springer International Publishing, Cham, 2015. ISBN 978-3-319-18567-5. doi: http://dx.doi.org/10.1007/978-3-319-18567-5_20.

About the Author

G. Gimbutienė earned her Bachelor (2011) and Master (2013) of Informatics at Vilnius University. She completed the Master's studies with an honours "Magna Cum Laude"diploma for her academic achievements. From 2013 to 2017 she was a PhD student at Vilnius University. She spent one semester as an Erasmus student at Aalborg university (Denmark) during academic year 2010-2011. Moreover, she worked for the European Organization of Nuclear Research (CERN) (Switzerland) as an intern during the summer of 2012. Her industrial experience includes application management at "Barclays Technology Centre Limited Lithuanian Branch"(Vilnius) during 2011-2012 and software development at "Nortal, UAB"(Vilnius) during 2012-2013.

Literatūra

[1] GKLS test functions generator implementation home page. `http://wwwinfo.deis.unical.it/~yaro/GKLS.html`.

[2] J. Calvin. On a global optimization algorithm for multivariate smooth functions. (Private communication.).

[3] J. Calvin. An adaptive univariate global optimization algorithm and its convergence rate under the Wiener measure. *Informatica*, 22(4):471–488, 2011.

[4] J. Calvin and A. Žilinskas. One-dimensional P-algorithm with convergence rate O(n-3+δ) for smooth functions. *Journal of Optimization Theory and Applications*, 106(2): 297–307, 2000. ISSN 0022-3239. doi: http://dx.doi.org/10.1023/A:1004699313526. URL `http://dx.doi.org/10.1023/A%3A1004699313526`.

[5] J. M. Calvin and A. Žilinskas. On a global optimization of bivariate smooth functions. *Journal of Optimization Theory and Applications*, 163:528–547, 2014.

[6] J. M. Calvin, Y. Chen, and A. Žilinskas. An adaptive univariate global optimization algorithm and its convergence rate for twice continuously differentiable functions. *Journal of Optimization Theory and Applications*, 155(2):628–636, 2012. ISSN 0022-3239. doi: http://dx.doi.org/10.1007/s10957-012-0060-3. URL `http://dx.doi.org/10.1007/s10957-012-0060-3`.

[7] Yu. M. Danilin and S. A. Pijavskij. An algorithm for finding the absolute minimum. *Theory of Optimal Decisions*, 2:25–37, 1967.

[8] J. M. Gablonsky. *Modifications of the DIRECT algorithm*. PhD thesis, North Caroline State University, Raleigh, North Carolina, 2001.

[9] J. M. Gablonsky and C. T. Kelley. A locally-biased form of the DIRECT algorithm. *Journal of Global Optimization*, 21(1):27–37, 2001.

[10] M. Gaviano, D. E. Kvasov, D. Lera, and Y. D. Sergeyev. Algorithm 829: Software for generation of classes of test functions with known local and global minima for global optimization. *ACM Transactions on Mathematical Software*, 29(4):469–480, 2003.

[11] D. R. Jones, C. D. Perttunen, and B. E. Stuckman. Lipschitzian optimization without the Lipschitz constant. *Journal of Optimization Theory and Applications*, 79(1):157–181, 1993.

[12] H. Kushner. A versatile stochastic model of a function of unknown and time-varying form. *Journal of Mathematical Analysis and Applications*, 5:150–167, 1962.

[13] D. E. Kvasov and Ya. D Sergeyev. Lipschitz gradients for global optimization in a one-point-based partitioning scheme. *Journal of Computational and Applied Mathematics*, 236(16):4042–4054, 2012.

[14] J. Mockus. On Bayes methods for seeking an extremum. *Avtomatika i Vychislitelnaja Technika*, 3:53–62, 1972. (In Russian.).

[15] R. Paulavičius, Y. D. Sergeyev, D. E. Kvasov, and J. Žilinskas. Globally-biased DISIMPL algorithm for expensive global optimization. *Journal of Global Optimization*, 59(2-3):545–567, 2014.

[16] S. Pijavskij. An algorithm for finding the absolute extremum of a function. *USSR Computational Mathematics and Mathematical Physics*, 12(4):57–67, 1972.

[17] Y. D. Sergeyev and D. E. Kvasov. Global search based on efficient diagonal partitions and a set of Lipschitz constants. *SIAM Journal on Optimization*, 16(3):910–937, 2006. doi: http://dx.doi.org/10.1137/040621132.

[18] Ya. D. Sergeyev. Efficient strategy for adaptive partition of N-dimensional intervals in the framework of diagonal algorithms. *Journal of Optimization Theory and Applications*, 107(1):145–168, 2000.

[19] Ya. D. Sergeyev. Efficient partition of N-dimensional intervals in the framework of one-point-based algorithms. *Journal of Optimization Theory and Applications*, 124(2): 503–510, 2005. ISSN 0022-3239. doi: http://dx.doi.org/10.1007/s10957-004-0948-7. URL `http://dx.doi.org/10.1007/s10957-004-0948-7`.

[20] Ya. D. Sergeyev and D. Kvasov. *Diagonal global optimization methods*. Fizmatlit, Moscow, 2008. (In Russian.).

[21] Ya. D. Sergeyev and D. Kvasov. A deterministic global optimization using smooth diagonal auxiliary functions. *Communications in Nonlinear Science and Numerical Simulation*, 21(1):99 – 111, 2015.

[22] B. O. Shubert. A sequential method seeking the global maximum of a function. *SIAM Journal on Numerical Analysis*, 9(3):379–388, 1972.

[23] A. Žilinskas. Axiomatic approach to statistical models and their use in multimodal optimization theory. *Mathematical Programming*, 22(1):104–116, 1982.

[24] A. Žilinskas. A statistical model for global optimization by means of select and clone. *Optimization: A Journal of Mathematical Programming and Operations Research*, 48 (1):117–135, 2000. doi: http://dx.doi.org/10.1080/02331930008844497.

Gražina Gimbutienė

STATISTINIAIS IR LIPŠICO TIKSLO FUNKCIJOS MODELIAIS PAGRĮSTI NEIŠKILOS GLOBALIOSIOS OPTIMIZACIJOS ALGORITMAI

Daktaro disertacija

Fiziniai mokslai (P000)
Informatika (09P)
Redaguota "Vertimų karaliai", MB

Gražina Gimbutienė

ALGORITHMS FOR NON-CONVEX GLOBAL OPTIMIZATION BASED ON THE STATISTICAL AND LIPSCHITZ OBJECTIVE FUNCTION MODELS

Doctoral dissertation

Physical sciences (P000)
Informatics (09P)
Editor Nijolė Požėraitytė