LMD mokslo darbai, 212–218
1998, Vilnius

Daugiamačių modelių struktūros priklausomybė nuo apibrėžimo srities

V. Šaltenis, V. Tiešis (MII)

1. Įvadas

Prastinant sudėtingus, įvairiose mokslo ir technikos srityse iškylančius daugiamačius modelius dažnai naudojama dekompozicija – skaidymas į paprastesnius, mažesnio matavimų skaičiaus modelius. Tokio skaidymo leistinumo kriterijai dažnai būna heuristiniai. V. Šaltenio pasiūlyta daugiamačių modelių struktūros analizė [2], plačiau naudota optimizavimo modeliuose, leidžia vertinti modelio prastinimo (tame tarpe ir dekompozicijos) sukeliamas pasekmes. Modelis yra suskaidomas, jei sąveika tarp kintamųjų ar jų grupių yra silpna. Deja, praktiniai modeliai dažniausiai turi nemažai neatmestinų kintamųjų sąveikų.

Straipsnyje siekiama panagrinėti, kaip keičiasi sąveika, keičiant modelio kintamųjų kitimo sritį. Parodoma, kad sričiai be galo mažėjant, sąveikos turi išnykti. Naudojantis testine dvimate funkcija iliustruojama, kaip dalinant sritį, sąveiką galima sumažinti iki praktiškai leistinų ribų.

2. Daugiamačių modelių struktūros charakteristikos

Daugelio kintamųjų funkcijos skaidymu į mažesnio kintamųjų skaičiaus funkcijų sumą, nagrinėtu Šaltenio ir Sobolio [2], [3], remsimės analizuodami daugiamačių modelių struktūrą.

Tegu funkcija $f(X)=f(x_1,\dots,x_n)$ paprastumo dėlei apibrėžta srityje K^n $(a_1 \le x_1 \le b_1,\dots,a_n \le x_n \le b_n): X \in K^n$. Kartais naudosime sutrumpintus funkcijų žymėjimus be argumentų, pvz. f vietoj $f(x_1,\dots,x_n)$.

Įveskime pažymėjimus funkcijos f apibrėžimo sričiai, kuri yra bazinių sričių $\Omega_1,\dots,\Omega_n$ Dekarto sandauga: $\Omega=\Omega_1\times\cdots\times\Omega_n$, ir sritims:

$$\Omega_{i_1\dots i_s}=\Omega_{i_1}\times\cdots\times\Omega_{i_s},\quad 1\le i_1<\dots<i_s\le n\ ,\quad s=1,\dots,n$$

$$\Omega_{(i)}=\Omega_1\times\cdots\times\Omega_{i-1}\times\Omega_{i+1}\times\cdots\times\Omega_n\ ,\quad i=1,\dots,n;$$

$$\Omega_{(ij)}=\Omega_1\times\cdots\times\Omega_{i-1}\times\Omega_{i+1}\times\cdots\times\Omega_{j-1}\times\Omega_{j+1}\times\cdots\times\Omega_n\,,$$

$i,j=1,\dots,n\,,\quad i<j\,.$

Bendru atveju sritis $\Omega_{(i_1...i_s)}$ apibrėžiama analogiškai. Atitinkamus šių sričių Lebego matus žymėsime $\mu, \mu_{i_1..i_s}$ ir $\mu_{(i_1..i_s)}$.

Naudosime indeksų grupes $i_1,...,i_s$, $1 \le i_1 < ... < i_s \le n$, $s = 1,...,n$ ir įvesime pažymėjimą sumai:

$$\hat{\sum} T_{i_1...i_s} = \sum_{s=1}^{n} \sum_{1 \le i_1 < .. < i_s \le n} .. \sum T_{i_1..i_s}$$

Jei $f \in L_2$, tai egzistuoja vienintelis f skaidymas į tarpusavyje ortogonalias centruotas dedamasias (žr. [3]):

$$f = f_0 + \hat{\sum} f_{i_1...i_s}(x_{i_1},...,x_{i_s}) \tag{1}$$

kur f_0 yra konstanta ir tenkinamos šios sąlygos:

$$\int_{\Omega_{i_k}} f_{i_1...i_s}(x_{i_1},...,x_{i_s})dx_{i_k} = 0, \quad 1 \le k \le s.$$

Įveskime pažymėjimą s kintamųjų funkcijai:

$$f^{i_1...i_s} = \frac{1}{\mu_{(i_1..i_s)}} \int_{\Omega_{(i_1...i_s)}} f.$$

Tada, suintegravus (1) srityje Ω, pastovi dedamoji bus lygi:

$$f_0 = \frac{1}{\mu} \int_{\Omega} f, \tag{2}$$

o suintegravę (1) srityse $\Omega_{(i)}$ ir $\Omega_{(ij)}$ gausime vienmates ir dvimates dedamasias:

$$f_i(x_i) = f^i - f_0, \quad i = 1,...,n,$$

$$f_{ij}(x_i, x_j) = f^{ij} - f_0 - f_i(x_i) - f_j(x_j), \quad i,j = 1,...,n, \quad i < j, \qquad \text{ir t. t.}$$

Struktūros charakteristikų sistema:

$$D = \hat{\sum} D_{i_1...i_s},$$

$$D_{i_1...i_s} = \frac{1}{\mu} \int_{\Omega} (f_{i_1...i_s})^2, \tag{3}$$

$$D = \frac{1}{\mu} \int_{\Omega} (f)^2 - (f_0)^2 \tag{4}$$

pasiūlyta, tirta (žr. [2]) ir daugiausiai naudota analizuojant daugiamačių optimizavimo uždavinių struktūrą. Struktūros charakteristikos D_i rodo atskirų kintamųjų, o $D_{i_1 \dots i_s}$ – jų grupių įtaką (sąveiką) aproksimuojant funkciją mažesnio kintamųjų skaičiaus centruotomis ortogonaliomis funkcijomis. Paprastai šios charakteristikos normalizuojamos (visų charakteristikų suma lygi vienetui).

Lentelėje 1 pateikti paprasčiausių funkcijų struktūros pavyzdžiai.

Lentelė 1. Paprasčiausių funkcijų struktūros charakteristikos, $\Omega_i = [0,1]$.

$f(X)$	D_1	D_2	D_{12} (sąveika)
$x_1 + x_2$	0.5	0.5	0
$x_1 x_2 - \frac{x_1 + x_2}{2}$	0	0	1

Paprastai šiuolaikiniuose sudėtinguose modeliuose galime gauti tik funkcijos $f(X)$ reikšmes tam tikruose taškuose X^j ($j=1,\dots,N$). Jei jie tolygiai pasiskirstę srityje Ω, galime naudoti Monte-Karlo metodą charakteristikų įvertinimui, naudodamiesi (2), (3) ir (4):

$$f_0 \approx \frac{1}{N} \sum_{j=1}^{N} f(X^j) \ ,$$

$$D + \left(f_0\right)^2 \approx \frac{1}{N} \sum_{j=1}^{N} \left(f \ (X^j)\right)^2 .$$

Vertinant $D_{i_1 \dots i_s}$ (pagal [3]):

$$D_{i_1 \dots i_s} + \left(f_0\right)^2 \approx \frac{1}{N} \sum_{j=1}^{N} f(Y^j, Z^j) f(Y^j, U^j)$$

naudosime poras taškų, kurių s tolygiai srityje $\Omega_{i_1 \dots i_s}$ pasiskirsčiusios kcordinatės $Y^j = (x_{i_1}, \dots, x_{i_s})$ sutampa, o $n - s$ likusios koordinatės Z^j ir U^j yra tolygiai pasiskirstę srityje $\Omega_{(i_1 \dots i_s)}$.

3. Sąveika be galo mažėjančioje srityje

Šiame paragrafe parodysime, kad apie nestacionarų tašką visada galime suformuoti pakankamai mažą sritį kurioje $\sum_{i=1}^{n} D_i / D$ būtų kiek norima arti vieneto. Tai yra, mažindami sritį galime kiek norima sumažinti kintamųjų tarpusavio sąveikas.

Tarkim, kad $f(X)$ yra dukart tolygiai diferencijuojama funkcija uždaroje srityje Ω. Taigi $f(X)$ antros išvestinės bus aprėžtos, ką naudosime skaičiuodami struktūrinių charakteristikų ribines išraiškas. Skaičiuodami integralus (2)–(4), funkciją $f(X)$ skleisime Teiloro eilute apie centrinį tašką X_c, $x_{ci} = 0,5(b_i+a_i)$:

$$f(X) = f(X_c) + \sum_i \frac{\partial f}{\partial x_i}\bigg|_{X_c} (x_i - x_{ci}) + \frac{1}{2}\sum_i\sum_j \frac{\partial f}{\partial x_i \partial x_j}\bigg|_{\xi(X)} (x_i - x_{ci})(x_j - x_{cj}) \qquad (5)$$

kur $\xi(X) \in [X,X_c]$. Čia ir toliau visų sumų kintamieji kinta nuo 1 iki n. Integruodami pointegralinių išraiškų liekamuosius narius, kuriuose yra išvestinės, priklausančios nuo pointegralinio kintamojo, naudosimės apibendrinta vidurinės reikšmės teorema, tas išvestines keisdami jų vidurine reikšme, kuri, kaip minėjome, yra aprėžta. Paskaičiuokime f_0:

$$f_0 = \frac{1}{\mu}\int_\Omega f(X) = f(X_c) + \sum_i \frac{\partial f}{\partial x_i}\bigg|_{X_c} \frac{1}{b_i - a_i}\int_{a_i}^{b_i} (x_i - x_{ci})dx_i +$$

$$+\frac{1}{2}\sum_i \frac{1}{b_i - a_i}\int_{a_i}^{b_i}\left[\frac{1}{\mu(i)}\int_{\Omega(i)} \frac{\partial^2 f}{\partial^2 x_i}\bigg|_{\xi(X)}\right](x_i - x_{ci})^2 dx_i + \qquad (6)$$

$$+\sum_i \frac{1}{b_i - a_i}\sum_j \frac{1}{b_j - a_j}\int_{a_i}^{b_i}\int_{a_j}^{b_j}\left[\frac{1}{\mu(ij)}\int_{\Omega(ij)} \frac{\partial^2 f}{\partial x_i \partial x_j}\bigg|_{\xi(X)}\right](x_i - x_{ci})(x_j - x_{cj})dx_i dx_j$$

Panagrinėkime formulės (6) kairės pusės dėmenis. Integralas antrame dėmenyje lygus nuliui. Trečiam dėmeniui taikome vidurinės reikšmės teoremą, pakeisdami tolydinę išvestinę jos vidurine reikšme: $\eta_i = \dfrac{\partial^2 f}{\partial^2 x_i}\bigg|_V, V \in \Omega$ ir tada šis dėmuo lygus $\dfrac{1}{2}\sum_i (b_i - a_i)^2 \eta_i$. Skaičiuodami ketvirtą dėmenį, vidurinės reikšmės teoremą galime taikyti atskiriems Ω_{ij} kvadrantams, kuriuose nesikeičia $(x_i - x_{ci})(x_j - x_{cj})$ ženklas. Pažymėkim $\eta_{ij}^k = \dfrac{\partial^2 f}{\partial x_i \partial x_j}\bigg|_V$, kur vektoriaus v ij-tos koordinatės priklauso k-tam srities Ω_{ij} kvadrantui, sunumeruotam įprasta tvarka prieš laikrodžio rodyklę. Tada ketvirtas dėmuo lygus

$$\frac{1}{64}\sum_i\sum_j (b_i - a_i)(b_j - a_j)(\eta_{ij}^1 + \eta_{ij}^3 - \eta_{ij}^2 - \eta_{ij}^4).$$

Pažymėkim $r_i\Delta = (b_i - a_i)$, kur r_i yra srities Ω kraštinės mastelis, o Δ yra srities dydžio koeficientas. Tada įstatę dėmenų reikšmes į (6) gauname

$$f_0 = f(X_c) + R_0\Delta^2, \tag{7}$$

kur R_0 priklauso nuo funkcijos f antros eilės išvestinių srityje Ω ir yra tolydinė bei aprėžta, kaip ir šios išvestinės.

Skaičiuodami D į išraišką (4) statome (5) bei (7) išraiškas ir integralus įvertiname analogiškai, kaip ir skaičiuodami f_0. Gauname:

$$D = \frac{1}{12}\sum_i \left(\left.\frac{\partial f}{\partial x_i}\right|_{X_c} \right)^2 (b_i - a_i)^2 + o(\Delta^2) \tag{8}$$

Analogiškai skaičiuojame dedamąją f_k :

$$f_k(x_k) = \frac{1}{\mu_{(k)}} \int_{\Omega_{(k)}} (f(X) - f_0)$$

$$= \left.\frac{\partial f}{\partial x_k}\right|_{X_c} (x_k - x_{ck}) + 0.5 \left.\frac{\partial^2 f}{\partial^2 x_k}\right|_{\xi(x_k)} (x_k - x_{ck})^2$$

$$+ 0.5 \left.\frac{\partial^2 f}{\partial^2 x_k}\right|_{\xi(x_k)} (x_k - x_{ck})^2 + (x_k - x_{ck})R_1(x_k)\Delta + R_2(x_k)\Delta^2 + R_0\Delta^2 \tag{9}$$

kur $R_1(x_k)$, $R_2(x_k)$ yra funkcijos tolydinės intervale $[a_k, b_k]$, priklausančios nuo funkcijos f antrųjų išvestinių ir yra aprėžtos kaip ir šios išvestinės; vidurinės reikšmės argumento $\xi(x_k) \in \Omega$ k-toji koordinatė priklauso intervalui $[x_k, x_{ck}]$.

Integruodami (9) ir funkcijoms $R_1(x_k)$, $R_2(x_k)$ taikydami vidurinės reikšmės teoremą gauname struktūrinę charakteristiką:

$$D_k = \frac{1}{\mu} \int_\Omega [f_k(x_k)]^2 = \left(\left.\frac{\partial f_k}{\partial x_k}\right|_{X_c} \right)^2 (b_k - a_k)^2 / 12 + o(\Delta^2) \tag{10}$$

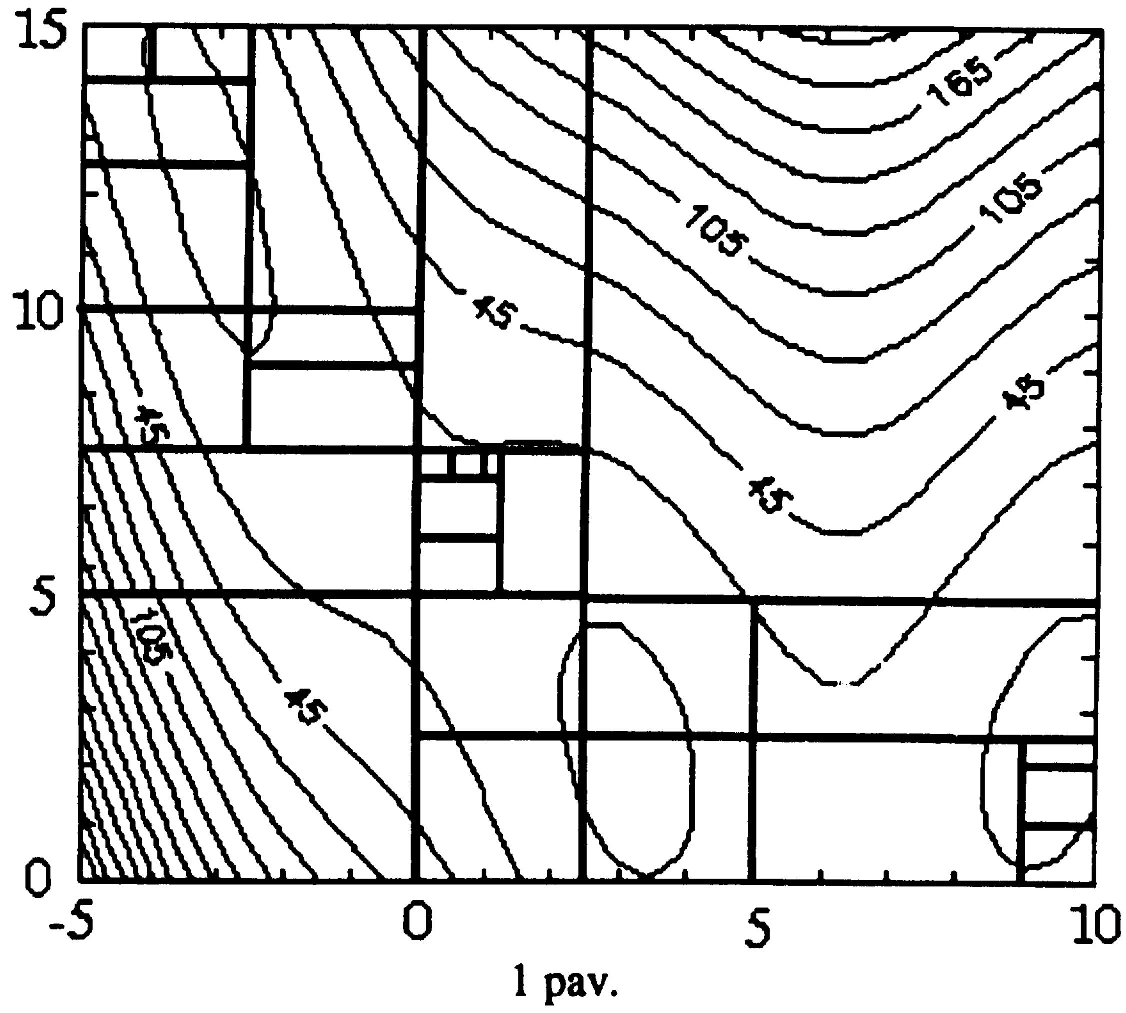
Dalindami (10) iš (8) prie $\sum_i \left(\left.\frac{\partial f}{\partial x_i}\right|_{X_c} \right) \neq 0$ gauname ką ir norėjome įrodyti:

$$\frac{D_k}{D} = \frac{\left(\left.\frac{\partial f_k}{\partial x_k}\right|_{X_c}\right)^2 r_k^2 + O(\Delta)}{\sum_i \left(\left.\frac{\partial f_i}{\partial x_i}\right|_{X_c}\right)^2 r_i^2 + O(\Delta)} \xrightarrow[\Delta \to 0]{} \frac{\left(\left.\frac{\partial f_k}{\partial x_k}\right|_{X_c}\right)^2 r_k^2}{\sum_i \left(\left.\frac{\partial f_i}{\partial x_i}\right|_{X_c}\right)^2 r_i^2}$$

ir $\sum_k \frac{D_k}{D} \xrightarrow[\Delta \to 0]{} 1$. Pastebėsime, kad ribiniu atveju kvadratinei sričiai Ω struktūrinės charakteristikos D_i proporcingos gradiento koordinačių kvadratams. Stačiakampėms sritims šios charakteristikos taip pat proporcingos atitinkamų kraštinių mastelių kvadratams.

4. Pavyzdys

Naudosimės optimizavime paplitusia dvimate Branino testine funkcija, žr. [1]. Jos lygio linijos (plonos) iliustruojamos 1 pav. Tokioje apibrėžimo srityje funkcijos kintamieji stipriai sąveikoti (sąveiką nusakanti charakteristika $D_{12} = 0.60$. Dalinant sritį į dalis buvo siekta, kad jose sąveikos neviršytų 0.03. Storesnėmis linijomis parodytas toks srities dalinimas į dekomponuojamas dalis.



1 pav.

5. Išvados

Pademonstruotos principinės galimybės suskaidyti funkcijos apibrėžimo sritį į dalis, kurių viduje beveik nėra atskirų kintamųjų tarpusavio sąveikos. Tai įgalina sudaryti paprastesnius daugiamačius modelius.

LITERATŪRA

[1] L.C.W. Dixon and G.P. Cziego. The global optimisation problem: An introduction, *Towards Global Optimisation*, Amsterdam: North-Holland, 1978, 2, p.1-15.
[2] V. Šaltenis, *Structure analysis of optimisation problems* (rusų klb.). V.:Mokslas, 1989, 123 p.
[3] I.M. Sobol', On sensitivity estimation for nonlinear mathematical models, *Matematitcheskoye modielirovanye*, 1990, 2(1), p.112 -118.

The impact to multidimensional models by the range of definition

V. Šaltenis, V. Tiešis

The aim of the research was to investigate ways for simplification of multidimensional models. The simplification mode based on the system of characteristics, that is similar to the dispersion analysis, was investigated. It was shown how characteristics may be used to divide the model into several more simple submodels, for example, linear models. The core of the simplification is the division of a range of definition where the characteristics serve as an indicator of accuracy and type of submodels.