LMD mokslo darbai, III tomas, 143–148

Elipsoidų ir simplekso metodų palyginimas

Eligijus LAURINAVIČIUS, Danutė RAŠKINIENĖ (LŽŪU)
e-mail: elis@info.lzua

Darbe nagrinėjamas tiesinio programavimo uždavinio

$$\mathbf{C}^{\mathrm{T}}\mathbf{X} = F \to \max \tag{1}$$

$$\mathbf{a}_j^{\mathrm{T}}\mathbf{X} \leqslant b_j,\ j \in \mathbf{M},\ \mathbf{M} = \{1, \ldots, m\} \tag{2}$$

sprendimas, panaudojant bazinio sprendinio elipsoidą ir nebazinio sprendinio atvirkštinę matricą. Čia

$$\mathbf{C}^{\mathrm{T}} = (c_1, c_2, \ldots, c_n),\ \mathbf{X}^{\mathrm{T}} = (x_1, x_2, \ldots, x_n),\ \mathbf{a}_j^{\mathrm{T}} = (a_{j1}, a_{j2}, \ldots, a_{jn}).$$

Pastebėsime, kad apribojimai $x_j \geqslant 0,\ i \in \mathbf{N},\ \mathbf{N} = \{1, \ldots, n\},\ \mathbf{N} \subset \mathbf{M}$, jei jie būtini, keičiami į apribojimus $-x_j \leqslant 0$. Uždavinio (1)–(2) sprendimui neįvedami jokie papildomi kintamieji ir nekeliamos kintamųjų neneigiamumo sąlygos, būtinos daugelyje žinomų tokio uždavinio sprendimo algoritmų.

Simplekso metodo pagrindas – iteratyvi procedūra, leidžianti, pagal tam tikrą taisyklę perskaičiuojant (1)–(2) koeficientus, per baigtinį žingsnių skaičių pasiekti tikslų uždavinio sprendinį. Sprendimui reikalingo maksimalaus žingsnių skaičiaus baigtinumas akivaizdus ($K_{\max} \leqslant C_m^n$), tačiau tų pačių matmenų (n, m) uždaviniuose būtinasis žingsnių skaičius gali būti labai skirtingas. Praktinių simplekso metodo taikymų statistika rodo, kad uždaviniui (1)–(2) išspręsti būtinas žingsnių skaičius retai kada daugiau nei keletą kartų viršija uždavinio matmenis. Tačiau yra uždavinių, kurie simplekso metodu neįveikiami per priimtiną sprendimui skirtą laiką. Pavyzdžiui, Klee-Minty uždavinys

$$\begin{array}{lllll}
2^{n-1}x_1 & + \ 2^{n-2}x_2 & + \ldots 2x_{n-1} + x_n & \to & \max \\
x_1 & & & \leqslant & 5 \\
4x_1 \ + & x_2 & & \leqslant & 25 \\
8x_1 \ + & 4x_2 + x_3 & & \leqslant & 125 \\
\ldots & \ldots & & & \ldots \\
2^n x_1 & + \ 2^{n-1}x_2 & + \ldots + 4x_{n-1} + x_n & \leqslant & 5^n
\end{array}$$

$$x_j \geqslant 0,\ i \in \mathbf{N}, \mathbf{N} = \{1, \ldots, n\}$$

simplekso metodu išsprendžiamas per 2^n žingsnių, pradedant sprendimą nuo nulinio bazinio sprendinio. Palyginti kuklių matmenų ($n = 50,\ m = 100$) Klee-Minty uždavinį tektų spręsti kelis dešimtmečius. Taip yra dėl to, kad simplekso metodas „žingsniuoja“ uždavinio leistinų sprendinių srities pakraščiu, kuriame būtinai yra optimalus sprendinys. Deja, neretai šis kelias būna ne pats trumpiausias. Sukurta kitokiais principais grindžiamų (1)–(2) uždavinio sprendimo metodų. Paminėtini apibrėžtų ir įbrėžtų elipsoidų, Karmarkaro (vidinio taško) metodai. Labiausiai paplitęs konkurentabilus simplekso metodui ir net efektyvesnis už jį yra polinomialus Karmarkaro metodas. Darbe pateikiamas metodas, skirtingai nuo apibrėžtų ir įbrėžtų elipsoidų, elipsoidą naudoja ne srities, kuriame yra optimalus sprendinys nustatymui, o krypties į optimalų sprendinį apskaičiavimui. Darbe pateikiamo metodo idėja realizuojama tokiais etapais: 1) iš bazinio sprendinio apribojimų konstruojamas elipsoidas; 2) optimizavimo kryptis, garantuojanti tikslo funkcijos didėjimą ir išliekanti leistinų sprendinių srityje, apskaičiuojama panaudojant elipsoidą; 3) tikslo funkcijos vektorius projektuojamas į nebazinio sprendinio poaibį, pasinaudojant nebazinio sprendinio atvirkštine matrica.

Bazinio sprendinio elipsoidas

Tarkime, kad rastas bazinis leistinas neoptimalus sprendinys. Bendrumo nesumažinsime sakydami, kad jis sudarytas iš pirmųjų n $(n < m)$ sistemos (2) apribojimų. Bazinį sprendinį pilnai nusako bazės apribojimų vektorių matrica $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n)$, jos atvirkštinė matrica $\mathbf{A}^{-1}$ bei vektoriai $\mathbf{B} = (b_1, b_2, \ldots, b_n)$ ir $\mathbf{C}$. Žinoma, kad bazinis sprendinys $\mathbf{X}_B = (\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{B}$ (kuriame $\mathbf{C}^{\mathrm{T}}\mathbf{X}_B = F_B$) optimalus, jeigu tenkinama sąlyga

$$\mathbf{A}^{-1}\mathbf{C} = \mathbf{P} \geqslant 0, \tag{3}$$

t.y. jei $p_i \geqslant 0$, $\forall i \in \mathbf{N}$, $\mathbf{N} \in \mathbf{M}$. Čia $\mathbf{P}^{\mathrm{T}} = (p_1, p_2, \ldots, p_n)$. Bazinio sprendinio elipsoidu vadinsime funkciją

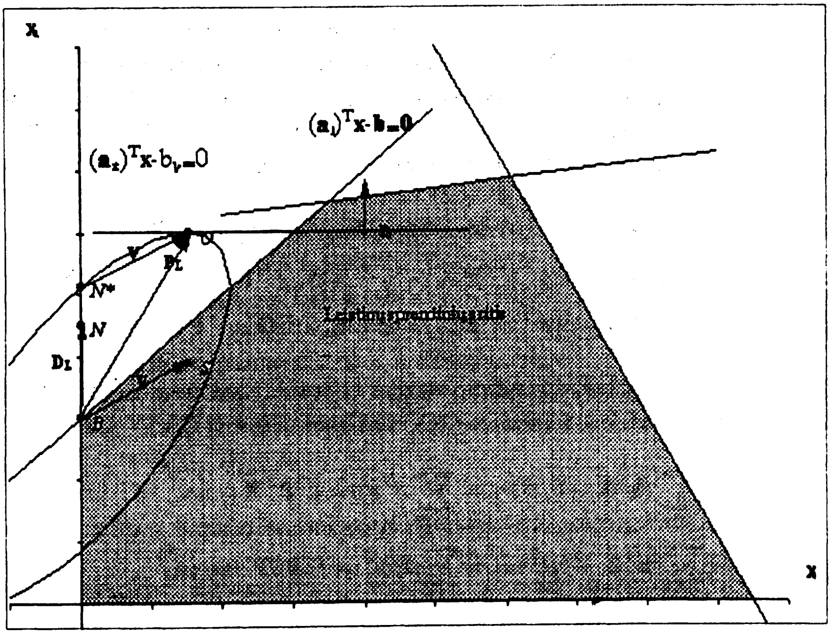
$$Q(\mathbf{X}) = \sum_{j \in N} (\mathbf{a}_j^{\mathrm{T}}\mathbf{X} - b_j)^2, \tag{4}$$

realizuojančią mažiausių kvadratų metodą visų bazinio sprendinio apribojimų plokštumų (hiperplokštumų) atžvilgiu. Nesunku įsitikinti, kad (4) – teigiamai apibrėžta kvadratinė forma, kurios $\nabla Q(\mathbf{X}) = \mathbf{A}\mathbf{A}^{\mathrm{T}}\mathbf{X} - \mathbf{A}\mathbf{B}$. Elipsoido centras $X = \mathbf{A}^{-1}\mathbf{B}$ – lygties $\nabla Q(\mathbf{X}) = 0$ sprendinys – sutampa su baziniu sprendiniu. Tegu $F_1 = \mathbf{C}^{\mathrm{T}}\mathbf{X} > F_B$. Tada uždavinio $Q(\mathbf{X}) + \lambda(\mathbf{C}^{\mathrm{T}}\mathbf{X} - F_1) \to$ min sprendinys (λ-Lagranžo daugiklis, $\lambda < 0$) – taško O, kuriame plokštuma F_1 liečia nenulinio tūrio elipsoidą $Q(\mathbf{X})$ yra:

$$X_O = \mathbf{X}_B \text{-} \lambda(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{A}^{-1}\mathbf{C} = \mathbf{X}_B + \beta(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P},\ \beta = -\lambda > 0.$$

Iš kitos pusės, X_O – lygties $\nabla Q(\mathbf{X}) = \beta\mathbf{C}$ sprendinys, t.y. spindulyje

$$X_O = \mathbf{X}_B + \beta(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P}, \quad \beta > 0$$



1 pav. Optimizavimo krypties vektoriaus V apskaičiavimo iliustracija.

elipsoido gradientas kolinearus tikslo funkcijos (1) vektoriui **C**. Vektorių $P_L = (\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P}$ vadinsime optimalaus sprendinio lietimo vektoriumi (1 pav.). Bazinis sprendinys optimalus, jei taške X_O tenkinama sąlyga $\mathbf{A}^{\mathrm{T}}\mathbf{X} - \mathbf{B} \geqslant 0$. Iš tikrųjų $(\mathbf{A}^{\mathrm{T}}\mathbf{X} - \mathbf{B})_O = \mathbf{A}^{\mathrm{T}}(\mathbf{X}_B + \beta(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P}) - \mathbf{B} = \mathbf{A}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P} + \beta\mathbf{A}^{\mathrm{T}}(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P}) - \mathbf{B} = \beta\,\mathbf{P} \geqslant 0$. T. y. tenkinama (3) sąlyga. Sąlyga $\mathbf{A}^{\mathrm{T}}\mathbf{X} - \mathbf{B} \geqslant 0$ reiškia, kad bet kuriame spindulio

$$X_O = \mathbf{X}_B + \beta(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P}, \quad \beta > 0$$

taške nė vienas apribojimas nėra tenkinamas griežtai. Tašką O vadinsime optimalaus sprendinio vaizdu elipsoide. Koks taškas būtų neoptimalaus sprendinio vaizdas elipsoide? Tarkime, kad k-asis ($k \in N$) bazinio sprendinio apribojimas tenkinamas griežtai, t.y. $\mathbf{a}_k^{\mathrm{T}} - b_k < 0$. Netenkinama sąlyga $\mathbf{A}^{\mathrm{T}}\mathbf{X} - \mathbf{B} \geqslant 0$, todėl bazinis sprendinys neoptimalus. Naudosime funkciją

$$Q(\mathbf{X})_{-k} = \sum_{\substack{j \in N \\ j \neq k}} (\mathbf{a}_j^{\mathrm{T}}\mathbf{X} - b_j)^2,$$

sudarytą tik iš griežtai netenkinamų apribojimų. $\nabla Q(\mathbf{X})_{-k} = \mathbf{A}_{-k}(\mathbf{A}^{\mathrm{T}})_{-k}\mathbf{X} - \mathbf{A}_{-k}\mathbf{B}_{-k}$.

Čia $\mathbf{A}_{-k} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_k = 0, \ldots, \mathbf{a}_n)$ – matrica **A**, kurioje vektorius $\mathbf{a}_k$ pakeistas nuliniu vektoriumi. $\mathbf{B}_{-k} = (b_1, b_2, \ldots, b_k = 0, \ldots, b_n)$ – vektorius B su k-ąja nuline komponente. Tada $(\nabla Q(\mathbf{X})_{-k})_O = \beta\mathbf{A}_{-k}\mathbf{P}_{-k}$. $(\mathbf{P}_{-k})^{\mathrm{T}} = (p_1, p_2, \ldots, p_k = 0, \ldots, p_n)$ – vektorius, kurio k-asis elementas nulinis. Vektorius $\beta\mathbf{A}_{-k}\mathbf{P}_{-k}$ – neoptimalų sprendinį charakterizuojančios plokštumos D vektorius. $\mathbf{X}_N$ – elipsoido $Q(\mathbf{X})$ lietimosi su plokštuma D taško **N** koordinačių vektorius: $X_N = \mathbf{X}_B + \beta(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P}_{-k}$. Tašką **N** vadinsime neoptimalaus sprendinio vaizdu elipsoide, o vektorių $\mathbf{D}_L = \beta(\mathbf{A}^{\mathrm{T}})^{-k}\mathbf{P}_{-k}$ – neoptimalaus sprendinio lietimo vektoriumi.

Apskaičiuosime funkcijos $Q(\mathbf{X})$ reikšmes taškuose O ir **N**, prieš tai apskaičiuodami atskirus funkcijos dėmenis. Tam įvesime pažymėjimą $(\mathbf{O}_{+j})^{\mathrm{T}} = (0, 0, \ldots, o_j = 1, \ldots, 0)$. T.y. $(\mathbf{O}_{+j})^{\mathrm{T}}$ – beveik nulinis vektorius su vienetine j-ąja komponente. Tada

$$
\begin{aligned}
(\mathbf{a}_j^{\mathrm{T}}\mathbf{X}-b_j)_O &= \mathbf{a}_j^{\mathrm{T}}(\mathbf{X}_B+\beta(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P})-b_j = (\mathbf{O}_{+j})^{\mathrm{T}}\mathbf{B}+\beta(\mathbf{O}_{+j})^{\mathrm{T}}\mathbf{P} - b_j = \beta p_j;\\
(\mathbf{a}_j^{\mathrm{T}}\mathbf{X}-b_j)_N &= \mathbf{a}_j^{\mathrm{T}}(\mathbf{X}_B+\beta(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P}_{-k})-b_j = \beta(\mathbf{O}_{+j})^{\mathrm{T}}\mathbf{P}_{-k} = \beta p_j,\ \forall j\in\mathbf{N},\ j\neq k;\\
(\mathbf{a}_k^{\mathrm{T}}\mathbf{X}-b_k)_N &= \mathbf{a}_k^{\mathrm{T}}(\mathbf{X}_B+\beta(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P}_{-k})-b_k = b_k+\beta(\mathbf{O}_{+j})^{\mathrm{T}}\mathbf{P}_{-k}-b_k = 0.
\end{aligned}
$$

Išvada: taškas **N** (neoptimalaus sprendinio vaizdas elipsoide) priklauso visoms griežtai tenkinamų apribojimų plokštumoms. Funkcijos $Q(\mathbf{X})$ reikšmes taškuose O ir N yra:

$$
\begin{aligned}
(Q(\mathbf{X}))_O &= \Big(\sum_{j\in N}(\mathbf{a}_j^{\mathrm{T}}\mathbf{X} - b_j)^2\Big)_O = \sum_{j\in N}\beta^2 p_j^2 = \beta^2\mathbf{P}^{\mathrm{T}}\mathbf{P} = \eta_O;\\
Q(\mathbf{X}))_N &= \Big(\sum_{\substack{j\in N\\ j\neq k}}(\mathbf{a}_j^{\mathrm{T}}\mathbf{X} - b_j)^2\Big)_N = \sum_{\substack{j\in N\\ j\neq k}}\beta^2 p_j^2 = \beta^2\mathbf{PTP} = \eta_N.
\end{aligned}
$$

Akivaizdu, kad $(Q(\mathbf{X}))_O > (Q(\mathbf{X}))_N$, t.y. taškas **N** priklauso mažesnio tūrio elipsoidui. Vektoriui $\beta(\mathbf{A}^{\mathrm{T}})^{-1}\mathbf{P}_{-k}$ parinkus daugiklį $\alpha > 0$, galima pasiekti, kad $(Q(\mathbf{X}))_O = (Q(\mathbf{X}))_N$. Gautume $\alpha^2 = (\eta_O/\eta_N) > 1$, kai $\eta_N \neq 0$.

Sudarysime naujos optimizavimo krypties vektorių $\mathbf{V} = \mathbf{P}_L - \alpha\mathbf{D}_L$, t.y. vektorių iš neoptimalaus į optimalų sprendinio vaizdą elipsoide. Galima įrodyti, kad vektorius $\mathbf{V} = \mathbf{P}_L - \alpha\mathbf{D}_L$, kai $\alpha^2 \in (1, \eta_O/\eta_N)$, tenkina sąlygas

$$\mathbf{C}^{\mathrm{T}}\mathbf{V} > 0 \tag{5}$$

$$\mathbf{a}_j^{\mathrm{T}}\mathbf{V} < 0, \quad j \in \mathbf{N} \tag{6}$$

t.y. garantuoja tikslo funkcijos (1) didėjimą, o taškas S ($X_S = X_B + \varepsilon\mathbf{V}$, su kiek norima mažu $\varepsilon > 0$) priklauso leistinų sprendinių sričiai.

Sprendimo metodas

Tegu $\mathbf{X}_B$ – leistinas neoptimalus bazinis uždavinio (1)–(2) sprendinys. Pradedant sudaryti naujo sprendinio bazę $\mathbf{A} = \mathbf{A}^{-1} = \mathbf{I}_n$, t.y. naujosios bazės matricos vienetinės. Naudosime vektorių **U**, parodantį naujos bazės matricų **A** ir $\mathbf{A}^{-1}$ užimtumą apribojimų vektoriais. Jei $u_k = 0$, tai k-tieji **A** ir $\mathbf{A}^{-1}$ stulpeliai laisvi, t.y. $\mathbf{a}_k = (0, 0, \ldots, a_k = 1, \ldots, 0)$. Jei $u_k = r$ ($r \in \mathbf{M}$), tai k-asis **A** stulpelis užimtas ir $\mathbf{a}_k = (a_{r1}, a_{r2}, \ldots, a_{rn})$. Pradedant sudaryti naujo sprendinio bazę $\mathbf{U} = 0$. Tegu $\mathbf{X}^{(z)}$ – sprendinys po z naujos bazės sudarymo žingsnių. $\mathbf{X}^{(0)} = \mathbf{X}_B$.

$z = 1$. Spidulyje $\mathbf{X} = \mathbf{X}^{(0)} + \mathbf{V}^{(0)*}t$, $t > 0$ ieškomas artimiausias apribojimas Jeigu $(b_j - \mathbf{a}_j^{\mathrm{T}}\mathbf{X}^{(0)})/\mathbf{a}_j^{\mathrm{T}}\mathbf{V} < 0$, $\forall j \in \mathbf{M}$, $j \notin \mathbf{N}^S$ ($\mathbf{N}^S$ – senosios bazės indeksų aibė), tai leistinų sprendinių sritis neapribota ir, vektoriui $\mathbf{V}$ tenkinant (5) sąlygą, tikslo funkcijos reikšmė neaprėžtai didėja. Tarkime, kad k-asis apribojimas – artimiausias, t.y. $t_k = t_{\min} > 0$. Tada $\mathbf{X}^{(1)} = \mathbf{X}^{(0)} + \mathbf{V}^* t_{\min}$ – leistinas nebazinis sprendinys, nes $\mathbf{a}_k \mathrm{T}\mathbf{X}^{(1)} - b_k = 0$ ir $\mathbf{a}_j^{\mathrm{T}}\mathbf{X}^{(1)} - b_j < 0,\ j \in \mathbf{M},\ j \neq k$. Tai seka iš (6) sąlygos. Be to $F^{(1)} = F^{(0)} + \mathbf{C}^{\mathrm{T}}\mathbf{V}^* t_{\min} > F^{(0)}$. Tai seka iš (6) sąlygos.

$\mathrm{d}^{\mathrm{T}} = \mathbf{A}^{-1}\mathbf{a}_k = (d_1, d_2, \ldots, d_n)$. Atsiras bent vienas $d_i \neq 0,\ i \in \mathbf{N}$. Priešingu atveju $\mathbf{a}_k = 0$, t.y. rastas perteklinis apribojimas (nes $\mathbf{OX} - b_k \leqslant 0$), kurį galima pašalinti iš apribojimų sistemos (2). Yra dvi apribojimo vektoriaus įvedimo į bazę galimybės: 1) pagal $\max(di)$, t.y. pasirenkant tokį laisvą matricos $\mathbf{A}$ stulpelį, į kurį įvedant minimizuojama $\mathbf{A}^{-1}$ skaičiavimo paklaida; 2) pagal $\max(c_i \frac{bi}{aki})$, t.y. laisvą matricos $\mathbf{A}$ stulpelį, į kurį įvedant maksimizuojamas tikslo funkcijos pokytis.

Tarkime, kad pasirinktas pirmas matricos $\mathbf{A}$ stulpelis. Tada

$$(A^{-1})^{(1)} = \begin{pmatrix} (A_n^{-1})^{(1)} & 0 \\ -a_{k-1} \cdot A_{(k)}^{-1} & I_{n-1} \end{pmatrix};$$

čia $(A_n^{-1})^{(1)} = a/a_{k1}$ – bazinio minoro matricos atvirkštinė matrica.

$$(-a_k)_{-1} \cdot (A_M^{-1})^{(1)} = (-a_{k2}/a_{k1},\ -a_{k3}/a_{k1}, \ldots, -a_{k2}/a_{k1}).$$

$z = 2, \ldots, n$. Siekiant išlaikyti sprendinį leistinų sprendinių srityje, nauja optimalaus sprendinio paieškos kryptis apskaičiuojama, tikslo funkcijos vektorių $\mathbf{C}$ projektuojant į $k = (n - z + 1)$-matę briauną, sudarytą iš z į naują bazę įvestų apribojimų vektorių.

Po z žingsnių nebazinio sprendinio atvirkštinė matrica $(A^{-1})^{(z)}$ – projektyvinė matrica, gaunama palaipsniui užpildant matricą $\mathbf{A}$ į naują bazę įtraukiamais stulpeliais, turi tokią struktūrą:

$$(A^{-1})^{(z)} = \begin{pmatrix} (A_n^{-1})^{(z)} & 0 \\ -a_{k-1} \cdot A_{(k)}^{-1} & I_{n-1} \end{pmatrix}. \tag{7}$$

Galima įrodyti, kad po z žingsnių gaunamas optimizavimo krypties vektorius $\mathbf{V}^{(z)} = \mathbf{C}^{\mathrm{T}}(\mathbf{O}_k, (A^{-1})^{(z)}\mathbf{C})$ patenkina (5) ir (6) sąlygas.

Trumpas uždavinio sprendimo elipsoidų metodu algoritmas:

1) jei $\mathbf{V} = 0$, toliau 4);

2) spindulyje $\mathbf{X} = \mathbf{X}^{(z)} + \mathbf{V}^{(z)*}t,\ t > 0$ ieškoma artimiausio apribojimo;

3) vektorius $\mathbf{V}$ apskaičiuojamas:

   a) naudojant bazinio sprendinio elipsoidą, jei $z = 1$;

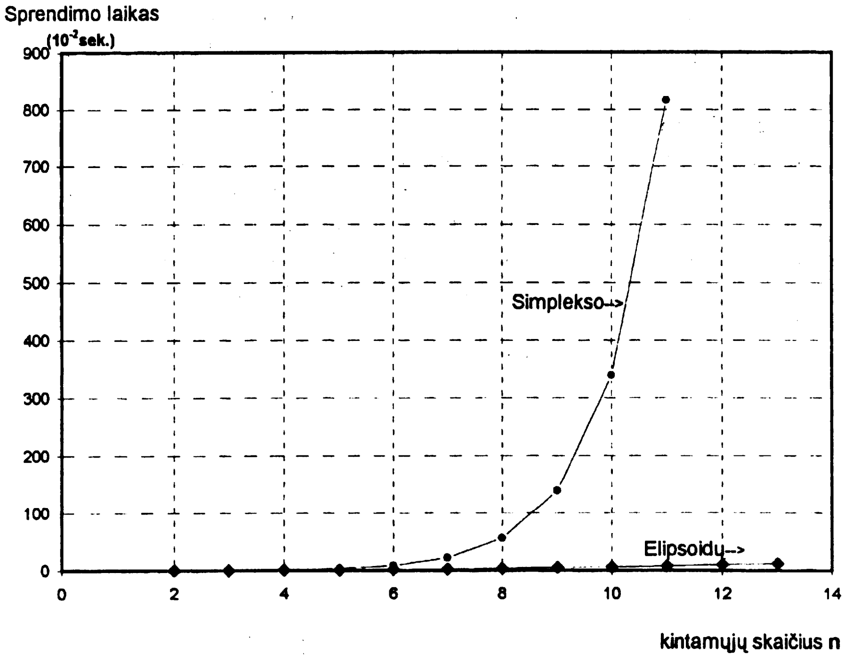
   b) naudojant nebazinio sprendinio projektyvinę matricą $\mathbf{A}^{-1}$, jei $1 < z \leqslant n$;

   c) toliau 1);

4) pabaiga.

Metodų palyginimas

Užprogramuotas elipsoidų metodas lygintas su simplekso metodu reikalingo uždaviniui sprendimo laiko požiūriu. Metodai lyginti, sprendžiant Klee-Minty uždavinį.



2 pav. Klee-Minty uždavinio sprendimo laikas elipsoidų ir simplekso metodu.

Pastebėtina, kad Klee-Minty uždavinys, o tuo pačiu ir sprendimo laikas, pilnai apibūdinamas kintamųjų skaičiumi n, su kuriuo susijęs apribojimų skaičius (visada $2n$) bei uždavinio koeficientų 5^n eilės reikšmėmis. Nenaudojant dvigubo tikslumo aritmetikos, elipsoidų metodu pavyko išspręsti Klee-Minty uždavinius visiems $n \leqslant 25$, t.y. kol uždavinio koeficientai pasiekė 5^{25} eilės reikšmes.

Literatūra

[1] Klee V., Minty G.J., How good is the simplex algorithm?, *Inequalities*, **111**, O.Shisha ed., Academic Press, N.Y. (1972).

Comparison of the ellipsoid and simplex methods

E. Laurinavičius, D. Raškinienė

The ellipsoid method is presented. The least squers functional of the basic solution linear constrains is used as ellipsoid. The Klee-Minty linear programming problem is used to compare the ellipsoid and simplex methods.