FIT ČVUT

MI-LOM

Lineární optimalizace a metody

Výpočetní složitost lineárního programování: elipsoidový algoritmus a Chačijanova věta

Evropský sociální fond
Praha & EU: Investujeme do vaší budoucnosti

Michal Černý, 2011

Výpočetní složitost lineárního programování: elipsoidový algoritmus a Chačijanova věta

Evropský sociální fond
Praha & EU: Investujeme do vaší budoucnosti

Označme $\mathbf{LP} := \{[\boldsymbol{A}, \boldsymbol{b}] : \boldsymbol{A}$ je racionální matice, $\boldsymbol{b}$ je racionální vektor a $\{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\} \neq \emptyset\}$. V této přednášce dokážeme ústřední větu lineárního programování:

Věta 1 (Chačijan, 1979)**.** $\mathbf{LP} \in \boldsymbol{P}$. □

Z přednášky o dualitě víme, že to má důležitý důsledek:

Důsledek 2. *Optimalisační verse úlohy lineárního programování je řešitelná v polynomiálním čase.* □

Konvence. V celém textu bude n označovat dimensi ($\boldsymbol{x} \in \mathbb{R}^n$) a m počet nerovnic v systému $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$.

Bez újmy na obecnosti můžeme předpokládat, že platí:

(i) je-li polyedr $\mathcal{F} := \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$ neprázdný, má vrchol;

(ii) matice $\boldsymbol{A}$ i vektor $\boldsymbol{b}$ jsou celočíselné;

(iii) $n \geqslant 2$.

Jako cvičení zdůvodněte, proč jsou uvedené předpoklady „bez újmy na obecnosti“.

1 Elipsoidy a E-norma

V této části shrneme známá fakta o positivně definitních maticích a elipsoidech; často si dovolíme se odvolat jen na geometrický názor.

Připomeňme, že symetrická matice $\boldsymbol{E}$ je *positivně semidefinitní*, jestliže pro každé $\boldsymbol{x}$ jest $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x} \geqslant 0$, a je *positivně definitní*, jestliže navíc platí, že rovnost nastává jen pro $\boldsymbol{x} = \boldsymbol{0}$. Následující lemma shrnuje základní vlastnosti positivně definitních matic, které jsou dobře známy z lineární algebry. Připomeňme, že vlastní čísla symetrických matic jsou reálná.

Lemma 3 (o positivní definitnosti)**.** *(a) Positivně definitní matice je regulární.*

(b) Inverse positivně definitní matice je positivně definitní.

(c) Symetrická matice je positivně semidefinitní, právě když má nezáporná vlastní čísla, a je positivně definitní, právě když má kladná vlastní čísla.

(d) K positivně definitní matici $\boldsymbol{E}$ existuje jediná positivně definitní matice $\boldsymbol{E}^{1/2}$ taková, že $(\boldsymbol{E}^{1/2})^2 = \boldsymbol{E}$.

(e) $(\boldsymbol{E}^{1/2})^{-1} = (\boldsymbol{E}^{-1})^{1/2}$ $(=: \boldsymbol{E}^{-1/2})$.

(f) Jsou-li $\lambda_1, \ldots, \lambda_n$ vlastní čísla positivně definitní matice $\boldsymbol{E}$ (rozměru $n \times n$), pak $\lambda_1^x, \ldots, \lambda_n^x$ jsou vlastní čísla matice $\boldsymbol{E}^x$ pro $x \in \{-2, -1, -\frac{1}{2}, \frac{1}{2}, 1, 2\}$. □

Z (c) plyne, že positivně definitní matice mají kladný determinant; determinant je totiž součin vlastních čísel.

Geometrický význam positivně definitních matic je patrný z následující definice.

Definice 4. ***Elipsoid*** *určený středem $\boldsymbol{z}$ a positivně definitní maticí $\boldsymbol{E}$ je množina*

$$\mathcal{E}(\boldsymbol{z}, \boldsymbol{E}) := \{\boldsymbol{x} : \ (\boldsymbol{x} - \boldsymbol{z})^{\mathrm{T}} \boldsymbol{E}^{-1} (\boldsymbol{x} - \boldsymbol{z}) \leqslant 1\}.$$

Hranicí *elipsoidu $\mathcal{E}(\boldsymbol{z}, \boldsymbol{E})$ rozumíme množinu $\{\boldsymbol{x} : \ (\boldsymbol{x} - \boldsymbol{z})^{\mathrm{T}} \boldsymbol{E}^{-1} (\boldsymbol{x} - \boldsymbol{z}) = 1\}$.* □

Nadále v této kapitole značí $\boldsymbol{E}$ positivně definitní matici.

Ze školy si každý patrně vybaví rovnici elipsy v rovině

$$\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} \leqslant 1, \tag{1}$$

kde $a_1 > 0$ a $a_2 > 0$ jsou délky poloos; to je speciální případ definice 4 s $\boldsymbol{E} = \left(\begin{smallmatrix} a_1^2 & 0 \\ 0 & a_2^2 \end{smallmatrix}\right)$. Elipsoidy s diagonální maticí $\boldsymbol{E}$ jsou elipsoidy, jejichž poloosy jsou rovnoběžné s osami souřadného systému.

Na elipsoid $\mathcal{E}(\boldsymbol{0}, \boldsymbol{E})$ se můžeme také dívat jako na obraz jednotkové koule $K := \mathcal{E}(\boldsymbol{0}, \boldsymbol{I}) = \{\boldsymbol{x} : \|\boldsymbol{x}\| \leqslant 1\}$ ($\|\boldsymbol{x}\| := \sqrt{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}}$ značí eukleidovskou normu) při zobrazení $f : \boldsymbol{x} \mapsto \boldsymbol{E}^{1/2} \boldsymbol{x}$. Platí totiž

$$\begin{aligned}
\boldsymbol{x} \in K &\longleftrightarrow \sqrt{\boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}} \leqslant 1 \\
&\longleftrightarrow \boldsymbol{x}^{\mathrm{T}} \boldsymbol{x} \leqslant 1 \\
&\longleftrightarrow \boldsymbol{x}^{\mathrm{T}} \boldsymbol{E}^{1/2} \boldsymbol{E}^{-1} \boldsymbol{E}^{1/2} \boldsymbol{x} \leqslant 1 \\
&\longleftrightarrow (\boldsymbol{E}^{1/2} \boldsymbol{x})^{\mathrm{T}} \boldsymbol{E}^{-1} (\boldsymbol{E}^{1/2} \boldsymbol{x}) \leqslant 1 \\
&\longleftrightarrow \boldsymbol{E}^{1/2} \boldsymbol{x} \in \mathcal{E}(\boldsymbol{0}, \boldsymbol{E}) \\
&\longleftrightarrow f(\boldsymbol{x}) \in \mathcal{E}(\boldsymbol{0}, \boldsymbol{E});
\end{aligned} \tag{2}$$

odtud

$$\mathcal{E}(\boldsymbol{0}, \boldsymbol{E}) = \{\boldsymbol{E}^{1/2} \boldsymbol{x} : \|\boldsymbol{x}\| \leqslant 1\}.$$

V řadě ekvivalencí (2) můžeme nerovnost „$\leqslant$“ nahradit rovností; pak obdržíme tvrzení

$$\mathit{hranice}(\mathcal{E}(\boldsymbol{0}, \boldsymbol{E})) = \{\boldsymbol{E}^{1/2} \boldsymbol{x} : \|\boldsymbol{x}\| = 1\}.$$

Jinými slovy: jestliže $\boldsymbol{x}$ probíhá povrch jednotkové koule, pak $\boldsymbol{E}^{1/2} \boldsymbol{x}$ probíhá povrch elipsoidu. Odtud plyne

$$\begin{aligned}
\text{délka nejdelší poloosy } \mathcal{E}(\boldsymbol{0}, \boldsymbol{E}) &= \max\{\|\boldsymbol{E}^{1/2} \boldsymbol{x}\| : \|\boldsymbol{x}\| = 1\}, \\
\text{délka nejkratší poloosy } \mathcal{E}(\boldsymbol{0}, \boldsymbol{E}) &= \min\{\|\boldsymbol{E}^{1/2} \boldsymbol{x}\| : \|\boldsymbol{x}\| = 1\}.
\end{aligned}$$

Nechá se ukázat, že *vlastní čísla matice $\boldsymbol{E}^{1/2}$ jsou délky poloos elipsy $\mathcal{E}(\boldsymbol{0}, \boldsymbol{E})$ a vlastní vektory matice $\boldsymbol{E}^{1/2}$ jsou směry poloos elipsy $\mathcal{E}(\boldsymbol{0}, \boldsymbol{E})$*. (Zůstáváme u intuitivního chápání významu termínu „poloosa elipsoidu“, protože jsme elipsoid nezaváděli geometricky.)

Symbolem $\Lambda(\boldsymbol{X})$ značíme největší vlastní číslo (symetrické) matice $\boldsymbol{X}$ a symbolem $\lambda(\boldsymbol{X})$ značíme nejmenší vlastní číslo $\boldsymbol{X}$. Všimněme si, že z lemmatu 3(f) plyne

$$\Lambda(\boldsymbol{E}) = \frac{1}{\lambda(\boldsymbol{E}^{-1})}, \quad \lambda(\boldsymbol{E}) = \frac{1}{\Lambda(\boldsymbol{E}^{-1})}; \tag{3}$$

tento vztah budeme často využívat.

Díky vztahu mezi délkami poloos a vlastními čísly můžeme psát

$$\sqrt{\Lambda(\boldsymbol{E})} = \Lambda(\boldsymbol{E}^{1/2}) = \max\{\|\boldsymbol{E}^{1/2}\boldsymbol{x}\| : \|\boldsymbol{x}\| = 1\},$$
$$\sqrt{\lambda(\boldsymbol{E})} = \lambda(\boldsymbol{E}^{1/2}) = \min\{\|\boldsymbol{E}^{1/2}\boldsymbol{x}\| : \|\boldsymbol{x}\| = 1\},$$

nebo ekvivalentně:

Lemma 5. $\Lambda(\boldsymbol{E}) = \max\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x} : \|\boldsymbol{x}\| = 1\}$ a $\lambda(\boldsymbol{E}) = \min\{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x} : \|\boldsymbol{x}\| = 1\}$. □

Důsledek 6. *Jestliže* $\|\boldsymbol{x}\| = 1$, *pak* $\lambda(\boldsymbol{E}) \leqslant \boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x} \leqslant \Lambda(\boldsymbol{E})$. □

Pro libovolný vektor $\boldsymbol{x} \neq \boldsymbol{0}$ dosazením $\boldsymbol{x} := \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ obdržíme vztah

$$\|\boldsymbol{x}\| \cdot \sqrt{\lambda(\boldsymbol{E})} \leqslant \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x}} \leqslant \|\boldsymbol{x}\| \cdot \sqrt{\Lambda(\boldsymbol{E})}, \tag{4}$$

který bude užitečný později.

Definice 7. ***E*-normou** *vektoru* $\boldsymbol{x}$ *rozumíme číslo* $\|\boldsymbol{x}\|_{\boldsymbol{E}} := \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{x}}$. □

Snadno se vidí, že $\|\cdot\|_{\boldsymbol{E}}$ je opravdu norma na $\mathbb{R}^n$.

Lemma 8 (o $\boldsymbol{E}$-normě)**.** *(a) Jednotkovou koulí* $\boldsymbol{E}$*-normy je elipsa* $\mathcal{E}(\boldsymbol{0}, \boldsymbol{E})$. *Tedy:* $\|\boldsymbol{x}\|_{\boldsymbol{E}} \leqslant 1$, *právě když* $\boldsymbol{x} \in \mathcal{E}(\boldsymbol{0}, \boldsymbol{E})$, *a* $\|\boldsymbol{x}\|_{\boldsymbol{E}} = 1$, *právě když* $\boldsymbol{x} \in$ *hranice*$(\mathcal{E}(\boldsymbol{0}, \boldsymbol{E}))$.

(b) $\|\boldsymbol{x}\|_{\boldsymbol{E}} = \|\boldsymbol{E}^{-1/2}\boldsymbol{x}\|$.

(c) $\frac{\|\boldsymbol{x}\|}{\sqrt{\Lambda(\boldsymbol{E})}} \leqslant \|\boldsymbol{x}\|_{\boldsymbol{E}} \leqslant \frac{\|\boldsymbol{x}\|}{\sqrt{\lambda(\boldsymbol{E})}}$.

(d) Pro $\alpha > 0$ *jest* $\|\boldsymbol{x}\|_{\alpha\boldsymbol{E}} = \frac{1}{\sqrt{\alpha}}\|\boldsymbol{x}\|_{\boldsymbol{E}}$.

Důkaz. (a) $\|\boldsymbol{x}\|_{\boldsymbol{E}} = 1$, právě když $\sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{x}} = 1$, právě když $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{x} = 1$, právě když $\boldsymbol{x} \in$ *hranice*$(\mathcal{E}(\boldsymbol{0}, \boldsymbol{E}))$ (poslední ekvivalence plyne z definice hranice elipsoidu).

(b) $\|\boldsymbol{x}\|_{\boldsymbol{E}} = \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{x}} = \sqrt{(\boldsymbol{E}^{-1/2}\boldsymbol{x})^{\mathrm{T}}\boldsymbol{E}^{-1/2}\boldsymbol{x}} = \|\boldsymbol{E}^{-1/2}\boldsymbol{x}\|$.

(c) Užijme (4) s $\boldsymbol{E} := \boldsymbol{E}^{-1}$; obdržíme $\|\boldsymbol{x}\| \cdot \sqrt{\lambda(\boldsymbol{E}^{-1})} \leqslant \|\boldsymbol{x}\|_{\boldsymbol{E}} \leqslant \|\boldsymbol{x}\| \cdot \sqrt{\Lambda(\boldsymbol{E}^{-1})}$. Tvrzení (c) plyne z (3).

(d) $\|\boldsymbol{x}\|_{\alpha\boldsymbol{E}} = \sqrt{\boldsymbol{x}^{\mathrm{T}}(\alpha\boldsymbol{E})^{-1}\boldsymbol{x}} = \frac{1}{\sqrt{\alpha}} \cdot \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}^{-1}\boldsymbol{x}} = \frac{1}{\sqrt{\alpha}}\|\boldsymbol{x}\|_{\boldsymbol{E}}$. □

Tvrzení (b) lemmatu říká, že $\boldsymbol{E}$-norma vektoru $\boldsymbol{x}$ je totéž co jeho eukleidovská délka, zobrazíme-li prostor pomocí $\boldsymbol{\xi} \mapsto \boldsymbol{E}^{-1/2}\boldsymbol{\xi}$ (to jest, zobrazíme-li elipsoid $\mathcal{E}(\boldsymbol{0}, \boldsymbol{E})$ na jednotkovou kouli).

Konečne si všimněme, že objem elipsoidu je úměrný součinu délek poloos. Délky poloos jsou vlastní čísla matice $\boldsymbol{E}^{1/2}$. A součin vlastních čísel je determinant. Označíme-li objem symbolem *vol*, můžeme říci:

Lemma 9. *V každé dimensi n existuje konstanta k_n taková, že pro libovolný elipsoid $E := \mathcal{E}(\boldsymbol{z}, \boldsymbol{E}) \subseteq \mathbb{R}^n$ platí $vol(E) = k_n \cdot \sqrt{\det \boldsymbol{E}}$.* □

Nakonec ještě připomeňme, že stopa matice je součet diagonálních prvků. Platí pozoruhodná věta, že stopa je rovna součtu vlastních čísel. Protože positivně (semi)definitní matice má nezáporná vlastní čísla, platí, že *stopa je horním odhadem na její největší vlastní číslo.*

2 Elipsoidové řezy

Představme si, že máme dán elipsoid E a rozřízneme jej (libovolnou) nadrovinou, která prochází jeho středem, na dvě části; jednu zahodíme. Hledáme nový elipsoid E', který obsahuje celou nezahozenou polovinu rozříznutého elipsoidu, a má co nejmenší objem. Střed a matici určující elipsoid E' lze vyjádřit explicitním vzorečkem.

Lemma 10 (o centrálních elipsoidových řezech)**.** *Nechť je dán elipsoid $E = \mathcal{E}(\boldsymbol{z}, \boldsymbol{E})$ v $\mathbb{R}^n$ a vektor $\boldsymbol{a} \neq \boldsymbol{0}$. Nejmenší (co do objemu) elipsoid $E' = \mathcal{E}(\boldsymbol{z}', \boldsymbol{E}')$, který obsahuje množinu*

$$E \cap \{(\boldsymbol{a} - \boldsymbol{z})^{\mathrm{T}} \boldsymbol{x} \leqslant 0\},$$

má střed v bodě

$$\boldsymbol{z}' = \boldsymbol{z} - \frac{1}{n+1} \cdot \frac{\boldsymbol{E}\boldsymbol{a}}{\sqrt{\boldsymbol{a}^{\mathrm{T}} \boldsymbol{E} \boldsymbol{a}}}$$

a je určen maticí

$$\boldsymbol{E}' = \frac{n^2}{n^2 - 1} \cdot \left(\boldsymbol{E} - \frac{2}{n+1} \frac{\boldsymbol{E}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{a}} \right).$$

Důkaz. Označme $\boldsymbol{e}_i := (0, \ldots, 0, 1, 0, \ldots, 0)^{\mathrm{T}}$, kde jednička je na i-tém místě. Hovoříme-li o vektoru $\boldsymbol{x}$, značíme jeho složky $x_1, x_2, \ldots, x_n$.

Můžeme předpokládat, že $\boldsymbol{z} = \boldsymbol{0}$. Buď $\boldsymbol{D} := \boldsymbol{E}^{1/2}$. Elipsoid $E = \mathcal{E}(\boldsymbol{0}, \boldsymbol{E})$ vznikl zobrazením $D : \boldsymbol{\xi} \mapsto \boldsymbol{D}\boldsymbol{\xi}$ z jednotkové koule. Myšlenka důkazu je tato: při zobrazení $D^{-1} : \boldsymbol{\xi} \mapsto \boldsymbol{D}^{-1}\boldsymbol{\xi}$ se E zobrazí na jednotkovou kouli. Na jednotkové kouli provedeme propočet a pak vše zobrazíme zpět pomocí D.

Následující pozorování říká, jak se při zobrazení D^{-1} chová nadrovina zadaná normálovým vektorem $\boldsymbol{a}$.

Pozorování A. $D^{-1}(\{\boldsymbol{x} : \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = 0\}) = \{\boldsymbol{x} : (\boldsymbol{D}\boldsymbol{a})^{\mathrm{T}}\boldsymbol{x} = 0\}$. *Důkaz.* Pro libovolné $\boldsymbol{y}$ platí

$$\begin{aligned} \boldsymbol{y} \in \{\boldsymbol{x} : \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = 0\} \quad &\longleftrightarrow \quad \boldsymbol{a}^{\mathrm{T}}\boldsymbol{y} = 0 \quad \longleftrightarrow \quad (\boldsymbol{D}\boldsymbol{a})^{\mathrm{T}}\boldsymbol{D}^{-1}\boldsymbol{y} = 0 \\ &\longleftrightarrow \quad \boldsymbol{D}^{-1}\boldsymbol{y} \in \{\boldsymbol{x} : (\boldsymbol{D}\boldsymbol{a})^{\mathrm{T}}\boldsymbol{x} = 0\} \quad \longleftrightarrow \quad D^{-1}(\boldsymbol{y}) \in \{\boldsymbol{x} : (\boldsymbol{D}\boldsymbol{a})^{\mathrm{T}}\boldsymbol{x} = 0\}. \end{aligned}$$

Tím je pozorování dokončeno.

Položme

$$\boldsymbol{\theta} := \frac{\boldsymbol{D}\boldsymbol{a}}{\|\boldsymbol{D}\boldsymbol{a}\|}.$$

Podle pozorování A víme, že při D^{-1} se nadrovina $\{\boldsymbol{x} : \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = 0\}$ zobrazí na nadrovinu $\{\boldsymbol{x} : (\boldsymbol{D}\boldsymbol{a})^{\mathrm{T}}\boldsymbol{x} = 0\} = \{\boldsymbol{x} : \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{x} = 0\}$; z toho plyne, že při zobrazení D se nadrovina $\{\boldsymbol{x} : \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{x} = 0\}$ zobrazí na $\{\boldsymbol{x} : \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = 0\}$. Najdeme-li nejmenší (co do objemu) elipsoid, který obsahuje množinu *jednotková koule* $\cap$ $\{\boldsymbol{x} : \boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{x} \leqslant 0\}$, a tento elipsoid zobrazíme pomocí D, zíkáme nejmenší

elipsoid, který obsahuje množinu $E \cap \{\boldsymbol{x} : \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} \leqslant 0\}$. (Regulární zobrazení D mění objemy jen o konstantní násobek, „o svůj determinant“, takže vlastnost *míti minimální objem* se při zobrazení D zachovává.)

Zřejmě je $\|\boldsymbol{\theta}\| = 1$, a tudíž prostor můžeme zrotovat jistým zobrazením $R : \boldsymbol{\xi} \mapsto \boldsymbol{R}\boldsymbol{\xi}$, kde $\boldsymbol{R}$ je ortogonální matice[1], tak, že

$$\boldsymbol{R}\boldsymbol{\theta} = \boldsymbol{e}_1. \tag{6}$$

Proto jsme ostatně do definice $\boldsymbol{\theta}$ dali normalisaci $\frac{1}{\|\boldsymbol{D}\boldsymbol{a}\|}$.

Řekněme, že jsme zobrazením D^{-1} a následně zobrazením R dosáhli toho, že elipsoid E se zobrazil na jednotkovou kouli a řezná nadrovina se zobrazila na $\{\boldsymbol{x} : \boldsymbol{e}_1\boldsymbol{x} = x_1 = 0\}$. Řešíme úlohu najít elipsoid $F = \mathcal{E}(\boldsymbol{s}, \boldsymbol{F})$ s co nejmenším objemem, který obsahuje množinu $\{\boldsymbol{x} : \|\boldsymbol{x}\| \leqslant 1, x_1 \leqslant 0\}$. Z obrázku 1 se nahlédne, že hranice F obsahuje body

$$-\boldsymbol{e}_1, \pm\boldsymbol{e}_2, \pm\boldsymbol{e}_3, \ldots, \pm\boldsymbol{e}_n, \tag{7}$$

F má poloosy rovnoběžné s osami souřadného systému $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ a že jeho střed $\boldsymbol{s}$ má tvar $\boldsymbol{s} = (s, 0, 0, \ldots, 0)^{\mathrm{T}}$. Hledáme čísla $s, d_1^2, d_2^2, \ldots, d_n^2$ taková, že elipsoid s hranicí popsanou rovnicí $(\boldsymbol{x} - s\boldsymbol{e}_1)^{\mathrm{T}}\boldsymbol{F}^{-1}(\boldsymbol{x} - s\boldsymbol{e}_1) = 1$, kde

$$\boldsymbol{F} := diag\left(d_1^2, d_2^2, \ldots, d_n^2\right),$$

má co nejmenší objem, přičemž hranice obsahuje body (7). Rovnici hranice můžeme přepsat do tvaru

$$\frac{(x_1 - s)^2}{d_1^2} + \frac{x_2^2}{d_2^2} + \frac{x_3^2}{d_3^2} + \cdots + \frac{x_n^2}{d_n^2} = 1. \tag{8}$$

Dosazením bodu $-\boldsymbol{e}_1$ do (8) dostaneme rovnici $\frac{(s+1)^2}{d_1^2} = 1$, z čehož jest

$$d_1^2 = (s+1)^2, \tag{9}$$

a dosazením $\boldsymbol{e}_i$ do (8) ($i = 2, 3, \ldots, n$) obdržíme $\frac{s^2}{d_1^2} + \frac{1}{d_i^2} = \frac{s^2}{(s+1)^2} + \frac{1}{d_i^2} = 1$. Odkud se vyjádří

$$d_i^2 = \frac{(s+1)^2}{2s+1}. \tag{10}$$

Objem V elipsoidu F je minimální, právě když V^2 je minimální, a podle lemmatu 9 pro jistou konstantu k_n, jež závisí jen na dimensi n, platí

$$V^2 = k_n^2 \cdot d_1^2 d_2^2 \cdots d_n^2 = k_n^2 \cdot \frac{(s+1)^{2n}}{(2s+1)^{n-1}}.$$

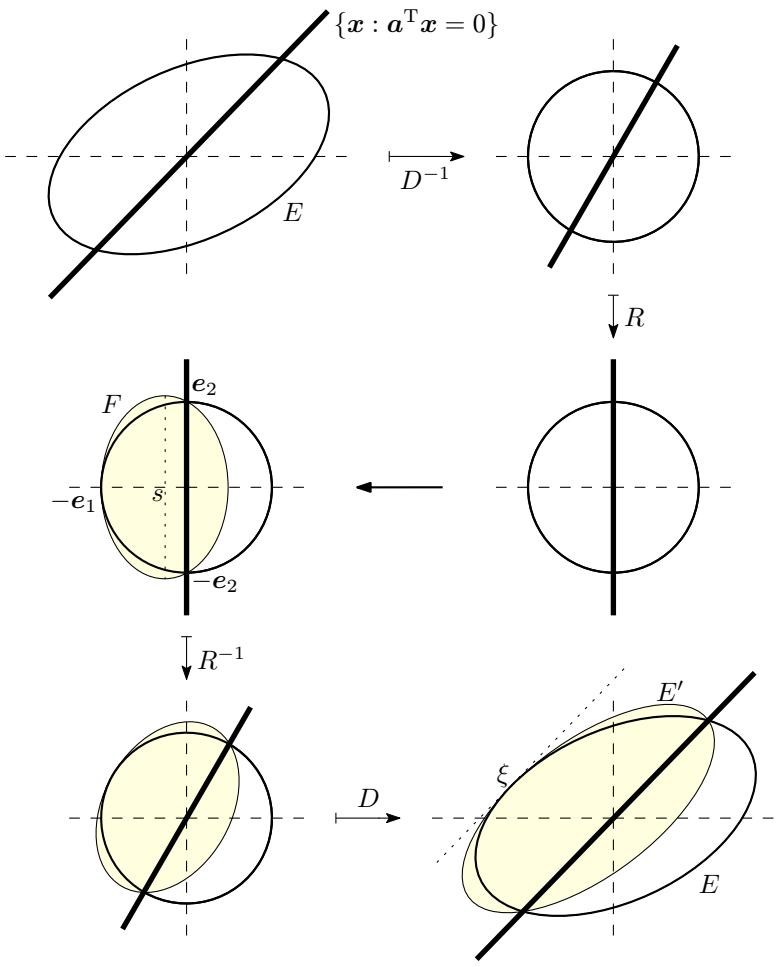
[1] *Ortogonální maticí* $\boldsymbol{R}$ rozumíme čtvercovou matici s řádky $\boldsymbol{r}_1^{\mathrm{T}}, \ldots, \boldsymbol{r}_n^{\mathrm{T}}$ splňující (a) pro $i \neq j$ je $\boldsymbol{r}_i^{\mathrm{T}}\boldsymbol{r}_j = 0$ (tedy: řádky jsou kolmé) a (b) $\|\boldsymbol{r}_i\| = 1$ pro každé i. To znamená, že řádky tvoří ortonormální basi prostoru $\mathbb{R}^n$. (Možná je čtenář zvyklý za ortogonální matici považovat libovolnou matici splňující (a), a matici splňující (a) i (b) nazývat *ortonormální*.) Z ortogonality matice $\boldsymbol{R}$ plyne $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} = \boldsymbol{I}$. Odtud

$$\boldsymbol{R}^{-1} = \boldsymbol{R}^{\mathrm{T}}. \tag{5}$$

Tento vztah bude užitečný později.

Příkladem ortogonálních matic jsou rotace; například matice, jež rotuje $\mathbb{R}^2$ o úhel θ, je tvaru

$$\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$



Obrázek 1: *Krájení elipsoidů středem.*

Konečně, s se spočte z $\frac{\mathrm{d}(V^2)}{\mathrm{d}s} = 0$, odkud vychází

$$s = -\frac{1}{n+1},$$

a z (9) a z (10) jest

$$d_i^2 = \begin{cases} \left(1 - \frac{1}{n+1}\right)^2 & = \frac{n^2}{n^2-1}\left(1 - \frac{2}{n+1}\right) \quad \text{pro } i = 1, \\ \frac{\left(1-\frac{1}{n+1}\right)^2}{1-\frac{2}{n+1}} & = \frac{n^2}{n^2-1} \quad \text{pro } i = 2, 3, \ldots, n, \end{cases}$$

takže

$$\boldsymbol{F} = \frac{n^2}{n^2-1}\left(\boldsymbol{I} - \mathit{diag}\left(\frac{2}{n+1}, 0, \ldots, 0\right)\right) = \frac{n^2}{n^2-1}\left(\boldsymbol{I} - \frac{2}{n+1}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\right)$$

a

$$\boldsymbol{s} = -\frac{1}{n+1}\boldsymbol{e}_1.$$

Nyní použijeme zobrazení $R^{-1} : \boldsymbol{\xi} \mapsto \boldsymbol{R}^{-1}\boldsymbol{\xi}$, které prostor zrotuje zpět ($\boldsymbol{e}_1 \mapsto \boldsymbol{\theta}$), a zobrazení D, kterým se od jednotkové koule přejde k původnímu elipsoidu E. Z (6) jest

$$\begin{aligned}\boldsymbol{z}' &= \boldsymbol{D}\boldsymbol{R}^{-1}\boldsymbol{s} = -\frac{1}{n+1}\boldsymbol{D}\boldsymbol{R}^{-1}\boldsymbol{e}_1 = -\frac{1}{n+1}\boldsymbol{D}\boldsymbol{\theta} \\ &= -\frac{1}{n+1}\boldsymbol{D}\frac{\boldsymbol{D}\boldsymbol{a}}{\|\boldsymbol{D}\boldsymbol{a}\|} = -\frac{1}{n+1}\frac{\boldsymbol{E}\boldsymbol{a}}{\sqrt{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{D}\boldsymbol{a}}} = -\frac{1}{n+1}\frac{\boldsymbol{E}\boldsymbol{a}}{\sqrt{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{a}}}.\end{aligned}$$

Porozování B. Při regulárním zobrazení $\Xi : \boldsymbol{\xi} \mapsto \boldsymbol{\Xi}\boldsymbol{\xi}$ se $\mathcal{E}(\mathbf{0}, \boldsymbol{\Omega})$ zobrazí na $\mathcal{E}(\mathbf{0}, \boldsymbol{\Xi}\boldsymbol{\Omega}\boldsymbol{\Xi}^{\mathrm{T}})$. Důkaz. Pro libovolné $\boldsymbol{y}$ jest

$$\begin{aligned}\boldsymbol{y} \in \mathcal{E}(\mathbf{0}, \boldsymbol{\Omega}) \quad &\longleftrightarrow \quad \boldsymbol{y}\boldsymbol{\Omega}^{-1}\boldsymbol{y} \leqslant 1 \\ &\longleftrightarrow \quad (\boldsymbol{\Xi}\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{\Xi}^{\mathrm{T}})^{-1}\boldsymbol{\Omega}^{-1}\boldsymbol{\Xi}^{-1}(\boldsymbol{\Xi}\boldsymbol{y}) \leqslant 1 \\ &\longleftrightarrow \quad (\boldsymbol{\Xi}\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{\Xi}\boldsymbol{\Omega}\boldsymbol{\Xi}^{\mathrm{T}})^{-1}(\boldsymbol{\Xi}\boldsymbol{y}) \leqslant 1 \\ &\longleftrightarrow \quad \boldsymbol{\Xi}\boldsymbol{y} \in \mathcal{E}(\mathbf{0}, \boldsymbol{\Xi}\boldsymbol{\Omega}\boldsymbol{\Xi}^{\mathrm{T}}) \\ &\longleftrightarrow \quad \Xi(\boldsymbol{y}) \in \mathcal{E}(\mathbf{0}, \boldsymbol{\Xi}\boldsymbol{\Omega}\boldsymbol{\Xi}^{\mathrm{T}}).\end{aligned}$$

Tím je pozorování dokončeno.

Při zobrazení R^{-1} a D elipsoidu F obdržíme

$$\begin{aligned}\boldsymbol{E}' &= \boldsymbol{D}\boldsymbol{R}^{-1}\boldsymbol{F}(\boldsymbol{R}^{-1})^{\mathrm{T}}\boldsymbol{D}^{\mathrm{T}} \\ &= \boldsymbol{D}\boldsymbol{R}^{-1}\boldsymbol{F}(\boldsymbol{R}^{-1})^{\mathrm{T}}\boldsymbol{D} \\ &= \boldsymbol{D}\boldsymbol{R}^{-1}\left(\frac{n^2}{n^2-1}\left(\boldsymbol{I} - \frac{2}{n+1}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\right)\right)(\boldsymbol{R}^{-1})^{\mathrm{T}}\boldsymbol{D} \\ &= \frac{n^2}{n^2-1}\Big(\boldsymbol{D}\underbrace{\boldsymbol{R}^{-1}\boldsymbol{I}(\boldsymbol{R}^{-1})^{\mathrm{T}}}_{=\boldsymbol{I}\text{ podle (5)}}\boldsymbol{D} - \frac{2}{n+1}\boldsymbol{D}\underbrace{\boldsymbol{R}^{-1}\boldsymbol{e}_1}_{=\boldsymbol{\theta}}\underbrace{\boldsymbol{e}_1^{\mathrm{T}}(\boldsymbol{R}^{-1})^{\mathrm{T}}}_{=\boldsymbol{\theta}^{\mathrm{T}}}\boldsymbol{D}\Big) \\ &= \frac{n^2}{n^2-1}\left(\boldsymbol{D}\boldsymbol{D} - \frac{2}{n+1}\boldsymbol{D}\frac{\boldsymbol{D}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{D}}{\|\boldsymbol{D}\boldsymbol{a}\|^2}\boldsymbol{D}\right) \\ &= \frac{n^2}{n^2-1}\left(\boldsymbol{E} - \frac{2}{n+1}\frac{\boldsymbol{E}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{a}}\right). \quad \square\end{aligned}$$

Geometrický náhled je následující: E' je nejmenší elipsoid, který obsahuje (i) řez elipsoidu E řeznou nadrovinou (tj. obsahuje množinu $E \cap \{\boldsymbol{x} : (\boldsymbol{a}-\boldsymbol{z})^{\mathrm{T}}\boldsymbol{x} = 0\}$) a (ii) bod ξ, který vznikne takto: rovnoběžně posuňme nadrovinu $\{\boldsymbol{x} : (\boldsymbol{a}-\boldsymbol{z})^{\mathrm{T}}\boldsymbol{x} = 0\}$ tak, aby byla tečnou elipsoidu E (viz tečkovanou čáru na obrázku 1); tečný bod je ξ.

Všimněme si ještě jednoho důležitého pozorování:

Důsledek 11. *Střed E' je uvnitř E.* □

Přichází důležitý odhad.

Lemma 12. *Buďte E, E' jako v lemmatu 10. Pak platí*

$$\frac{vol(E')}{vol(E)} < e^{-\frac{1}{2(n+1)}}.$$

Důkaz. Podle lemmatu 9 jest $\frac{vol(E')}{vol(E)} = \sqrt{\frac{\det \boldsymbol{E}'}{\det \boldsymbol{E}}}$. Pokračujme ve značení z předchozího lemmatu. Stačí vyšetřit případ $\boldsymbol{z} = \boldsymbol{0}$, $\boldsymbol{E} = \boldsymbol{I}$ a $\boldsymbol{a} = \boldsymbol{e}_1$; obecný případ se dostane pomocí regulárních zobrazení, a ta nemění poměry objemů. Pak

$$\begin{aligned}
\frac{vol(E')}{vol(E)} &= \sqrt{\frac{\det \boldsymbol{F}}{\det \boldsymbol{I}}} = \sqrt{\det \boldsymbol{F}} = \sqrt{d_1^2 \cdot d_2^2 \cdots d_n^2} = \sqrt{\left(\frac{n^2}{n^2-1}\right)^n \cdot \left(1 - \frac{2}{n+1}\right)} \\
&= \sqrt{\left(\frac{n^2}{n^2-1}\right)^{n-1} \cdot \frac{n^2}{(n-1)(n+1)} \cdot \frac{n-1}{n+1}} = \sqrt{\left(\frac{n^2}{n^2-1}\right)^{n-1}} \cdot \frac{n}{n+1} \\
&= \left(1 + \frac{n^2}{n^2-1} - 1\right)^{\frac{n-1}{2}} \cdot \left(1 + \frac{n}{n+1} - 1\right) = \left(1 + \frac{1}{n^2-1}\right)^{\frac{n-1}{2}} \cdot \left(1 - \frac{1}{n+1}\right) \\
&< \exp\left(\frac{n-1}{2} \cdot \frac{1}{n^2-1} - \frac{1}{n+1}\right) = \exp\left(\frac{1}{2(n+1)} - \frac{1}{n+1}\right) = \exp\left(-\frac{1}{2(n+1)}\right),
\end{aligned}$$

kde jsme využili toho, že $1 + \xi < e^{\xi}$, je-li $\xi \neq 0$ (jednou s $\xi := \frac{1}{n^2-1}$, podruhé s $\xi := -\frac{1}{n+1}$). □

Poznámka. Všimněme si, že první vztah z tvrzení lemmatu 10 obsahuje odmocninu. Není jej tedy možné spočítat (vyčíslit pomocí Turingova stroje) přesně, pracujeme-li v racionální aritmetice. Přesto budeme potřebovat tento vztah vyčíslit; později popíšeme strategii, jak počítat jen přibližně. Mimo jiné i popis této strategie činí Chačijanovu větu netriviální.

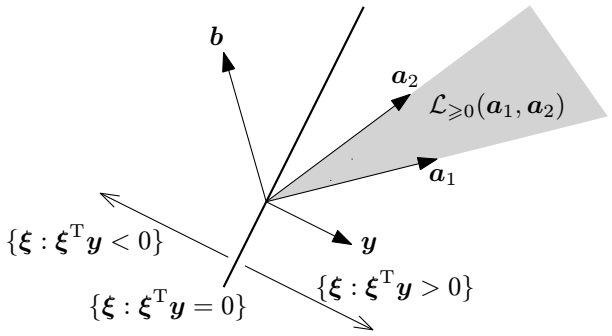
3 Farkasovo lemma

Z geometrie polyedrů budeme potřebovat jedno jednoduché, a přitom výtěžné tvrzení, známé jako *Farkasovo lemma*. Důkaz se snadno provede pomocí věty o dualitě; necháváme jej čtenáři jako cvičení.

Označme symbolem $\mathcal{L}_{\geqslant 0}$ systém nezáporných lineární kombinací (též někdy: systém *konických kombinací*) daných vektorů $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$:

$$\mathcal{L}_{\geqslant 0}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k) := \left\{ \sum_{i=1}^{k} \alpha_i \boldsymbol{a}_i : \ \alpha_i \geqslant 0, \ i = 1, 2, \ldots, k \right\}.$$

Někdy se této množině též říká *kužel* či *konus* generovaný vektory $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k$.



Obrázek 2: *Farkasovo lemma.*

Farkasovo lemma (zhruba) říká: $\boldsymbol{b} \notin \mathcal{L}_{\geqslant 0}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k)$*, právě když existuje nadrovina, která bod* $\boldsymbol{b}$ *od množiny* $\mathcal{L}_{\geqslant 0}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_k)$ *odděluje.* V tvrzení (a) následujícího lemmatu to řekněme to přesněji. Tvrzení (b) je jiná formulace, která se bude hodit později.

Lemma 13 (Farkas)**.** *(a) Neexistuje* $\boldsymbol{x} \geqslant \boldsymbol{0}$ *splňující soustavu* $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ *(jinými slovy:* $\boldsymbol{b} \notin \mathcal{L}_{\geqslant 0}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$*, kde* $\boldsymbol{a}_i$ *jsou sloupce* $\boldsymbol{A}$*), právě když existuje vektor* $\boldsymbol{y}$ *splňující* $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \geqslant \boldsymbol{0}$ *a* $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} < 0$ *(jinými slovy: nadrovina s normálou* $\boldsymbol{y}$ *odděluje* $\boldsymbol{b}$ *od* $\mathcal{L}_{\geqslant 0}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$*).*

(b) Soustava nerovností $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ *nemá řešení, právě když existuje vektor* $\boldsymbol{y} \geqslant \boldsymbol{0}$ *splňující* $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$ *a* $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} < 0$*; dokonce lze* $\boldsymbol{y}$ *volit tak, že* $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} = -1$*.* □

Poznámka. Tvrzení (a) nahlédneme neformálně takto: $\boldsymbol{b} \notin \mathcal{L}_{\geqslant}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$, právě když (z názoru) můžeme prostor $\mathbb{R}^n$ rozříznout nadrovinou na dva poloprostory tak, že v jednom poloprostoru je bod $\boldsymbol{b}$ a v druhém poloprostoru jsou všechny body $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$, a tudíž i množina $\mathcal{L}_{\geqslant 0}(\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n)$. (Ilustrace v rovině je zachycena na obrázku 2.) „Lze rozříznout“ znamená: existuje nadrovina $\{\boldsymbol{\xi} : \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{y} = 0\}$ rozdělující $\mathbb{R}^n$ na dva poloprostory $P^- := \{\boldsymbol{\xi} : \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} < 0\}$ a $P^+ := \{\boldsymbol{\xi} : \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{y} \geqslant 0\}$ takové, že $\boldsymbol{b} \in P^-$ a pro všechna $i = 1, 2, \ldots, n$ je $\boldsymbol{a}_i \in P^+$. To je jen jinými slovy vyřčené tvrzení (a): $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} < 0$ a $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \geqslant \boldsymbol{0}$.

Tvrzení (b) je důsledkem (a). Buď

$$\boldsymbol{A}' := (\boldsymbol{A} \;\; -\boldsymbol{A} \;\; \boldsymbol{I}). \tag{11}$$

Jak ukážeme v pozorování A, platí

$$\text{systém } \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b} \text{ má řešení} \quad \longleftrightarrow \quad \text{systém } \boldsymbol{A}'\boldsymbol{x}' = \boldsymbol{b} \text{ má nezáporné řešení.} \tag{12}$$

Tedy,

$$\begin{aligned} \text{systém } \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b} \text{ nemá řešení} \quad &\overset{(12)}{\longleftrightarrow} \quad \text{systém } \boldsymbol{A}'\boldsymbol{x}' = \boldsymbol{b} \text{ nemá řešení } \boldsymbol{x}' \geqslant \boldsymbol{0} \\ &\overset{(a)}{\longleftrightarrow} \quad \text{existuje } \boldsymbol{y} \text{ tak, že } (\boldsymbol{A}')^{\mathrm{T}}\boldsymbol{y} \geqslant \boldsymbol{0} \text{ a } \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} < 0 \\ &\overset{\text{poz. B}}{\longleftrightarrow} \quad \text{existuje } \boldsymbol{y} \geqslant \boldsymbol{0} \text{ tak, že } \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0} \text{ a } \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} < 0. \end{aligned} \tag{13}$$

Tím je (b) dokázáno; zbývá si už jen uvědomit, že vzhledem k tomu, že $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$, pro libovolné $k > 0$ platí $\boldsymbol{A}^{\mathrm{T}}(k \cdot \boldsymbol{y}) = \boldsymbol{0}$, a tedy stačí přenásobit $\boldsymbol{y}$ vhodnou konstantou k tak, aby $\boldsymbol{b}^{\mathrm{T}}(k \cdot \boldsymbol{y}) = -1$.

Dokažme pozorování A a B.

Pozorování A. Platí (12). Důkaz. Nechť $\boldsymbol{x}'$ je tvaru $\begin{pmatrix} \boldsymbol{x}'_{\boldsymbol{A}} \\ \boldsymbol{x}'_{-\boldsymbol{A}} \\ \boldsymbol{x}'_{\boldsymbol{I}} \end{pmatrix}$ a jednotlivé složky odpovídají komponentám $\boldsymbol{A}$, $-\boldsymbol{A}$, $\boldsymbol{I}$ matice $\boldsymbol{A}'$. Má-li systém $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ řešení $\boldsymbol{x}$, je

$$(\boldsymbol{x}'_{\boldsymbol{A}})_i := \max\{(\boldsymbol{x})_i, 0\}, \quad (\boldsymbol{x}'_{-\boldsymbol{A}})_i := \max\{-(\boldsymbol{x})_i, 0\}, \quad (\boldsymbol{x}'_{\boldsymbol{I}})_i := (\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})_i$$

nezáporným řešením systému $\boldsymbol{A}'\boldsymbol{x}' = \boldsymbol{b}$. A naopak, má-li systém $\boldsymbol{A}'\boldsymbol{x}' = \boldsymbol{b}$ nezáporné řešení, pak má i nezáporné řešení $\boldsymbol{x}'$ takové, že pro každé i je alespoň jedno z čísel $(\boldsymbol{x}'_{\boldsymbol{A}})_i$ a $(\boldsymbol{x}'_{-\boldsymbol{A}})_i$ nulové. Pak $(\boldsymbol{x})_i$ definované vztahem

$$(\boldsymbol{x})_i := \begin{cases} (\boldsymbol{x}'_{\boldsymbol{A}})_i, & \text{je-li } (\boldsymbol{x}'_{\boldsymbol{A}})_i > 0, \\ -(\boldsymbol{x}'_{-\boldsymbol{A}})_i, & \text{je-li } (\boldsymbol{x}'_{-\boldsymbol{A}})_i > 0, \\ 0 & \text{jinak} \end{cases}$$

je řešením systému $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$. Tím je pozorování A dokončeno.

Pozorování B. Platí poslední ekvivalence v (13). Důkaz. Nerovnost $(\boldsymbol{A}')^{\mathrm{T}}\boldsymbol{y} \geqslant \boldsymbol{0}$, vzhledem k (11), říká

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \geqslant \boldsymbol{0}, \quad -\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \geqslant \boldsymbol{0}, \quad \boldsymbol{I}\boldsymbol{y} \geqslant \boldsymbol{0}.$$

Tato trojice nerovností je ekvivalentní s $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$, $\boldsymbol{y} \geqslant \boldsymbol{0}$. Tím je pozorování dokončeno.

4 Velikost zápisu řešení soustav lineárních rovnic

Symbolem $\sigma(\cdot)$ značíme počet bitů k zápisu racionálního čísla, racionálního vektoru či racionální matice. Z první přednášky víme, že pro celočíselnou matici $\boldsymbol{A}$ platí

$$\sigma(\det \boldsymbol{A}) < \sigma(\boldsymbol{A}) \tag{14}$$

(zde $\sigma(\det \boldsymbol{A})$ značí počet bitů na zápis celého čísla a $\sigma(\boldsymbol{A})$ značí počet bitů na zápis celočíselné matice).

Pozorování 14. *(a) Pro $k, l \in \mathbb{N}$ je $\sigma(k^l) \leqslant l \cdot \sigma(k)$.*

(b) Buď $\boldsymbol{x} = (\frac{p_1}{q_1}, \frac{p_2}{q_2}, \ldots, \frac{p_n}{q_n})$, kde pro každé i platí, že je-li $p_i \neq 0$, jsou čísla $p_i \in \mathbb{Z}$ a $q_i \in \mathbb{N} \setminus \{0\}$ nesoudělná. Platí $\prod_{i=1}^n q_i \leqslant 2^{\sigma(\boldsymbol{x})}$.

(c) Pro racionální vektor $\boldsymbol{x}$ platí $\|\boldsymbol{x}\| \leqslant 2^{\sigma(\boldsymbol{x})}$.

Důkaz (a) je zřejmý; (b) plyne z odhadu

$$2^{\sigma(\boldsymbol{x})} = 2^{n+\sum_{i=1}^n (2+\lceil \log_2(|p_i|+1)\rceil + \lceil \log_2(q_i+1)\rceil)} \geqslant 2^{\sum_{i=1}^n \log_2(q_i+1)} = \prod_{i=1}^n (q_i+1) \geqslant \prod_{i=1}^n q_i$$

a (c) plyne z odhadu

$$\begin{aligned} 2^{\sigma(\boldsymbol{x})} &= 2^{n+\sum_{i=1}^n (2+\lceil \log_2(|p_i|+1)\rceil + \lceil \log_2(q_i+1)\rceil)} \geqslant 2^{\sum_{i=1}^n \log_2(|p_i|+1)} \\ &= \prod_{i=1}^n (|p_i|+1) \overset{(*)}{\geqslant} \sum_{i=1}^n |p_i| = \|(p_1, p_2, \ldots, p_n)\|_1 \geqslant \|(p_1, p_2, \ldots, p_n)\| \geqslant \|\boldsymbol{x}\|. \end{aligned}$$

Užili jsme známé nerovnosti mezi součtovou normou, zde značenou $\|\cdot\|_1$, a eukleidovskou normou. Nerovnost $(*)$ se nahlédne tak, že se závorky ve výrazu $\prod_{i=1}^n (|p_i|+1)$ roznásobí: ve vzniklém (dlouhém) součtu se vyskytují všechny sčítance $|p_i|$. □

Vstupem pro problém **LP** je celočíselná matice $\boldsymbol{A}$ a celočíselný vektor $\boldsymbol{b}$: počet bitů k jeho zápisu označíme

$$L := \sigma((\boldsymbol{A}\ \boldsymbol{b})). \tag{15}$$

Ukážeme, že problém **LP** je rozhodnutelný v čase $poly(L)$. Nalezneme polynom v proměnných n (dimense), m (počet nerovností) a L; ovšem zřejmě platí

$$m \leqslant L, \quad n \leqslant L,$$

a tak se každý polynom v n, m a L nechá omezit polynomem v L.

Již jsme konstatovali, že celočíselná matice s malou velikostí zápisu má determinant s malou velikostí zápisu. Důsledkem je, že má-li soustava lineárních rovnic řešení, pak má i řešení s omezenou velikostí zápisu.[2]

Lemma 15. *Má-li soustava rovnic $\boldsymbol{Ax} = \boldsymbol{b}$ řešení, pak má řešení $\boldsymbol{x}$ splňující $\sigma(\boldsymbol{x}) \leqslant 3nL$.*

Důkaz. Jestliže $\boldsymbol{Ax} = \boldsymbol{b}$ má řešení, můžeme ze soustavy $\boldsymbol{Ax} = \boldsymbol{b}$ vybrat maximální (co do počtu rovnic) systém nezávislých rovnic; dostaneme tak soustavu

$$\boldsymbol{B}\boldsymbol{x}_{\boldsymbol{B}} + \boldsymbol{C}\boldsymbol{x}_{\boldsymbol{C}} = \boldsymbol{b}', \tag{16}$$

kde $\boldsymbol{B}$ je regulární matice rozměru $r \times r$ a $r \leqslant n$ (a $\boldsymbol{C}$ může být i prázdná matice). Položíme-li $\boldsymbol{x}_{\boldsymbol{C}} = \boldsymbol{0}$, můžeme $\boldsymbol{x}_{\boldsymbol{B}}$ jednoznačně dopočítat Cramerovým pravidlem

$$(\boldsymbol{x}_{\boldsymbol{B}})_j = \frac{\det \boldsymbol{B}_j}{\det \boldsymbol{B}},$$

[2] Všimněme si, že věta *má-li systém lineárních rovnic $\boldsymbol{Ax} = \boldsymbol{b}$ s celočíselnými koeficienty řešení, pak má racionální řešení s polynomiální velikostí zápisu* plyne z toho, že Gaussova eliminace počítá v polynomiálním čase. Následující lemma toto pozorování zesiluje do tvaru, který ukazuje explicitní horní mez na velikost zápisu řešení. Ideu, na níž je postaven důkaz lemmatu, jsme již užili v důkazu polynomiality Gaussovy eliminace.

kde $\boldsymbol{B}_j$ je matice, která z $\boldsymbol{B}$ vznikla nahrazením j-tého sloupce vektorem $\boldsymbol{b}'$. Tedy podle definice velikosti popisu racionálního čísla a podle (14) dostáváme

$$\begin{aligned}\sigma((\boldsymbol{x}_{\boldsymbol{B}})_j) &= 2 + \sigma(\det(\boldsymbol{B}_j)) + \sigma(\det(\boldsymbol{B})) \\ &= \underbrace{1 + \sigma(\det(\boldsymbol{B}_j))}_{\leqslant \sigma(\boldsymbol{B}_j)} + \underbrace{1 + \sigma(\det(\boldsymbol{B}))}_{\leqslant \sigma(\boldsymbol{B})} \leqslant \underbrace{\sigma(\boldsymbol{B}_j)}_{\leqslant L} + \underbrace{\sigma(\boldsymbol{B})}_{\leqslant L} \leqslant 2L\end{aligned}$$

(velikost matice $\boldsymbol{B}$ i $\boldsymbol{B}_j$ je nanejvýš L, protože L jsme definovali jako velikost matice $(\boldsymbol{A}\ \boldsymbol{b})$).

Řešení $\boldsymbol{x} = \binom{\boldsymbol{x}_{\boldsymbol{B}}}{\boldsymbol{0}}$ splňuje $\sigma(\boldsymbol{x}) \leqslant 3nL$. Jest totiž

$$\begin{aligned}\sigma(\boldsymbol{x}) &= n + \sum_{j=1}^{r} \sigma((\boldsymbol{x}_{\boldsymbol{B}})_j) + \sum_{j=r+1}^{n} \sigma(0) \\ &= n + \underbrace{\sum_{j=1}^{r} \sigma((\boldsymbol{x}_{\boldsymbol{B}})_j)}_{\leqslant 2nL} + 4(n-r) \leqslant 5n + 2nL \leqslant nL + 2nL = 3nL.\end{aligned}$$

Velikost popisu nuly zapsané jako racionální číslo $+\frac{0}{1}$ je podle definice $\sigma(\cdot)$ rovna 4. Zřejmě lze bez újmy na obecnosti předpokládat, že $L \geqslant 5$. □

Důsledek 16. *Každý vrchol polyedru* $\{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$ *má velikost popisu nanejvýš* $3nL$. □

5 Ořezávací lemma

Dokázali jsme, že k danému polyedru vždy můžeme přidat systém nerovností, které zaručí jeho omezenost, a přitom se vlastnost „býti neprázdný“ zachovává.

Důsledek 17 (o ořezání)**.** *Soustava nerovnic* $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ *má řešení, právě když soustava nerovnic*

$$\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}, \quad -2^{3nL}\boldsymbol{1} \leqslant \boldsymbol{x} \leqslant 2^{3nL}\boldsymbol{1} \tag{17}$$

má řešení. Navíc, velikost zápisu soustavy (17) *je omezena polynomem (v proměnných* n *a* L*) a lze ji zkonstruovat v polynomiálním čase.* □

Věc je snadná. Předpokládali jsme, že je-li polyedr neprázdný, má vrchol. Z předchozí části víme, že platí: je-li polyedr neprázdný, má vrchol, jenž má velikost popisu nanejvýš $3nL$, a tudíž každá z jeho složek je (v absolutní hodnotě) nanejvýš 2^{3nL}. Polyedr (17) je očividně omezený a velikost popisu soustavy (17) představuje jen polynomiální (v n, L) zvětšení oproti velikosti popisu L původní soustavy.

6 Nafukovací lemma

Na konci této kapitoly budeme moci prohlásit: *bez újmy na obecnosti lze předpokládat, že polyedr* $\{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$ *je buďto prázdný, anebo je omezený a má plnou dimensi.* Například soustava nerovnic

$$x_1 + x_2 \leqslant 1, \quad -x_1 - x_2 \leqslant -1 \tag{18}$$

definuje neomezený polyedr, který nemá plnou dimensi: je to totiž přímka (jednodimensionální objekt) v rovině ($n = 2$). Soustavu (18) nahradíme jinou soustavou, jež definuje polyedr, který je *dvoudimensionální* a *omezený*. Obecně: popíšeme transformaci, jak k systému $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ sestrojit systém $\boldsymbol{A}'\boldsymbol{x} \leqslant \boldsymbol{b}'$ takový, že

(a) systém $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ má řešení, právě když systém $\boldsymbol{A}'\boldsymbol{x} \leqslant \boldsymbol{b}'$ má řešení,

(b) má-li systém $\boldsymbol{A}'\boldsymbol{x} \leqslant \boldsymbol{b}'$ řešení, pak polyedr $\{\boldsymbol{x} : \boldsymbol{A}'\boldsymbol{x} \leqslant \boldsymbol{b}'\}$ je omezený a má plnou dimensi,

(c) systém $\boldsymbol{A}'\boldsymbol{x} \leqslant \boldsymbol{b}'$ lze zkonstruovat v polynomiálním čase.

Omezeností jsme se zabývali v předchozí části; nyní ukažme, jak zajistit plnou dimensi.

Lemma 18 (o nafouknutí)**.** *$\mathcal{F} := \{x : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\} \neq \emptyset$, právě když*

$$\mathcal{F}_\varepsilon := \{x : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b} + \boldsymbol{1}\varepsilon\} \neq \emptyset, \tag{19}$$

kde $\varepsilon = \frac{1}{2m}2^{-6mL}$. Navíc, je-li $\mathcal{F}_\varepsilon \neq \emptyset$, pak má plnou dimensi. □

Nechť si čtenář všimne, že přechodem od soustavy $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ k soustavě $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b} + \boldsymbol{1}\varepsilon$ dojde jen k polynomiálnímu zvětšení velikosti zápisu soustavy. Ještě je nutné každou nerovnost ze systému $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b} + \boldsymbol{1}\varepsilon$ přenásobit nejmenším společným násobkem jmenovatelů zlomků, abychom obdrželi systém s celočíselnými koeficienty. To ovšem představuje opět jen polynomiální zvětšení.

Důkaz lemmatu 18. Zřejmě $\mathcal{F} \subseteq \mathcal{F}_\varepsilon$, takže z $\mathcal{F} \neq \emptyset$ plyne i $\mathcal{F}_\varepsilon \neq \emptyset$. Ukážeme, že z $\mathcal{F} = \emptyset$ plyne i $\mathcal{F}_\varepsilon = \emptyset$. Opřeme se o Farkasovo lemma (lemma 13).

Nechť $\mathcal{F} = \emptyset$. Díky lemmatu 13(b) předpoklad $\mathcal{F} = \emptyset$ znamená, že systém (ne)rovnic

$$\boldsymbol{y} \geqslant \boldsymbol{0}, \quad \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}, \quad \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} = -1 \tag{20}$$

má řešení $\boldsymbol{y}$. Tento systém definuje neprázdný polyedr v prvním kvadrantu; má proto vrchol. Podle důsledku 16 existuje řešení (vrchol) systému (20) $\boldsymbol{y}$ s velikostí zápisu nanejvýš $6mL$ (užíváme toho, že velikost zápisu soustavy $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}, \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} = -1$ je nanejvýš $2L$ a počet proměnných je m, takže v tvrzení důsledku v roli dimense vystupuje číslo m namísto čísla n a v roli odhadu na velikost zápisu systému rovnic vystupuje číslo $2L$ namísto čísla L).

Pak jest

$$(\boldsymbol{b} + \boldsymbol{1}\varepsilon)^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} + \sum_{i=1}^{m} \varepsilon \cdot (\boldsymbol{y})_i < 0,$$

protože $\boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} = -1$ a

$$\sum_{i=1}^{m} \varepsilon \cdot (\boldsymbol{y})_i \leqslant \sum_{i=1}^{m} \underbrace{\frac{1}{2m}2^{-6mL}}_{=\varepsilon} \cdot \underbrace{2^{6mL}}_{\geqslant (\boldsymbol{y})_i} = \frac{1}{2}.$$

Fakt $(\boldsymbol{b} + \boldsymbol{1}\varepsilon)^{\mathrm{T}}\boldsymbol{y} < \boldsymbol{0}$ spolu s tím, že $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$ a $\boldsymbol{y} \geqslant \boldsymbol{0}$, podle lemmatu 13(b) impikuje, že systém $\boldsymbol{A}\boldsymbol{x} \leqslant (\boldsymbol{b} + \boldsymbol{1}\varepsilon)$ nemá řešení, a proto $\mathcal{F}_\varepsilon = \emptyset$.

Tvrzení o plné dimensi je zřejmé: je-li $\boldsymbol{x}$ řešením systému $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$, pak $\boldsymbol{x}$ splňuje $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b} < \boldsymbol{b} + \boldsymbol{1}\varepsilon$; všechny nerovnosti definující $\mathcal{F}_\varepsilon$ jsou splněny jako ostré, a tedy $\boldsymbol{x}$ je vnitřním bodem polyedru $\mathcal{F}_\varepsilon$, kterému lze uvnitř $\mathcal{F}_\varepsilon$ opsat kouli nenulového objemu. □

7 Elipsoidový algoritmus (Šorova-Judinova-Němirovského verse)

Body (a) – (c) shrnují, co již víme.

(a) Můžeme předpokládat, že polyedr $\mathcal{F} = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$ je omezený. Tudíž, je-li neprázdný, je konvexním obalem svých vrcholů.

(b) Podle důsledku 16 má každý jeho vrchol velikost zápisu nanejvýš $3nL$, takže podle pozorování 14(c) je $\mathcal{F} \subseteq \{\boldsymbol{x} : \|\boldsymbol{x}\| \leqslant 2^{3nL}\}$ (celý polyedr je uzavřen ve velké kouli).

(c) Můžeme předpokládat, že buďto $\mathcal{F} = \emptyset$, anebo má $\mathcal{F}$ plnou dimensi.

(d) Snadno dokážeme, že je-li $\mathcal{F} \neq \emptyset$, pak $\mathcal{F}$ má objem alespoň 2^{-7n^2L}.

Než se pustíme do důkazu (d), stojí za to shrnout, jak se dosáhne platnosti (a) a (c) *zároveň*. Nechť je dán systém $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$, o němž dopředu víme jen to, že je buďto prázdný anebo má vrchol.

Krok 1. Vyčíslíme velikost zápisu tohoto systému L_0 a přejdeme k systému (17) (s $L := L_0$). Tím splníme (a).

Krok 2. Vyčíslíme velikost zápisu L_1 systému z kroku 1 a počet jeho nerovností označíme nerovností m_1 ($= m_0 + 2n$, kde m_0 je počet nerovností původního systému $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$). Pak přejdeme k systému (19) (s $L := L_1$ a $m := m_1$). Všimněme si, že ε-nafouknutí (19) zachová omezenost polyedru díky přítomnosti nerovností $-2^{3nL_0}\mathbf{1} \leqslant \boldsymbol{x} \leqslant 2^{3nL_0}\mathbf{1}$ z kroku 1, takže nyní platí (a) i (c).

Zbývá již jen vzniklý systém přenásobit tak, aby jeho koeficienty byly celočíselné.

Tím jsme dospěli k finálnímu systému. Užíváme-li v (b), (d) a při následné analýze symbolů L a m, jde vždy o velikost zápisu a počet nerovností finálního systému.

A nyní prokažme (d).

Lemma 19. *Je-li $\mathcal{F} \neq \emptyset$ a má plnou dimensi, pak $vol(\mathcal{F}) \geqslant 2^{-7n^2L}$.*

Důkaz. Protože předpokládáme, že polyedr $\mathcal{F}$ je omezený a má v $\mathbb{R}^n$ dimensi n, má alespoň $n+1$ vrcholů $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n+1}$, a tedy

$$conv(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n+1}) \subseteq \mathcal{F}.$$

Platí známý vzoreček[3]:

$$vol(conv(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n+1})) = \frac{1}{n!}\left|\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_{n+1} \end{pmatrix}\right|.$$

Pozorování. Je-li $\boldsymbol{A} = (\frac{p_{ij}}{q_{ij}})_{ij}$ regulární matice racionálních čísel, existuje celé číslo p a přirozené číslo q tak, že $\det \boldsymbol{A} = \frac{p}{q}$ *a* $q \leqslant \prod_{i,j} q_{ij}$. *Důkaz* plyne přímo z definice determinantu: determinant matice $\boldsymbol{A}$ je definován jako jistý součet racionálních čísel, která mají společný jmenovatel nanejvýš $\prod_{i,j} q_{ij}$. Tím je pozorování dokázáno.

Pohleďme na racionální číslo

$$\frac{p}{q} := \left|\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_{n+1} \end{pmatrix}\right|,$$

[3]Vzoreček v sobě skrývá zajímavou geometrii, nechť si ji čtenář rozebere alespoň pro $n = 2$ a $n = 3$. Vzoreček lze číst tak, že rovnoběžnostěn, jehož hrany jsou určeny vektory $\boldsymbol{x}_1 - \boldsymbol{x}_{n+1}$, $\boldsymbol{x}_2 - \boldsymbol{x}_{n+1}, \ldots$, $\boldsymbol{x}_n - \boldsymbol{x}_{n+1}$, lze rozkrájet na $n!$ simplexů. (Simplexem v dimensi n rozumíme konvexní obal $n+1$ bodů, které neleží ve stejné nadrovině; simplexem v $\mathbb{R}^2$ je trojúhelník, simplexem v $\mathbb{R}^3$ je čtyřstěn.)

kde přirozená čísla p a q jsou nesoudělná. Podle pozorování je

$$q \leqslant [1 \cdot \underbrace{\mathit{jmenovatel}((\boldsymbol{x}_1)_1) \cdot \mathit{jmenovatel}((\boldsymbol{x}_1)_2) \cdots \mathit{jmenovatel}((\boldsymbol{x}_1)_n)}_{\leqslant 2^{3nL}}] \cdots$$
$$\cdots [1 \cdot \underbrace{\mathit{jmenovatel}((\boldsymbol{x}_{n+1})_1) \cdot \mathit{jmenovatel}((\boldsymbol{x}_{n+1})_2) \cdots \mathit{jmenovatel}((\boldsymbol{x}_{n+1})_n)}_{\leqslant 2^{3nL}}]$$
$$\leqslant 2^{(n+1)\cdot 3nL} \leqslant 2^{6n^2L}.$$

Využili jsme toho, každý vrchol $\boldsymbol{x}_i$ má velikost zápisu nanejvýš $3nL$ (důsledek 16), a pozorování 14(b).

Determinant je nenulový (díky předpokladu o dimensi), takže $p \geqslant 1$; užijeme-li dále odhadu $n! \leqslant n^n = 2^{n \log_2 n} \leqslant 2^{n^2}$, dostáváme

$$\mathit{vol}(\mathcal{F}) \geqslant \frac{1}{n!}\frac{p}{q} \geqslant \frac{1}{2^{n^2}}\frac{p}{q} \geqslant \frac{1}{2^{n^2}}\frac{1}{2^{6n^2L}} \geqslant 2^{-7n^2L}. \quad \square$$

Popis algoritmu

Budiž dána celočíselná matice $\boldsymbol{A}$ a celočíselný vektor $\boldsymbol{b}$ takové, že polyedr $\mathcal{F} := \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$ je omezený, a je-li neprázdný, má plnou dimensi. Cílem je rozhodnout, zdali $\mathcal{F} = \emptyset$ nebo nikoliv.

Nejprve prozkoumejme myšlenku tzv. elipsoidového algoritmu, který na konci šedesátých let popsali Naum Šor, David Judin a Arkadij Němirovskij. Až později se budeme zabývat jeho speciální versí, která prokazuje větu $\mathbf{LP} \in \boldsymbol{P}$.

Budeme konstruovat posloupnost elipsoidů

$$E_i = \mathcal{E}(\boldsymbol{z}_i, \boldsymbol{E}_i),$$

které representujeme středy $\boldsymbol{z}_i$ a maticemi $\boldsymbol{E}_i$. Za E_0 vezměme kouli se středem v nule a poloměrem 2^{3nL}:

$$E_0 := \mathcal{E}(\boldsymbol{z}_0 := \boldsymbol{0}, \boldsymbol{E}_0 := (2^{3nL})^2\boldsymbol{I}).$$

Podle bodu (b) máme zaručeno, že $\mathcal{F} \subseteq E_0$.

Iterace i. Z minulého kroku máme elipsoid $E_{i-1} = \mathcal{E}(\boldsymbol{z}_{i-1}, \boldsymbol{E}_{i-1})$. Otestujeme, zdali

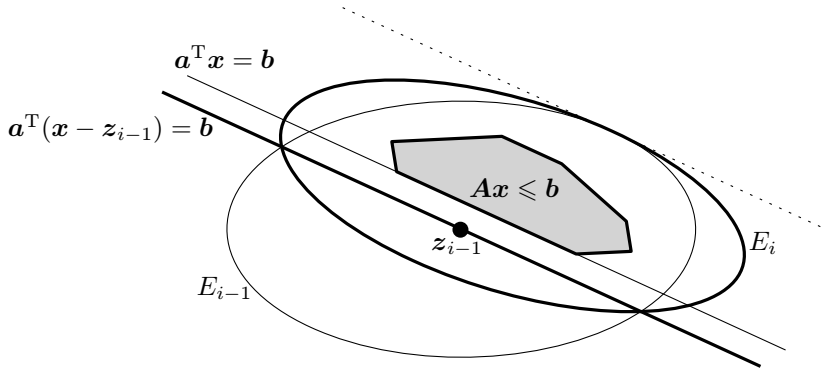
$$\boldsymbol{A}\boldsymbol{z}_{i-1} \leqslant \boldsymbol{b};$$

pokud ano, máme řešení, skončíme a odpovíme „ano, řešení existuje“. Jinak existuje alespoň jedna nerovnice $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} \leqslant b$ v systému $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$, která není vektorem $\boldsymbol{z}_{i-1}$ splněna: platí $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{z}_{i-1} > b$. (Je-li jich víc, vybereme libovolnou.)

Zkonstruujeme eliposid E_i tak, že z elipsoidu E_{i-1} polovinu uřízneme: povedeme řez nadrovinou procházející středem $\boldsymbol{z}_{i-1}$, rovnoběžnou s nadrovinou $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = b$, a zahodíme tu polovinu elipsoidu E_{i-1}, kde platí $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} > b$. Elipsoidem E_i bude nejmenší (co do objemu) elipsoid obsahující nezahozenou polovinu E_{i-1}. Použijeme na to lemma 10:

$$\boldsymbol{z}_i := \boldsymbol{z}_{i-1} - \frac{1}{n+1} \cdot \frac{\boldsymbol{E}_{i-1}\boldsymbol{a}}{\sqrt{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}\boldsymbol{a}}}, \tag{21}$$

$$\boldsymbol{E}_i := \frac{n^2}{n^2-1} \cdot \left(\boldsymbol{E}_{i-1} - \frac{2}{n+1}\frac{\boldsymbol{E}_{i-1}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}\boldsymbol{a}}\right). \tag{22}$$



Obrázek 3: *Krok elipsoidového algoritmu.*

Všimněme si, že jsme uřízli tu polovinu, kde nemůže být žádné řešení; proto $\mathcal{F} \subseteq E_{i-1}$ implikuje $\mathcal{F} \subseteq E_i$.

Později se budeme zabývat aritmetikou — při výpočtu na Turingově stroji si totiž nemůžeme dovolit, aby čísla v $\boldsymbol{z}_i$ a $\boldsymbol{E}_i$ byla příliš veliká co do velikosti zápisu. Pak by totiž i banální operace typu součet či součin mohly zabrat exponenciální čas. Ve vzorci (21) se dokonce vyskytuje odmocnina — to může být iracionální číslo, které se ve formě dvojkového rozvoje nedá na Turingově stroji representovat vůbec. Řešení těchto problémů je podstatou následujících kapitol. Ještě předtím si však stojí za to všimnout, že elipsoidová metoda počítá rychle, nemusíme-li se zabývat velikostí zápisu čísel.

8 Algoritmus pracuje správně a rychle, můžeme-li (teoreticky) representovat reálná čísla přesně

Na chvíli předpokládejme, že máme hypotetický počítač, který dokáže pracovat s reálnými čísly (s neomezenou přesností). Nahlédneme, že za tohoto předpokladu má náš algoritmus *polynomiální počet iterací*. Uvědomme si totiž, že jakmile je objem elipsoidu příliš malý (z lemmatu 19 máme k disposici dolní odhad na objem polyedru, je-li neprázdný), můžeme výpočet ukončit, neboť systém $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ už nemůže mít řešení.

Tvrzení 20. *Nechť je dán polyedr $\mathcal{F} := \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$ (kde $\boldsymbol{A}, \boldsymbol{b}$ jsou celočíselné), který je omezený a buď je prázdný, anebo má plnou dimensi. Ve výpočetním modelu, který dokáže representovat reálná čísla a přesně s nimi provádět operace $+, -, \cdot, \div, \sqrt{\ }$, platí, že nenajde-li elipsoidový algoritmus řešení soustavy $\boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}$ s celočíselnými koeficienty do $(64n^3L)$-té iterace, pak $\mathcal{F} = \emptyset$.*

Důkaz. Označme $R := 2^{3nL}$. Podle (a) a (b) víme, že $vol(E_0) \leqslant (2R)^n$, a podle lemmatu 12 je

$$vol(E_i) < e^{-\frac{i}{2n+2}} \cdot (2R)^n.$$

Díky lemmatu 19 máme dolní odhad na objem $\mathcal{F}$. Dosáhneme-li N-té iterace s

$$N := 64n^3L \quad (= poly(L)),$$

vede předpoklad $vol(\mathcal{F}) > 0$ ke sporu:

$$\begin{aligned} 2^{-7n^2L} &\overset{\text{lemma 19}}{\leqslant} vol(\mathcal{F}) \overset{(\forall i)\mathcal{F}\subseteq E_i}{\leqslant} vol(E_N) \overset{\text{lemma 12}}{<} e^{-\frac{N}{2n+2}} \cdot (2R)^n \\ &< 2^{-\frac{64n^3L}{2n+2}} (2 \cdot 2^{3nL})^n \leqslant 2^{-\frac{64n^3L}{4n}} \cdot 2^n \cdot 2^{3n^2L} = 2^{-16n^2L+n+3n^2L} \\ &= 2^{-7n^2L-n(6nL-1)} \overset{6nL>1}{<} 2^{-7n^2L}. \end{aligned}$$

Protože jsme předpokládali, že buď $\mathcal{F} = \emptyset$ anebo $\mathcal{F}$ má plnou dimensi, musí být $\mathcal{F} = \emptyset$. □

V $(64n^3L)$-té iteraci můžeme skončit a prohlásit „řešení neexistuje“.

Je jasné, že najde-li algoritmus nějaké řešení, pak $\mathcal{F} \neq \emptyset$. Tedy, dokázali jsme, že $\mathcal{F} = \emptyset$, *právě když* algoritmus skončí s odpovědí „řešení neexistuje“. Můžeme říci, že pro popsaný výpočetní model platí: algoritmus má polynomiální počet iterací a počítá správně.

9 Zaokrouhlování

Je třeba zajistit, aby výpočetní čas *uvnitř iterace* byl omezen polynomem v L. Prostředkem k řešení tohoto problému bude zaokrouhlování, či přesněji *vyčíslování vzorců (21) a (21) jen na omezený počet míst za dvojkovou čárkou*. Tak zajistíme, že veškeré aritmetické výpočty budou probíhat nad *krátkými* čísly, a tudíž v polynomiálním čase. (Nechť si čtenář rozmyslí, že všechny potřebné operace — sčítání, odčítání, násobení, dělení, vyčíslení odmocniny na $poly(L)$ míst — lze provést v racionální aritmetice dostatečně rychle). Vyčíslení výrazů (21) a (22) na omezený počet míst ovšem může vést k deformaci: může dojít k nežádoucímu posunutí nebo pootočení elipsoidu či ke změně jeho tvaru. Pak by se mohlo stát, že $\mathcal{F} \not\subseteq E_i$; algoritmus by pak nemusel počítat správně. Dokonce by mohlo dojít k tomu, že matice, jež má representovat elipsoid, nebude positivně definitní. Ukážeme, že zaokrouhlování lze provést tak, aby žádný z citovaných problémů nenastal. Přitom bude třeba zajistit, aby příliš nevzrostl počet iterací; ukáže se, že počet iterací se nanejvýš zdvojnásobí.

K dosažení vytčeného cíle budeme potřebovat trochu techniky.

Definice 21. *ε-**nafouknutí** množiny $A \subseteq \mathbb{R}^n$ je množina*

$$\mathcal{B}(A, \varepsilon) := \{\boldsymbol{x} : \text{existuje } \boldsymbol{y} \in A \text{ s } \|\boldsymbol{x} - \boldsymbol{y}\| \leqslant \varepsilon\}$$

*a k-**natáhnutí** elipsoidu $E := \mathcal{E}(\boldsymbol{z}, \boldsymbol{E})$ je*

$$k \bullet E := \mathcal{E}(\boldsymbol{z}, k^2 \cdot \boldsymbol{E}). \quad \square$$

Všimněme si rozdílu mezi $\mathcal{B}(E, \varepsilon)$ a $(1 + \varepsilon) \bullet E$. Množinu $\mathcal{B}(E, \varepsilon)$ si můžeme představovat jako obalení elipsoidu E vrstvou tlustou ε,[4] zatímco $(1 + \varepsilon)$-natáhnutí znamená, že elipsoid zvětšíme (s $\varepsilon > 0$) nebo zmenšíme (s $\varepsilon < 0$) tak, že délky všech poloos zvětšíme (zmenšíme) $(1 + \varepsilon)$-krát, přičemž získáme opět elipsoid.

Pozorování 22. *Nechť $E := \mathcal{E}(\boldsymbol{z}, \boldsymbol{E})$ a $\epsilon \geqslant 0$.*

(a)

$$\left(1 + \frac{\epsilon}{\sqrt{\Lambda(\boldsymbol{E})}}\right) \bullet E \subseteq \mathcal{B}(E, \epsilon) \subseteq \left(1 + \frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}}\right) \bullet E. \tag{23}$$

[4] Obecně množina $\mathcal{B}(E, \varepsilon)$ není elipsoidem: nechť si čtenář nakreslí např. množinu $\mathcal{B}(\mathcal{E}(\boldsymbol{0}, \boldsymbol{E}), 5)$, kde $\boldsymbol{E} = \left(\begin{smallmatrix} 100 & 0 \\ 0 & \frac{1}{100} \end{smallmatrix}\right)$.

(b) $vol(k \bullet E) = k^n \cdot vol(E)$.

(c) *Je-li* $\varphi \geqslant 0$, *pak* $\mathcal{B}(\mathcal{B}(E,\epsilon),\varphi) \subseteq \mathcal{B}(E,\epsilon+\varphi)$.

(d) $x \in k \bullet E$, *právě když* $\|\boldsymbol{x}\|_{\boldsymbol{E}} \leqslant k$. □

Nejprve, kvůli intuici, nahlédněme geometrii tvrzení (a); důkaz podáme záhy. Představme si, že elipsoid E nafoukneme např. o $\epsilon = 1$ cm; dostaneme tak množinu $\mathcal{B}(E,\epsilon)$. Chceme jí vepsat co největší k-natáhnutí elipsoidu E. Je patrné, že k-natáhneme-li E tak, že nejdelší poloosa se prodlouží o 1 cm (to odpovídá $k = 1 + \frac{\epsilon}{\text{délka nejdelší poloosy}}$), ostatní poloosy (kratší než ta nejdelší) se prodlouží o méně než 1 cm. Takže

$$\left(1 + \frac{\epsilon}{\sqrt{\Lambda(\boldsymbol{E})}}\right) \bullet E \subseteq \mathcal{B}(E,\epsilon).$$

Naopak, chceme-li množině $\mathcal{B}(E,\epsilon)$ opsat co nejmenší elipsoid získaný z E operací k-natáhnutí, je třeba volit k tak, aby se nejkratší poloosa elipsoidu E prodloužila alespoň o 1 cm ($k = 1 + \frac{\epsilon}{\text{délka nejkratší poloosy}}$); poloosy delší, než je ta nejkratší, se pak prodlouží o více než 1 cm. Z toho

$$\mathcal{B}(E,\epsilon) \subseteq \left(1 + \frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}}\right).$$

Tolik náhled. Zejména v případě druhé inkluse v (23) je ovšem intuice ošemetná, protože při hodně „placatém“ elipsoidu E může být množina $\mathcal{B}(E,\epsilon)$ hodně „oválovitá“.

Důkaz pozorování 22. (a) Nejprve prokažme druhou inklusi. Buď $\boldsymbol{x} \in \mathcal{B}(E,\epsilon)$; existuje tedy $\boldsymbol{y} \in E$ s $\|\boldsymbol{x}-\boldsymbol{y}\| \leqslant \varepsilon$. Protože $\boldsymbol{y} \in E$, jest $\|\boldsymbol{y}\|_{\boldsymbol{E}} \leqslant 1$. Chceme ukázat, že $\boldsymbol{x} \in \left(1 + \frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}}\right) \bullet E$, což je jinými slovy

$$\|\boldsymbol{x}\|_{\left(1+\frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}}\right)^2 \cdot \boldsymbol{E}} \leqslant 1.$$

Platí

$$\|\boldsymbol{x}\|_{\left(1+\frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}}\right)^2 \cdot \boldsymbol{E}} = \frac{1}{1+\frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}}} \cdot \|\boldsymbol{x}\|_{\boldsymbol{E}} \leqslant \frac{1}{1+\frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}}} \cdot \Big(\|\boldsymbol{x}-\boldsymbol{y}\|_{\boldsymbol{E}} + \underbrace{\|\boldsymbol{y}\|_{\boldsymbol{E}}}_{\leqslant 1}\Big)$$

$$\overset{\text{lemma 8(c)}}{\leqslant} \frac{1}{1+\frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}}} \cdot \Big(\underbrace{\frac{\|\boldsymbol{x}-\boldsymbol{y}\|}{\sqrt{\lambda(\boldsymbol{E})}}}_{\leqslant \frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}}} + 1\Big) \leqslant \frac{1}{1+\frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}}} \cdot \left(\frac{\epsilon}{\sqrt{\lambda(\boldsymbol{E})}} + 1\right) = 1.$$

První inkluse. Nechť $\boldsymbol{y}$ je bod na hranici elipsoidu $\left(1 + \frac{\epsilon}{\sqrt{\Lambda(\boldsymbol{E})}}\right) \bullet E$; ukážeme, že leží v $\mathcal{B}(E,\epsilon)$. Uvažme bod $\boldsymbol{x}$ na hranici elipsoidu E, který se při $k := (1 + \frac{\epsilon}{\sqrt{\Lambda(\boldsymbol{E})}})$-natahování elipsoidu E zobrazí na $\boldsymbol{y}$. (Body $\boldsymbol{y}$, $\boldsymbol{x}$ a střed elipsoidu leží na jedné přímce.) Pak $\boldsymbol{y} = k \cdot \boldsymbol{x}$ a

$$\|\boldsymbol{y}-\boldsymbol{x}\| = (k-1) \cdot \|\boldsymbol{x}\| \leqslant (k-1)\sqrt{\Lambda(\boldsymbol{E})} = \epsilon.$$

(Nerovnost $\|\boldsymbol{x}\| \leqslant \sqrt{\Lambda(\boldsymbol{E})}$ je jasná, neboť $\sqrt{\Lambda(\boldsymbol{E})}$ je délka nejdelší poloosy elipsoidu E.) Bod $\boldsymbol{x} \in E$ prokazuje, že $\boldsymbol{y} \in \mathcal{B}(E,\epsilon)$, neboť $\|\boldsymbol{y}-\boldsymbol{x}\| \leqslant \epsilon$.

Tvrzení (b) je zřejmé z toho, že determinant je součin vlastních čísel.

Tvrzení (c) je použití definice nafouknutí. Máme-li bod $\boldsymbol{z} \in \mathcal{B}(\mathcal{B}(E,\epsilon),\varphi)$, znamená to, že existuje bod $\boldsymbol{x} \in E$ a bod $\boldsymbol{y}$, které splňují $\|\boldsymbol{x}-\boldsymbol{y}\| \leqslant \epsilon$ a $\|\boldsymbol{y}-\boldsymbol{z}\| \leqslant \varphi$. Z trojúhelníhové nerovnosti plyne $\|\boldsymbol{x}-\boldsymbol{z}\| \leqslant \|\boldsymbol{x}-\boldsymbol{y}\| + \|\boldsymbol{y}-\boldsymbol{z}\| \leqslant \epsilon+\varphi$; odtud $\boldsymbol{z} \in \mathcal{B}(E,\epsilon+\varphi)$.

Tvrzení (d) je jasné z lemmatu 8(d) (str. 4). □

Zkoumejme, co se s elipsoidem stane, když místo přesných výpočtů (21) a (22) budeme počítat pouze na χ dvojkových míst za dvojkovou čárkou (říkejme tomu „zaokrouhlední na χ míst“). Již jsme konstatovali, že zaokrouhlením v (21) se může elipsoid posunout tak, že by mohlo dojít k porušení invariantu $\mathcal{F} \subseteq E_i$, a zaokrouhlením v (22) se elipsoid může zdeformovat. Míru škody způsobené zaokrouhlením kvantifikuje následující věta. Z tvrzení (a) se může zdát, že se elipsoid díky zaokrouhlení může poměrně dost v některém směru roztáhnout anebo smrštit; (b) však zaručuje, že vezmeme-li „zaokrouhlený“ elipsoid a trochu jej nafoukneme, budeme mít $\mathcal{F} \subseteq E_i$ zajištěno. Tvrzení (c) zaručuje, že zaokrouhlením se objem elipsoidu příliš nezvětší, takže nakonec budeme moci prokázat polynomiální počet iterací pomocí podobného argumentu jako na straně 16.

Ačkoliv pro účely důkazu $\mathbf{LP} \in \boldsymbol{P}$ je následující věta jen pomocným lemmatem, konstatujme, že má velký význam sama o sobě. Je to věta z numerické lineární algebry, oboru, který se (mimo jiné) zabývá dopadem malých změn v maticích na jejich vlastnosti.[5]

Konvence. Budeme-li hovořit nadále o zaokrouhlení reálného čísla x na χ (dvojkových) míst, máme na mysli, že spočteme libovolné racionální číslo ξ takové, že $|\xi - x| < 2^{-\chi}$. Hovoříme-li o zaokrouhlení čísel v symetrické matici, vždy vyčíslujeme její složky tak, aby výsledná matice byla symetrická.

Věta 23 (o zaokrouhlování)**.** *Buď* $E := \mathcal{E}(\boldsymbol{z},\boldsymbol{E}) \subseteq \mathbb{R}^n$, $0 < \varepsilon \leqslant 1$ *a*

$$\chi \geqslant 3n + |\log_2 \sqrt{\Lambda(\boldsymbol{E})}| + |\log_2 \lambda(\boldsymbol{E})| + \log_2 \tfrac{1}{\varepsilon}. \tag{24}$$

Nechť $\check{\boldsymbol{E}}$ *vznikne z* $\boldsymbol{E}$ *zaokrouhlením čísel v matici* $\boldsymbol{E}$ *na* χ *míst a nechť* $\check{\boldsymbol{z}}$ *vznikne zaokrouhlením čísel v* $\boldsymbol{z}$ *na* χ *míst. Položme* $\check{E} := \mathcal{E}(\check{\boldsymbol{z}},\check{\boldsymbol{E}})$*. Pak*

[5] Zaokrouhlování používáme v běžném životě velmi často, říkáme třeba, že $\pi = 3.14$, a intuitivně rozumíme tomu, že se nedopouštíme velkého zkreslení. Otázka, jak se změní vlastnosti matic, pokud čísla v nich změníme třeba na χ-tém místě za desetinnou čárkou, je však značně netriviální, a tato zdánlivě nevýznamná změna může mít významný dopad. Může se změnit hodnost, může se výrazně změnit inversní matice, vlastní čísla, vlastnosti zobrazení $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ nebo prostor řešení soustavy $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$. Jeden notoricky známý příklad (viz [**?**]) je

$$\boldsymbol{A}_1 = \begin{pmatrix} 2 & 6 \\ 2 & 6.00001 \end{pmatrix}, \quad \boldsymbol{b}_1 = \begin{pmatrix} 8 \\ 8.00001 \end{pmatrix},$$

$$\boldsymbol{A}_2 = \begin{pmatrix} 2 & 6 \\ 2 & 5.99999 \end{pmatrix}, \quad \boldsymbol{b}_2 = \begin{pmatrix} 8 \\ 8.00002 \end{pmatrix}.$$

Matice $\boldsymbol{A}_1$ a $\boldsymbol{A}_2$ se „téměř“ neliší, vektory $\boldsymbol{b}_1$ a $\boldsymbol{b}_2$ se rovněž „téměř“ neliší (zakrouhlením se veškeré rozdíly smažou), a přitom rovnice $\boldsymbol{A}_1\boldsymbol{x} = \boldsymbol{b}_1$ a $\boldsymbol{A}_2\boldsymbol{x} = \boldsymbol{b}_2$ mají odlišná řešení

$$\boldsymbol{A}_1^{-1}\boldsymbol{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{A}_2^{-1}\boldsymbol{b}_2 = \begin{pmatrix} 10 \\ -2 \end{pmatrix}.$$

Je to tím, že výpočet inversní matice je velmi numericky citlivý:

$$\boldsymbol{A}_1^{-1} = \begin{pmatrix} 300000.5 & -300000 \\ -100000 & 100000 \end{pmatrix}, \quad \boldsymbol{A}_2^{-1} = \begin{pmatrix} -300000 & 300000 \\ 100000 & -100000 \end{pmatrix}.$$

Všimněme si, že funkce $\frac{1}{\det \left(\begin{smallmatrix} 2 & 6 \\ 2 & x \end{smallmatrix}\right)}$ má uvnitř intervalu $[5.99999, 6.00001]$ singularitu.

(a) $\lambda(\check{\boldsymbol{E}}) \geqslant \frac{1}{2}\lambda(\boldsymbol{E})$ a $\Lambda(\check{\boldsymbol{E}}) \leqslant 2\Lambda(\boldsymbol{E})$; *speciálně,* $\check{\boldsymbol{E}}$ *je positivně definitní matice a* $\check{E}$ *je tudíž korektně definovaný elipsoid,*

(b) $E \subseteq \mathcal{B}(\check{E}, \varepsilon)$,

(c) $\frac{vol(\check{E})}{vol(E)} \leqslant 1 + \varepsilon$.

Důkaz (a). Všechna čísla v matici $\check{\boldsymbol{E}} - \boldsymbol{E}$ jsou díky zaokrouhlení v absolutní hodnotě $\leqslant 2^{-\chi}$. Vezmeme-li $\boldsymbol{x}$ s $\|\boldsymbol{x}\| = 1$, dostaneme

$$|\boldsymbol{x}^{\mathrm{T}}\check{\boldsymbol{E}}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x}| = |\boldsymbol{x}^{\mathrm{T}}(\check{\boldsymbol{E}} - \boldsymbol{E})\boldsymbol{x}| \leqslant n^2 2^{-\chi} \overset{(\spadesuit)}{\leqslant} \underbrace{\frac{n^2}{2^{3n}}}_{\leqslant \frac{1}{2}} \cdot \lambda(\boldsymbol{E}) \leqslant \frac{1}{2}\lambda(\boldsymbol{E}).^{6}$$

Rozeberme nerovnost (♠) podrobněji: z (24) jest

$$n^2 2^{-\chi} \leqslant n^2 2^{-(3n + |\log_2 \sqrt{\Lambda(\boldsymbol{E})}| + |\log_2 \lambda(\boldsymbol{E})| + \log_2(1/\varepsilon))} \leqslant \frac{n^2}{2^{3n}} \cdot 2^{-|\log_2 \lambda(\boldsymbol{E})|}.$$

Je-li $\lambda(\boldsymbol{E}) < 1$, pak $|\log_2 \lambda(\boldsymbol{E})| = -\log_2 \lambda(\boldsymbol{E})$ a nerovnost (♠) platí; je-li $\lambda(\boldsymbol{E}) \geqslant 1$, pak $|\log_2 \lambda(\boldsymbol{E})| = \log_2 \lambda(\boldsymbol{E})$ a dostáváme

$$\frac{n^2}{2^{3n}} \cdot 2^{-|\log_2 \lambda(\boldsymbol{E})|} = \frac{n^2}{2^{3n}} \cdot 2^{\log_2 \frac{1}{\lambda(\boldsymbol{E})}} = \frac{n^2}{2^{3n}} \cdot \frac{1}{\lambda(\boldsymbol{E})} \leqslant \frac{n^2}{2^{3n}}\lambda(\boldsymbol{E})$$

a nerovnost (♠) opět platí.

Z odhadu $\lambda(\boldsymbol{E}) \leqslant \boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x} \leqslant \Lambda(\boldsymbol{E})$ (viz důsledek 6, str. 4) a faktu, že $\boldsymbol{x}^{\mathrm{T}}\check{\boldsymbol{E}}\boldsymbol{x}$ a $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}\boldsymbol{x}$ se liší nanejvýš o $\frac{1}{2}\lambda(\boldsymbol{E})$, obdržíme

$$\lambda(\boldsymbol{E}) - \frac{1}{2}\lambda(\boldsymbol{E}) \leqslant \boldsymbol{x}^{\mathrm{T}}\check{\boldsymbol{E}}\boldsymbol{x} \leqslant \Lambda(\boldsymbol{E}) + \frac{1}{2}\lambda(\boldsymbol{E}),$$

z čehož plyne

$$\frac{1}{2}\lambda(\boldsymbol{E}) \leqslant \boldsymbol{x}^{\mathrm{T}}\check{\boldsymbol{E}}\boldsymbol{x} \leqslant 2\Lambda(\boldsymbol{E})$$

a (a) je dokázáno. Poznámka o positivní definitnosti je zřejmá z toho, že matice je i po zaokrouhlení symetrická a má kladná vlastní čísla.

(b). Budiž $\boldsymbol{D} := \boldsymbol{E}^{1/2}$.

Pozorování A. Všechna vlastní čísla matice $\boldsymbol{D}^{-1}\check{\boldsymbol{E}}\boldsymbol{D}^{-1}$ *se od jedničky liší nanejvýš o* $\frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}$. *Speciálně, matice* $\boldsymbol{D}^{-1}\check{\boldsymbol{E}}\boldsymbol{D}^{-1}$ *je positivně definitní. Důkaz.* Vezměme $\boldsymbol{x}$ s $\|\boldsymbol{x}\| = 1$. Zřejmě

$$\|\boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}^{-1}\| \overset{\text{lemma 8(b)}}{=} \|\boldsymbol{x}\|_{\boldsymbol{E}} \overset{\text{lemma 8(c)}}{\leqslant} \frac{1}{\sqrt{\lambda(\boldsymbol{E})}}.$$

Pak

$$\begin{aligned} |\boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}^{-1}\check{\boldsymbol{E}}\boldsymbol{D}^{-1}\boldsymbol{x} \;-\; 1| &= |\boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}^{-1}\check{\boldsymbol{E}}\boldsymbol{D}^{-1}\boldsymbol{x} - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}^{-1}\boldsymbol{E}\boldsymbol{D}^{-1}\boldsymbol{x}| \\ &= |\underbrace{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}^{-1}}_{\|\boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}^{-1}\| \leqslant \frac{1}{\sqrt{\lambda(\boldsymbol{E})}}} \cdot \underbrace{(\check{\boldsymbol{E}} - \boldsymbol{E})}_{|(\check{\boldsymbol{E}} - \boldsymbol{E})_{ij}| \leqslant 2^{-\chi}} \cdot \underbrace{\boldsymbol{D}^{-1}\boldsymbol{x}}_{\|\boldsymbol{D}^{-1}\boldsymbol{x}\| \leqslant \frac{1}{\sqrt{\lambda(\boldsymbol{E})}}}| \leqslant n^2 2^{-\chi} \cdot \frac{1}{\lambda(\boldsymbol{E})}, \end{aligned}$$

[6] Odhad $n^2 2^{-\chi}$ by bylo snadné zlepšit na $n \cdot 2^{-\chi}$.

čímž je pozorování dokázáno.

Pozorování B. $\mathcal{E}(\boldsymbol{z},\boldsymbol{E}) \subseteq \left(1+\frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}\right) \bullet \mathcal{E}(\boldsymbol{z},\check{\boldsymbol{E}})$. *Důkaz.* Bez újmy na obecnosti nechť je $\boldsymbol{z}=\boldsymbol{0}$. Vezměme $\boldsymbol{x} \in \mathcal{E}(\boldsymbol{0},\boldsymbol{E})$; odtud $\|\boldsymbol{x}\|_{\boldsymbol{E}} = \|\boldsymbol{D}^{-1}\boldsymbol{x}\| \leqslant 1$ (lemma 8(a) a (b)).

$$\|\boldsymbol{x}\|^2_{\check{\boldsymbol{E}}} = \boldsymbol{x}^{\mathrm{T}}\check{\boldsymbol{E}}^{-1}\boldsymbol{x} = \underbrace{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}^{-1}}_{\|\boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}^{-1}\|\leqslant 1} (\boldsymbol{D}\check{\boldsymbol{E}}^{-1}\boldsymbol{D}) \underbrace{\boldsymbol{D}^{-1}\boldsymbol{x}}_{\|\boldsymbol{D}^{-1}\boldsymbol{x}\|\leqslant 1} \overset{\substack{\text{pozorování A,}\\ \text{důsl. 6}}}{\leqslant}{}^{7} 1+\frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})} \leqslant \left(1+\frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}\right)^2,$$

takže $\|\boldsymbol{x}\|_{\check{\boldsymbol{E}}} \leqslant 1+\frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}$. Elipsoid $\mathcal{E}(\boldsymbol{0},\check{\boldsymbol{E}})$ tudíž stačí $\left(1+\frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}\right)$-nafouknout: $\boldsymbol{x} \in \left(1+\frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}\right)\bullet \mathcal{E}(\boldsymbol{0},\check{\boldsymbol{E}})$ a pozorování B je dokázáno.

Použijeme pozorování 22, konkrétně první inklusi v (23) s $\epsilon := \frac{\sqrt{\Lambda(\check{\boldsymbol{E}})}\cdot n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}$ (symbolem ϵ myslíme ϵ vyskytující se ve vztahu (23), nikoliv ε z předpokladu této věty); dostaneme tak

$$\begin{aligned}\left(1+\frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}\right)\bullet \mathcal{E}(\boldsymbol{z},\check{\boldsymbol{E}}) &= \left(1+\tfrac{\epsilon}{\sqrt{\Lambda(\check{\boldsymbol{E}})}}\right)\bullet \mathcal{E}(\boldsymbol{z},\check{\boldsymbol{E}}) \\ &\overset{(23)}{\subseteq} \mathcal{B}(\mathcal{E}(\boldsymbol{z},\check{\boldsymbol{E}}), \tfrac{\sqrt{\Lambda(\check{\boldsymbol{E}})}\cdot n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}) \overset{(a)}{\subseteq} \mathcal{B}(\mathcal{E}(\boldsymbol{z},\check{\boldsymbol{E}}), \tfrac{\sqrt{2\Lambda(\boldsymbol{E})}\cdot n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}).\end{aligned} \tag{25}$$

Dále si uvědomme, že zaokrouhlením $\boldsymbol{z}$ na $\check{\boldsymbol{z}}$ máme $\|\check{\boldsymbol{z}}-\boldsymbol{z}\| \leqslant n\cdot 2^{-\chi}$ (zaokrouhlení $\boldsymbol{z}$ způsobí „malé posunutí elipsoidu“), pročež musíme ještě trochu nafouknout:

$$\mathcal{B}(\mathcal{E}(\boldsymbol{z},\check{\boldsymbol{E}}), \tfrac{\sqrt{2\Lambda(\boldsymbol{E})}n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}) \subseteq \mathcal{B}(\mathcal{E}(\check{\boldsymbol{z}},\check{\boldsymbol{E}}), \tfrac{\sqrt{2\Lambda(\boldsymbol{E})}n^2 2^{-\chi}}{\lambda(\boldsymbol{E})} + n2^{-\chi}). \tag{26}$$

Konečně se nahlédne, že z (24) je[8] $\varepsilon \geqslant \frac{\sqrt{2\Lambda(\boldsymbol{E})}n^2 2^{-\chi}}{\lambda(\boldsymbol{E})} + n2^{-\chi}$, takže

$$\mathcal{B}(\mathcal{E}(z,\check{\boldsymbol{E}}), \tfrac{\sqrt{2\Lambda(\boldsymbol{E})}n^2 2^{-\chi}}{\lambda(\boldsymbol{E})} + n2^{-\chi}) \subseteq \mathcal{B}(\mathcal{E}(z,\check{\boldsymbol{E}}),\varepsilon). \tag{27}$$

Zřetězením tvrzení pozorování B, (25), (26) a (27) obdržíme $\mathcal{E}(\boldsymbol{z},\boldsymbol{E}) \subseteq \mathcal{B}(\mathcal{E}(\check{\boldsymbol{z}},\check{\boldsymbol{E}}),\varepsilon)$ a (b) je dokázáno.

(c).

$$\frac{vol(\check{E})}{vol(E)} = \sqrt{\frac{\det\check{\boldsymbol{E}}}{\det\boldsymbol{E}}} = \sqrt{\det\boldsymbol{D}^{-1}\check{\boldsymbol{E}}\boldsymbol{D}^{-1}} \overset{\text{pozorování A}}{\leqslant} \left(1+\frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}\right)^{\frac{n}{2}} \overset{(\clubsuit)}{\leqslant} 1+\varepsilon.$$

[7] Důsledek tvrdí, že pro $\|\boldsymbol{y}\|=1$ jest $\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{D}\check{\boldsymbol{E}}^{-1}\boldsymbol{D})\boldsymbol{y} \leqslant \Lambda(\boldsymbol{D}\check{\boldsymbol{E}}^{-1}\boldsymbol{D})$ (podle pozorování A je matice $\boldsymbol{D}\check{\boldsymbol{E}}^{-1}\boldsymbol{D}$ positivně definitní); odtud plyne, že pro libovolné $\boldsymbol{y}$ splňující $\|\boldsymbol{y}\| \leqslant 1$ (a tedy i pro $\boldsymbol{y} := \boldsymbol{D}^{-1}\boldsymbol{x}$) je také $\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{D}\check{\boldsymbol{E}}^{-1}\boldsymbol{D})\boldsymbol{y} \leqslant \Lambda(\boldsymbol{D}\check{\boldsymbol{E}}^{-1}\boldsymbol{D})$. Podle pozorování A je $\Lambda(\boldsymbol{D}\check{\boldsymbol{E}}^{-1}\boldsymbol{D}) \leqslant 1+\frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}$.

[8] Dokažme tvrzení $\varepsilon \geqslant \frac{\sqrt{2\Lambda(\boldsymbol{E})}n^2 2^{-\chi}}{\lambda(\boldsymbol{E})} + n2^{-\chi}$ podrobněji, konkrétně (zajímavý) případ s $\lambda(\boldsymbol{E})<1$ a $\Lambda(\boldsymbol{E})>1$ (je-li někde opačná nerovnost, úvaha je velmi podobná). V tomto případě je $|\log_2\sqrt{\Lambda(\boldsymbol{E})}| = \log_2\sqrt{\Lambda(\boldsymbol{E})}$ a $|\log_2\lambda(\boldsymbol{E})| = -\log_2\lambda(\boldsymbol{E})$, takže definice (24) má tvar $\chi \geqslant 3n+\log_2\sqrt{\Lambda(\boldsymbol{E})}-\log_2\lambda(\boldsymbol{E})-\log_2\varepsilon$, odkud

$$\varepsilon \geqslant 2^{3n-\chi+\log_2\sqrt{\Lambda(\boldsymbol{E})}-\log_2\lambda(\boldsymbol{E})} = 2^{3n}\cdot 2^{-\chi}\cdot\frac{\sqrt{\Lambda(\boldsymbol{E})}}{\lambda(\boldsymbol{E})} \overset{2^{3n}\geqslant 3n^2}{\geqslant} 2^{-\chi}\frac{\sqrt{\Lambda(\boldsymbol{E})}}{\lambda(\boldsymbol{E})}\cdot 3n^2$$

$$\overset{3\geqslant 2\sqrt{2}}{\geqslant} 2^{-\chi}\Big(\sqrt{2}\frac{\sqrt{\Lambda(\boldsymbol{E})}}{\lambda(\boldsymbol{E})}n^2 + \underbrace{\sqrt{2}\frac{\sqrt{\Lambda(\boldsymbol{E})}}{\lambda(\boldsymbol{E})}}_{\geqslant 1}\underbrace{n^2}_{\geqslant n}\Big) \geqslant \frac{\sqrt{2\Lambda(\boldsymbol{E})}n^2 2^{-\chi}}{\lambda(\boldsymbol{E})} + n2^{-\chi}.$$

Využili jsme toho, že determinant je součin vlastních čísel; zbývá dokázat nerovnost (♣). Řekněme, že $\lambda(\boldsymbol{E}) < 1$, a tedy výraz $|\log_2 \lambda(\boldsymbol{E})|$ v definici χ lze psát jako $-\log_2 \lambda(\boldsymbol{E})$ (zajímavý je právě případ s malým $\lambda(\boldsymbol{E})$, nicméně případ $\lambda(\boldsymbol{E}) \geqslant 1$ by se rozebral podobně). Nerovnost (24) tak dostane tvar

$$\chi \geqslant 3n + |\log_2 \sqrt{\Lambda(\boldsymbol{E})}| - \log_2 \lambda(\boldsymbol{E}) - \log_2 \varepsilon. \tag{28}$$

S pomocí binomické věty obdržíme odhad[9]

$$\begin{aligned}
\left(1 + \frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}\right)^{\frac{n}{2}} &\leqslant \left(1 + \frac{n^2 2^{-\chi}}{\lambda(\boldsymbol{E})}\right)^{n} = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{n^2 \cdot 2^{-\chi}}{\lambda(\boldsymbol{E})}\right)^{k} \\
&= 1 + \sum_{k=1}^{n} \binom{n}{k} \left(\frac{n^2 \cdot 2^{-\chi}}{\lambda(\boldsymbol{E})}\right)^{k} = 1 + \sum_{k=1}^{n} \binom{n}{k} 2^{2k \log_2 n - k \log_2 \lambda(\boldsymbol{E}) - k\chi} \\
&\overset{(28)}{\leqslant} 1 + \sum_{k=1}^{n} \binom{n}{k} 2^{2k \log_2 n - k \log_2 \lambda(\boldsymbol{E}) - 3kn - k|\log_2 \sqrt{\Lambda(\boldsymbol{E})}| + k \log_2 \lambda(\boldsymbol{E}) + k \log_2 \varepsilon} \\
&\leqslant \cdots
\end{aligned}$$

(poslední výraz zvětšíme vypuštěním členu $2k \log_2 n - 2kn$ [ten je $\leqslant 0$] a členu $-k|\log_2 \sqrt{\Lambda(\boldsymbol{E})}|$ [ten je též $\leqslant 0$] v exponentu a pokračujeme v psaní nerovnosti)

$$\begin{aligned}
\cdots \leqslant 1 + \sum_{k=1}^{n} \binom{n}{k} 2^{-kn + k \log_2 \varepsilon} = 1 + \sum_{k=1}^{n} \binom{n}{k} 2^{-kn} \varepsilon^k &\overset{\varepsilon \leqslant 1}{\leqslant} 1 + \sum_{k=1}^{n} \binom{n}{k} 2^{-kn} \varepsilon \\
&= 1 + \varepsilon \cdot \Big[\underbrace{n \cdot 2^{-n}}_{\leqslant \frac{1}{2}} + \underbrace{\sum_{k=2}^{n} \underbrace{\binom{n}{k}}_{\leqslant 2^n} 2^{-kn}}_{\leqslant (1/2)^k} \Big] \leqslant 1 + \varepsilon;
\end{aligned}$$

poslední hranatou závorku jsme odhadli začátkem geometrické řady, jejíž součet je 1. □

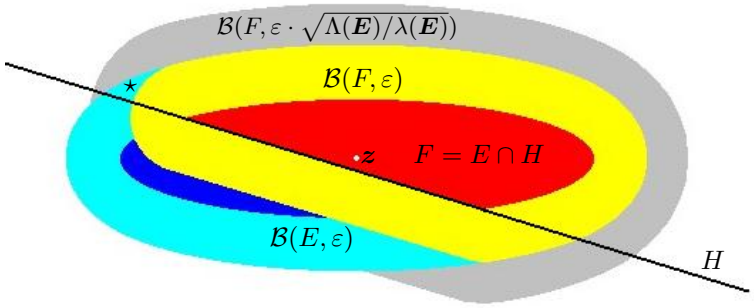
Význam následujícího lemmatu ilustrujme obrázkem 4. Lemma popisuje, co se stane, když poloprostorem H uřízneme část množiny $\mathcal{B}(E, \varepsilon)$, kde E je elipsoid a $\varepsilon > 0$. Nechť v H leží střed elipsoidu E; označme $F := E \cap H$ část elipsoidu a $Z := \mathcal{B}(E, \varepsilon) \cap H$ část množiny $\mathcal{B}(E, \varepsilon)$, které nám zůstaly po uříznutí. Ptáme se, zdali ε-nafouknutí F nám pokryje celou množinu Z. Prostor označený hvězdičkou v obrázku ukazuje, že odpověď na tuto otázku je záporná; chceme-li nafouknutím F pokrýt celou množinu Z, musíme nafukovat o více než o ε. O kolik, to říká následující lemma.

Lemma 24. *Budiž $E := \mathcal{E}(\boldsymbol{z}, \boldsymbol{E})$, H poloprostor (tj. podmnožina $\mathbb{R}^n$ tvaru $\{\boldsymbol{x} : \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} \leqslant b\}$) obsahující bod $\boldsymbol{z}$ a $\varepsilon \geqslant 0$. Pak*

$$\mathcal{B}(E, \varepsilon) \cap H \subseteq \mathcal{B}(E \cap H, \varepsilon \cdot \sqrt{\tfrac{\Lambda(\boldsymbol{E})}{\lambda(\boldsymbol{E})}}).$$

Důkaz. Bez újmy na obecnosti $\boldsymbol{z} = \boldsymbol{0}$. Buď $\boldsymbol{x} \in \mathcal{B}(E, \varepsilon) \cap H$. Je-li i $\boldsymbol{x} \in E$, je situace triviální. Zajímavý případ je $\boldsymbol{x} \notin E$. Pak z faktu $\boldsymbol{x} \in \mathcal{B}(E, \varepsilon)$ podle druhé inkluse v (23), str. 17 plyne

[9] Pro jistotu připomínáme obvyklou konvenci o prioritě oprátorů: například $\log a + b = b + \log a$, nikoliv $\log(a+b)$.



Obrázek 4: *K lemmatu 24.*

$\boldsymbol{x} \in \left(1 + \frac{\varepsilon}{\sqrt{\lambda(\boldsymbol{E})}}\right) \bullet E$, odkud

$$\|\boldsymbol{x}\|_{\boldsymbol{E}} \leqslant 1 + \frac{\varepsilon}{\sqrt{\lambda(\boldsymbol{E})}} \tag{29}$$

podle lemmatu 22(d).

Položme $\boldsymbol{v} := \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|_{\boldsymbol{E}}}$. Dokážeme, že $\boldsymbol{v} \in E \cap H$. Jest $\boldsymbol{v} \in H$, protože $\frac{1}{\|\boldsymbol{x}\|_{\boldsymbol{E}}} < 1$ ($\boldsymbol{x} \notin E$ totiž znamená, že $\|\boldsymbol{x}\|_{\boldsymbol{E}} > 1$), a tudíž $\boldsymbol{v}$ leží na úsečce spojující $\boldsymbol{0}$ (střed E) a $\boldsymbol{x}$, a celá tato úsečka je v H. Že $\boldsymbol{v} \in E$, je jasné díky tomu, že $\|\boldsymbol{v}\|_{\boldsymbol{E}} = 1$, a jednotkovou $\boldsymbol{E}$-normu mají právě body, které leží na hranici E.

Dále platí odhad

$$\|\boldsymbol{x} - \boldsymbol{v}\| \overset{\text{definice } v}{=} \left(1 - \frac{1}{\|\boldsymbol{x}\|_{\boldsymbol{E}}}\right) \cdot \|\boldsymbol{x}\| = \underbrace{(\|\boldsymbol{x}\|_{\boldsymbol{E}} - 1)}_{\leqslant \frac{\varepsilon}{\sqrt{\lambda(\boldsymbol{E})}} \text{ podle (29)}} \cdot \underbrace{\frac{\|\boldsymbol{x}\|}{\|\boldsymbol{x}\|_{\boldsymbol{E}}}}_{\leqslant \sqrt{\Lambda(\boldsymbol{E})} \text{ [lem. 8(c)]}} \leqslant \varepsilon \cdot \sqrt{\frac{\Lambda(\boldsymbol{E})}{\lambda(\boldsymbol{E})}}.$$

Ukázali jsme, že $\boldsymbol{v} \in E$ a $\|\boldsymbol{x} - \boldsymbol{v}\| \leqslant \varepsilon \cdot \sqrt{\frac{\Lambda(\boldsymbol{E})}{\lambda(\boldsymbol{E})}}$, a tak $\boldsymbol{x} \in \mathcal{B}(E \cap H, \varepsilon \cdot \sqrt{\frac{\Lambda(\boldsymbol{E})}{\lambda(\boldsymbol{E})}})$. □

10 Elipsoidová metoda (Chačijanova verse)

Chačijanův algoritmus je elipsoidový algoritmus se zaokrouhlováním. Skoro vše je stejné jako v kapitole 7 — máme dánu celočíselnou matici $\boldsymbol{A}$ a celočíselný vektor $\boldsymbol{b}$ a bez újmy na obecnosti předpokládáme, že polyedr $\mathcal{F} = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$ je omezený a buďto prázdný anebo má plnou dimensi.

Modifikujeme elipsoidový algoritmus z kapitoly 7 takto: položíme

$$N := 128n^3L, \quad \chi := 5N^2$$

a při výpočtu (21) a (22) instrukce přiřazení „$\alpha := \beta$“ nechť znamená „α vznikne zaokrouhlením β na χ míst“ (připomeňme zde konvenci o zaokrouhlování ze str. 19). Věci se tím zkomplikují: nemáme obecně zaručeno, že $\mathcal{F} \subseteq E_i$, protože zaokrouhlením čísel ve vektoru $\boldsymbol{z}$ a matici $\boldsymbol{E}$ dojde k jistému zkreslení. Tvrzení (c) následujícího lemmatu říká, že toto zkreslení není, alespoň během prvních N iterací, velké. Konečně věta 26 zaručuje, že zaokrouhlování nevadí.

V části 7 nám k rozhodnutí stačilo $N' := 64n^3L$ iterací; uvidíme, že za přítomnosti zaokrouhlování postačí $N = 2N'$ iterací.

Až dokážeme tvrzení (a) věty 26, budeme mít zajištěno, že aritmetické operace se budou provádět pouze s polynomiálně dlouhými čísly (zaokrouhlením jsme zajistili, že za dvojkovou čárkou budou mít čísla jen omezený počet míst; dále ještě dokážeme, že *čísla nebudou příliš narůstat ani před dvojkovou čárkou*), pročež provedení každé iterace trvá jen polynomiální čas.

Základní metodou následujících důkazů je indukce. Tvrzení, která dokážeme, jsou odhady na chyby vznikající zaokrouhlováním. Mohlo by se stát, že ačkoliv se zaokrouhlením v jednom kroku dopustíme jen malé chyby, další zaokrouhlování v dalších iteracích by mohlo původně malou chybu zvětšovat. Učiníme odhad na velikost chyby v jednom kroku a indukcí tento odhad přeneseme do kroků dalších, čímž budeme mít chybu pod kontrolou. Nebudeme sice trvat přímo na $\mathcal{F} \subseteq E_i$ jako v části 7, ale ukážeme, že při zaokrouhlování se zachovává invariant $\mathcal{F} \subseteq 2 \bullet E_i$.

Důkazy, ačkoliv dost technické, se opírají o základní nápad: představme si, že máme z minulé iterace spočten elipsoid E_{i-1}; nejprve zjistíme, jak by vypadal elipsoid $\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true})$ — to je elipsoid, který by vznikl rozříznutím E_{i-1}, *kdybychom uměli počítat s reálnými čísly přesně, bez zaokrouhlování*. Poté v druhém kroku odhadneme, co se s elisposidem $\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true})$ stane díky zaokrouhlení na χ míst. Zaokrouhlíme-li číslo ξ na χ míst za dvojkovou čárkou, změní se nanejvýš o hodnotu $2^{-\chi}$. To je základní odhad.

Máme-li na začátku i-té iterace z minulé iterace elipsoid $E_{i-1} = \mathcal{E}(\boldsymbol{z}_{i-1}, \boldsymbol{E}_{i-1})$, označme $\boldsymbol{z}^{true}$ a $\boldsymbol{E}^{true}$ přesné hodnoty dané vzorci (21) a (22); $\boldsymbol{z}_i$ a $\boldsymbol{E}_i$ vzniknou ze $\boldsymbol{z}^{true}$ a $\boldsymbol{E}^{true}$ zaokrouhlením na χ dvojkových míst. Samozřejmě se $\boldsymbol{z}^{true}$ a $\boldsymbol{E}^{true}$ nikde nevyčíslují algoritmicky; potřebujeme je jen pro analýzu.

Stejně jako v kapitole 7 označme poloměr koule E_0

$$R := 2^{3nL}.$$

Všimněme si, že se může stát, že ačkoliv se objem elipsoidů zmenšuje, délka nejdelší poloosy roste. Krájíme-li totiž elipsoidy opakovaně ve stejném směru, může pak elipsoid připomínat tenčící a prodlužující se špejli. V následujícím lemmatu podáme na prodlužování nejdelší poloosy alespoň hrubý odhad. Odhadneme též tempo zkracování nejkratší poloosy.

Lemma 25. *Buď $i \leqslant N$.*

(a) $\lambda(\boldsymbol{E}_i) \geqslant 8^{-i}R^2$, $\Lambda(\boldsymbol{E}_i) \leqslant 8^i R^2$.

(b) $\|\boldsymbol{z}_i\| \leqslant 8^i R$.

(c) $\mathcal{F} \subseteq \mathcal{B}(E_i, 2^{-2N-4N^2+4Ni})$.

Důkaz (a). Postupujme indukcí; zřejmě pro $i = 0$ je tvrzení pravda, neboť $\mathcal{F}$ je uvnitř koule E_0, jež má poloměr R. Máme-li dokázáno pro $i-1$, dokažme tvrzení pro i.

Budeme používat větu 23(a) s volbou $\varepsilon = 1$. Ověřme, že předpoklad věty je pro elipsoid E_{i-1} splněn: využijeme faktu $i - 1 \leqslant N$ a platnosti tvrzení (a) tohoto lemmatu pro $i - 1$. Máme ukázat, že platí nerovnost (✠) ve vztahu

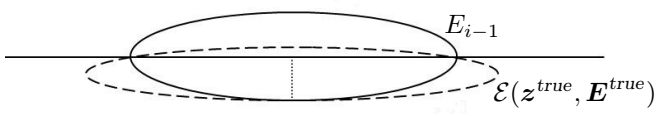
$$5N^2 = \chi \overset{(✠)}{\geqslant} \overbrace{\underbrace{3n}_{\leqslant 3NnL} + \underbrace{|\log_2 \sqrt{\Lambda(\boldsymbol{E}_{i-1})}|}_{\leqslant \frac{1}{2}\log_2(8^N 2^{6nL}) \leqslant 5NnL} + \underbrace{|\log_2 \lambda(\boldsymbol{E}_{i-1})|}_{\leqslant \log_2(8^N 2^{6nL}) \leqslant 9NnL}}^{=:Q} + \underbrace{\log_2 \tfrac{1}{\varepsilon}}_{=0}. \tag{30}$$

Z naznačených odhadů je $Q \leqslant 17NnL \leqslant 5N \cdot 4nL \leqslant 5N^2$, protože z definice N jest $4nL \leqslant 128n^3L = N$. Tím je předpoklad věty 23 ověřen.

Přistupme k indukci. Elipsoid $\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true})$ obsahuje přinejmenším polovinu elipsoidu E_{i-1}; z geometrického názoru, jejž ilustruje obrázek 5, se nahlédne, že jeho nejkratší poloosa nemůže být menší než polovina nejkratší poloosy elipsoidu E_{i-1}. Tedy

$$\sqrt{\lambda(\boldsymbol{E}^{true})} \geqslant \frac{1}{2}\sqrt{\lambda(\boldsymbol{E}_{i-1})} \geqslant \frac{1}{2}\sqrt{8^{-(i-1)}R^2} = \sqrt{2 \cdot 8^{-i}R^2},$$

a odtud $\lambda(\boldsymbol{E}^{true}) \geqslant 2 \cdot 8^{-i}R^2$.



Obrázek 5: *Intuice, jak se nejvíce může zkrátit nejkratší poloosa.*

Z tvrzení věty 23 (a) (připomeňme: zaokrouhlením $\boldsymbol{E}^{true}$ na $\boldsymbol{E}_i$ se nejmenší vlastní číslo zmenší přinejhorším na polovinu) je $\lambda(\boldsymbol{E}_i) \geqslant 8^{-i}R^2$ a první odhad v tvrzení (a) tohoto lemmatu je pro i dokázán.

Druhé tvrzení v (a). Z kroku $i-1$ víme, že $\Lambda(\boldsymbol{E}_{i-1}) \leqslant 8^{i-1}R^2$. Ze vztahu (22) s užitím lemmatu 5 (str. 4) získáme odhad

$$\begin{aligned}
\Lambda(\boldsymbol{E}^{true}) &= \max_{\|\boldsymbol{x}\|=1} \boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}^{true}\boldsymbol{x} \\
&= \max_{\|\boldsymbol{x}\|=1} \boldsymbol{x}^{\mathrm{T}}\left(\frac{n^2}{n^2-1}\boldsymbol{E}_{i-1} - \frac{2n^2}{(n^2-1)(n+1)} \cdot \frac{\boldsymbol{E}_{i-1}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}\boldsymbol{a}}\right)\boldsymbol{x} \\
&\leqslant \underbrace{\frac{n^2}{n^2-1}}_{\leqslant 2} \underbrace{\max_{\|\boldsymbol{x}\|=1}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}_{i-1}\boldsymbol{x})}_{=\Lambda(\boldsymbol{E}_{i-1})} + \underbrace{\frac{2n^2}{(n^2-1)(n+1)}}_{\leqslant 2} \max_{\|\boldsymbol{x}\|=1} \boldsymbol{x}^{\mathrm{T}}\frac{\boldsymbol{E}_{i-1}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}\boldsymbol{a}}\boldsymbol{x} \\
&\leqslant 2\Lambda(\boldsymbol{E}_{i-1}) + 2\Lambda\left(\frac{\boldsymbol{E}_{i-1}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}\boldsymbol{a}}\right) \overset{(\clubsuit)}{\leqslant} 2\Lambda(\boldsymbol{E}_{i-1}) + 2\Lambda(\boldsymbol{E}_{i-1}) \\
&= 4\Lambda(\boldsymbol{E}_{i-1}) \leqslant 4 \cdot 8^{i-1}R^2 = \frac{1}{2}8^iR^2.
\end{aligned}$$

Věta 23(a) říká, že zaokrouhlením se největší vlastní číslo nanejvýš zdvojnásobí, takže druhá nerovnost tvrzení (a) je dokázána.

Zbývá ještě prokázat nerovnost (♣). Označme $\boldsymbol{D} := \boldsymbol{E}_{i-1}^{1/2}$. Z (4) (str. 4) víme, že pro libovolné $\boldsymbol{x}$ platí

$$\|\boldsymbol{D}\boldsymbol{x}\| = \sqrt{\boldsymbol{x}^{\mathrm{T}}\boldsymbol{E}_{i-1}\boldsymbol{x}} \leqslant \Lambda(\boldsymbol{D}) \cdot \|\boldsymbol{x}\|. \tag{31}$$

Označíme-li $\boldsymbol{\beta} := \boldsymbol{E}_{i-1}\boldsymbol{a}$, pak $\frac{\boldsymbol{E}_{i-1}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}\boldsymbol{a}} = \frac{\boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}}}{\|\boldsymbol{D}\boldsymbol{a}\|^2}$. Tvrdíme, že

$$\Lambda(\boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}}) = \|\boldsymbol{E}_{i-1}\boldsymbol{a}\|^2. \tag{32}$$

Až to dokážeme, zavedeme $\boldsymbol{c} := \boldsymbol{D}\boldsymbol{a}$ a obdržíme

$$\begin{aligned}
\Lambda\left(\frac{\boldsymbol{E}_{i-1}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{E}_{i-1}\boldsymbol{a}}\right) &= \frac{\Lambda(\boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}})}{\|\boldsymbol{D}\boldsymbol{a}\|^2} = \left(\frac{\|\boldsymbol{E}_{i-1}\boldsymbol{a}\|}{\|\boldsymbol{D}\boldsymbol{a}\|}\right)^2 = \left(\frac{\|\boldsymbol{D}\cdot\boldsymbol{D}\boldsymbol{a}\|}{\|\boldsymbol{D}\boldsymbol{a}\|}\right)^2 \\
&= \left(\frac{\|\boldsymbol{D}\boldsymbol{c}\|}{\|\boldsymbol{c}\|}\right)^2 \overset{(31)}{\leqslant} \left(\frac{\Lambda(\boldsymbol{D})\cdot\|\boldsymbol{c}\|}{\|\boldsymbol{c}\|}\right)^2 = (\Lambda(\boldsymbol{D}))^2 = \Lambda(\boldsymbol{E}_{i-1}),
\end{aligned}$$

čímž je nerovnost (♣) dokázána.

Zbývá ukázat (32). Bez újmy na obecnosti předpokládejme, že $\boldsymbol{\beta}$ je tvaru $(\beta, 0, \ldots, 0)^{\mathrm{T}}$, kde $\beta \neq 0$. Toho lze dosáhnout vhodnou rotací prostoru, a rotace nemění délky poloos elipsoidů (= vlastní čísla) ani eukleidovskou normu. Pak matice $\boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}}$ obsahuje jediné nenulové číslo β^2 na posici $(1,1)$. Je tedy β^2 jejím jediným kladným, a tudíž největším vlastním číslem: $\Lambda(\boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}}) = \beta^2$. Ovšem $\|\boldsymbol{\beta}\| = \|\boldsymbol{E}_{i-1}\boldsymbol{a}\|$, a ježto $\|\boldsymbol{\beta}\| = \beta$, jest $\beta = \|\boldsymbol{E}_{i-1}\boldsymbol{a}\|$, a tedy $\Lambda(\boldsymbol{\beta}\boldsymbol{\beta}^{\mathrm{T}}) = \beta^2 = \|\boldsymbol{E}_{i-1}\boldsymbol{a}\|^2$. Tím je důkaz (32) dokončen.

(b). Toto je triviálně pravda pro $\boldsymbol{z}_0 = \mathbf{0}$; dále postupujeme indukcí. Cestu ze $\boldsymbol{z}_{i-1}$ do $\boldsymbol{z}_i$ si můžeme představit tak, že nejprve uděláme krok do $\boldsymbol{z}^{true}$ a pak zaokrouhlením druhý krok do $\boldsymbol{z}_i$:

$$\begin{aligned}\|\boldsymbol{z}_i\| &= \|\boldsymbol{z}_i - \boldsymbol{z}^{true} + \boldsymbol{z}^{true} - \boldsymbol{z}_{i-1} + \boldsymbol{z}_{i-1}\| \\ &\leqslant \|\boldsymbol{z}_i - \boldsymbol{z}^{true}\| + \|\boldsymbol{z}^{true} - \boldsymbol{z}_{i-1}\| + \|\boldsymbol{z}_{i-1}\| \leqslant n \cdot 2^{-\chi} + 8^{\frac{i-1}{2}} R + 8^{i-1} R \leqslant 8^i R.\end{aligned}$$

Použili jsme toho, že $\|\boldsymbol{z}_i - \boldsymbol{z}^{true}\| \leqslant n \cdot 2^{-\chi}$ díky zaokrouhlení, $\|\boldsymbol{z}_{i-1}\| \leqslant 8^{i-1}R$ z indukčního předpokladu a $\|\boldsymbol{z}^{true} - \boldsymbol{z}_{i-1}\| \leqslant 8^{\frac{i-1}{2}} R$, což plyne z toho, že bod $\boldsymbol{z}^{true}$ leží uvnitř elipsoidu E_{i-1} (viz důsledek 11), pročež jeho eukleidovská vzdálenost od středu E_{i-1} nemůže být větší než nejdelší poloosa E_{i-1}, a ta má délku nanejvýš $\sqrt{8^{i-1}R^2}$ (podle (a)).

(c). Opět indukcí. Pro $i = 0$ je tvrzení zřejmé, neboť E_0 jsme volili tak, že $\mathcal{F} \subseteq E_0$. Označme řezný poloprostor, kterým půlíme elipsoid E_{i-1}, jako H (viz lemma 10):

$$H := \{\boldsymbol{x} : (\boldsymbol{a} - \boldsymbol{z}_{i-1})^{\mathrm{T}} \boldsymbol{x} \leqslant 0\}.$$

Položme dále

$$\varepsilon_i := 2^{-2N-4N^2+4Ni}. \tag{33}$$

Máme-li dokázáno pro $i-1$, platí posloupnost inklusí

$$\begin{aligned}\mathcal{F} &\overset{\text{(i)}}{\subseteq} \mathcal{B}(E_{i-1}, \varepsilon_{i-1}) \cap H \\ &\overset{\text{(ii)}}{\subseteq} \mathcal{B}(E_{i-1} \cap H, \varepsilon_{i-1} \cdot \sqrt{\tfrac{8^{i-1}}{8^{-(i-1)}}}) \\ &= \mathcal{B}(E_{i-1} \cap H, \varepsilon_{i-1} \cdot 8^{i-1}) \\ &\overset{\text{(iii)}}{\subseteq} \mathcal{B}(\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true}), \varepsilon_{i-1} \cdot 8^{i-1}) \\ &\overset{\text{(iv)}}{\subseteq} \mathcal{B}(E_i, \varepsilon_{i-1} \cdot 8^{i-1} + 2^{-2N-4N^2}) \\ &= \mathcal{B}(E_i, 2^{-2N-4N^2+4N(i-1)+3(i-1)} + 2^{-2N-4N^2}) \\ &\overset{\text{(v)}}{\subseteq} \mathcal{B}(E_i, 2^{-2N-4N^2+4Ni}) = \mathcal{B}(E_i, \varepsilon_i).\end{aligned}$$

Inkluse (i) je indukční předpoklad spolu s pozorováním, že uřízneme-li poloprostorem H polovinu množiny $\mathcal{B}(E_{i-1}, \varepsilon_{i-1})$ a zahodíme ji, nemohli jsme zahodit žádný bod v $\mathcal{F}$. (ii) je použití lemmatu 24 spolu s dolním odhadem na nejmenší vlastní číslo a horním odhadem na největší vlastní číslo z bodu (a) tohoto lemmatu. Inkluse (iii) je zřejmá z lemmatu 10 — elipsoid $\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true})$ byl zkonstruován tak, aby obsahoval $E_{i-1} \cap H$, a tato vlastnost se zachovává i při $(\varepsilon_{i-1} \cdot 8^{i-1})$-nafouknutí.

K (iv). Použijeme větu 23, kde za proměnnou ε v předpokladu věty dosazujeme číslo

$$e := 2^{-4N^2-2N}.$$

Je třeba ověřit, že naše $\chi = 5N^2$ splňuje předpoklad věty: máme tedy nahlédnout, že platí

$$5N^2 \geqslant Q - \log_2 \frac{1}{2^{-4N^2-2N}} = Q + 4N^2 + 2N,$$

kde Q je stejné jako ve vztahu (30). Užijeme-li odhadu $Q \leqslant 17NnL$ uvedeného v (30), dostáváme

$$Q + 4N^2 + 2N \leqslant 17NnL + 4N^2 + 2N \leqslant 19NnL + 4N^2 \leqslant N^2 + 4N^2 = 5N^2 = \chi, \qquad (34)$$

protože z definice N je $19NnL \leqslant N^2$ (to je vidět z toho, že $19nL \leqslant 128n^3L = N$). Tím je předpoklad věty 23 ověřen.

Tvrzení věty 23(b) říká, že $\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true}) \subseteq \mathcal{B}(E_i, e) = \mathcal{B}(E_i, 2^{-4N^2-2N})$, a z toho plyne i

$$\begin{aligned}\mathcal{B}(\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true}), \varepsilon_{i-1} \cdot 8^{i-1}) &\subseteq \mathcal{B}(\mathcal{B}(E_i, 2^{-4N^2-2N}), \varepsilon_{i-1} \cdot 8^{i-1})) \\ &\overset{(\spadesuit)}{\subseteq} \mathcal{B}(E_i, \varepsilon_{i-1} \cdot 8^{i-1} + 2^{-4N^2-2N}),\end{aligned}$$

což je (iv). Inkluse (♠) plyne z pozorování 22(c).

Pro (v) jen potřebujeme dokázat

$$2^{-2N-4N^2+4N(i-1)+3(i-1)} + 2^{-2N-4N^2} \leqslant 2^{-2N-4N^2+4Ni}.$$

To je ale snadné, běžnými úpravami se dostane, že je to ekvivalentní s nerovností

$$2^{4Ni} \cdot 2^{-4N+3i-1} + 1 \leqslant 2^{4Ni},$$

což s $i = N$ evidentně platí, a tím spíš s $i \leqslant N$. □

Nyní jsme již ve finále. Následující věta říká, že jestliže jsme v N-té iteraci a dosud jsme nenalezli žádné řešení, pak $\mathcal{F} = \emptyset$; můžeme tedy skončit a prohlásit „ne, řešení neexistuje“. Je to tedy věta o korektnosti a zároveň věta o polynomialitě Chačijanova algoritmu.

Věta 26 (Chačijan)**.** *Buď $\boldsymbol{A}$ celočíselná matice a $\boldsymbol{b}$ celočíselný vektor. Nechť polyedr $\mathcal{F} = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\}$ je buďto prázdný, anebo je omezený a má plnou dimensi. Pro výpočet Chačijanova algoritmu platí:*

(a) Pro $i \leqslant N$ platí, že čísla v maticích $\boldsymbol{E}_i$ a vektorech $\boldsymbol{z}_i$ mají nanejvýš polynomiální velikost zápisu. (Takže výpočetní čas v jedné iteraci je nanejvýš poly(L).)

(b) Algoritmus dosáhne N-té iterace, aniž by nalezl bod v $\mathcal{F}$, právě když $\mathcal{F} = \emptyset$.

Důkaz (a) plyne z toho, že (i) zaokrouhlujeme na χ dvojkových míst, takže nám počet číslic za dvojkovou čárkou nenarůstá, a (ii) z toho, že délka čísel nenarůstá příliš rychle ani před dvojkovou čárkou. Tvrzení (ii) je třeba dokázat.

Důkaz (ii) pro $\boldsymbol{z}_i$ je snadný: podle lemmatu 25(b) je norma $\boldsymbol{z}_i$ malá, a tudíž i čísla ve vektoru $\boldsymbol{z}_i$ jsou malá.

Důkaz (ii) pro $\boldsymbol{E}_i$. Zajímají nás velká čísla, takže pro účely tohoto důkazu řekněme, že všechna čísla v matici $\boldsymbol{E}_i$ jsou $\geqslant 1$. Stopa (=součet prvků na diagonále) matice $\boldsymbol{E}_i^2$ je rovna $\sum_{j,k=1}^{n} (\boldsymbol{E}_i)_{jk}^2$. To znamená, že stopa matice $\boldsymbol{E}_i^2$ je horní odhad na absolutní hodnotu libovolného z čísel v matici $\boldsymbol{E}$. Stopa $\boldsymbol{E}_i^2$ je rovna součtu vlastních čísel matice $\boldsymbol{E}_i^2$, a to jsou druhé mocniny vlastních čísel $\boldsymbol{E}_i$. Na ně máme horní odhad podle lemmatu 25(a). Jest tedy $stopa(\boldsymbol{E}_i^2) \leqslant n \cdot (8^i R^2)^2 \leqslant n \cdot 16^N R^4$, a toto číslo lze zapsat pomocí $O(\log n + N + \log R) = poly(L)$ bitů, kde jsme užili faktu $\log_2 R = O(nL)$.

(b). Jedna implikace je zřejmá: je-li $\mathcal{F} = \emptyset$, žádný správný algoritmus (milovník Excelu promine) nemůže žádné řešení nikdy najít, takže ani ten náš během N iterací. Obtížná je implikace *jestliže jsme dosáhli N-té iterace* (tedy: právě jsme spočetli $\boldsymbol{z}_N$ a $\boldsymbol{E}_N$), *pak* $\mathcal{F} = \emptyset$. Dokazujme sporem. Platí

$$\frac{vol(E_{i+1})}{vol(E_i)} = \frac{vol(E_{i+1})}{vol(\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true}))} \cdot \frac{vol(\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true}))}{vol(E_i)} \leqslant \left(1 + \frac{1}{8n}\right) e^{-\frac{1}{2(n+1)}} < e^{-\frac{1}{8n}}, \quad (35)$$

kde jsme použili větu[10] 23 s $\varepsilon := \frac{1}{8n}$; její tvrzení (c) dává odhad

$$\frac{vol(\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true}))}{vol(E_i)} \leqslant 1 + \frac{1}{8n}.$$

Věta 12 zaručuje

$$\frac{vol(E_{i+1})}{vol(\mathcal{E}(\boldsymbol{z}^{true}, \boldsymbol{E}^{true}))} \leqslant e^{-\frac{1}{2(n+1)}}.$$

Dále je $1 + \frac{1}{8n} < e^{\frac{1}{8n}}$ (užíváme $1 + x < e^x$ pro $x \neq 0$) a zřejmě

$$e^{\frac{1}{8n} - \frac{1}{2n+2}} < e^{-\frac{1}{8n}},$$

čímž je platnost (35) dokázána.

Předpokládejme, že jsme dosáhli N-té iterace. Platí

$$vol(E_N) \leqslant (2R)^n e^{-\frac{N}{8n}};$$

koule E_0 totiž má poloměr R, odkud $vol(E_0) \leqslant (2R)^n$.

Podle lemmatu 25(c) s $i = N$ máme

$$\mathcal{F} \subseteq \mathcal{B}(E_N, 2^{-2N-4N^2+4N^2}) = \mathcal{B}(E_N, 2^{-2N}). \quad (36)$$

Inkluse (36) spolu s pozorováním 22 (konkrétně druhá inkluse v (23) s $\epsilon := 2^{-2N}$ a s odhadem na nejmenší vlastní číslo $\lambda(\boldsymbol{E}_N) \geqslant 8^{-N} R^2$ podle lemmatu 25(a)) dává

$$\mathcal{F} \subseteq \mathcal{B}(E_N, 2^{-2N}) \subseteq (1 + 2^{-2N} \cdot \tfrac{8^{N/2}}{R}) \bullet E_N \subseteq 2 \bullet E_N$$

($2^{-2N} \frac{8^{N/2}}{R} = \frac{2^{-N/2}}{R}$ je triviálně menší než 1), takže máme zaručeno, že

$$\mathcal{F} \subseteq 2 \bullet E_N.$$

Předpokládejme, že $vol(\mathcal{F}) > 0$; podle lemmatu 19 jest $vol(\mathcal{F}) \geqslant 2^{-7n^2 L}$. Pak

$$\begin{aligned} 2^{-7n^2 L} &\leqslant vol(\mathcal{F}) \leqslant vol(2 \bullet E_N) \overset{\text{poz. } 22(b)}{=} 2^n \cdot vol(E_N) < 2^n (2R)^n e^{-\frac{N}{8n}} \\ &< 2^n (2R)^n 2^{-\frac{N}{8n}} = 2^n (2 \cdot 2^{3nL})^n 2^{-\frac{128 n^3 L}{8n}} = 2^{2n + 3n^2 L - 16 n^2 L} \\ &= 2^{-7n^2 L - n(6nL - 2)} \overset{6nL > 2}{<} 2^{-7n^2 L} \end{aligned}$$

a máme spor. Protože jsme předpokládali, že $\mathcal{F}$ je buď prázdný anebo má plnou dimensi, a druhý případ vede ke sporu, musí být $\mathcal{F} = \emptyset$. □

Dokázali jsme kýžený výsledek: $\mathbf{LP} \in \boldsymbol{P}$. Důkaz Chačijanovy věty je dokončen.

[10] Není těžké nahlédnout, že naše volba $\chi = 5N^2$ zaručuje pro $\varepsilon := \frac{1}{8n}$ platnost předpokladu věty 23: s vědomím, že $\frac{1}{8n} \geqslant 2^{-4N^2 - 2N}$, lze přímo použít odhad (34) z bodu (iv) části (c) důkazu lemmatu 25.