

Reálne korene polynomickej rovnice

Pri mnohých úlohách nepotrebujeme nájsť všetky body nejakej algebraickej variety, špeciálne komplexné korene nás často nezaujímajú. Väčšinou chceme nájsť reálne korene, často sa dokonca uspokojíme aj s ich dostatočnou aproximáciou. Ukázali sme si, že problém hľadania riešení sústavy polynomických rovníc (v prípade, že ich je konečne veľa) vieme zredukovať na hľadanie koreňov jednej polynomickej rovnice s jednou neznámou. Stačia nám teda metódy, ktoré nám pomôžu nájsť korene takejto rovnice.

Numerická matematika ponúka niekoľko postupov hľadania koreňov (presnejšie ich aproximácií), každá z nich má nejaké obmedzenia. Všeobecná Newtonova metóda nájde jeden koreň, i to v prípade, že máme dobrý prvý odhad. Metóda Bezierovho orezávania nájde všetky korene na vopred zvolenom intervale, funguje ale len pre polynómy (na rozdiel od Newtonovej metódy, ktorá je použiteľná pri akejkoľvek diferencovateľnej funkcii). Takéto orezávanie je pomerne spoľahlivé, problémy sa môžu vyskytnúť pri nestabilných situáciách (viacnásobný koreň). V prípade, že máme korene rovnice vopred separované, čiže máme zoznam intervalov taký, že v každom intervale sa nachádza práve jeden koreň, použiteľná metóda na nájdenie koreňa je aj jednoduché binárne delenie intervalu.

V tejto kapitole si uvedieme tvrdenia a postupy na predspracovanie polynomickej rovnice, takže potom bude možné korene dohľadať numerickými metódami.

1. Ohraničenie koreňov

VERA 1.1. *Nech $f \in \mathbb{R}[x]$, $f = f_m x^m + \dots + f_1 x + f_0$, $f_m \neq 0$. Ak*

$$(10) \quad |a| \geq 2 \sum_{i=0}^m \frac{|f_i|}{|f_m|},$$

potom $f(a)$ a $f_m a^m$ majú rovnaké znamienko.

Dôkaz. Najprv si všimnime, že z (10) vyplýva, že $|a| > 2$. Odtiaľ ukážeme, že podiel $f(a)/f_m a^m$ je kladný.

$$\begin{aligned} \frac{f(a)}{f_m a^m} &= 1 + \sum_{i=0}^{m-1} \frac{f_i}{f_m} a^{i-m} \geq 1 - \left(\sum_{i=0}^{m-1} \frac{|f_i|}{|f_m|} |a|^{i-m} \right) \\ &\geq 1 - \left(\sum_{i=0}^{m-1} \frac{|f_i|}{|f_m|} \right) \left(\frac{1}{|a|} + \dots + \frac{1}{|a|^m} \right) \\ &\geq 1 - \frac{1}{2} \left(1 + \frac{1}{|a|} + \dots + \frac{1}{|a|^{m-1}} \right) > 1 - \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{|a|}} > 0. \end{aligned}$$

□

DÔSLEDOK. Všetky reálne korene rovnice $f_m x^m + \dots + f_1 x + f_0 = 0$, ($f_m \neq 0$) sa nachádzajú v intervale

$$\left(-2 \sum_{i=0}^m \frac{|f_i|}{|f_m|}, 2 \sum_{i=0}^m \frac{|f_i|}{|f_m|} \right).$$

Existuje viacero ohraničení pre reálne korene, aj omnoho lepšie, pre nás však stačí takéto jednoduché. Podstatné je, že vieme explicitne napísať interval obsahujúci všetky reálne korene.

2. Sturmova veta

Vo zvyšku kapitoly budeme používať nasledovné označenie: ak f, g sú polynómy z $k[x]$, tak $\text{rem}(f, g)$ označuje zvyšok po delení polynómu f polynómom g , t.j. jediný taký polynóm $r \in k[x]$, že

$$f = qg + r, \quad \text{kde } \deg r < \deg g.$$

DEFINÍCIA 2.1. *Sturmova postupnosť polynómu* $p \in \mathbb{R}[x]$ je postupnosť $(p_0, p_1, \dots, p_k)$ polynómov, kde

$$\begin{aligned} p_0 &= p, \\ p_1 &= p', \\ p_i &= -\text{rem}(p_{i-2}, p_{i-1}), \quad \text{pre } i \geq 2, \end{aligned}$$

kde p_k je posledný nenulový člen tejto postupnosti zvyškov (t.j. $p_k \mid p_{k-1}$).

V Sturmovej postupnosti zrejme platí

$$(11) \quad p_{i-1} = q_{i-1} p_i - p_{i+1}.$$

DEFINÍCIA 2.2. Nech $\mathbf{a} = (a_0, \dots, a_k)$ je postupnosť nenulových reálnych čísel. Počet znamienkových zmien $\text{var}(\mathbf{a})$ v postupnosti $\mathbf{a}$ je definovaný

$$\begin{aligned} \text{var}(a_0) &= 0 \\ \text{var}(a_0, \dots, a_i) &= \begin{cases} 1 + \text{var}(a_0, \dots, a_{i-1}), & \text{ak } a_{i-1} a_i < 0 \\ \text{var}(a_0, \dots, a_{i-1}), & \text{ak } a_{i-1} a_i > 0. \end{cases} \end{aligned}$$

Ak $\mathbf{a} = (a_0, \dots, a_k)$ je postupnosť reálnych čísel, potom $\text{var}(\mathbf{a})$ definujeme ako počet znamienkových zmien v postupnosti, ktorú získame z $\mathbf{a}$ vynechaním všetkých núl.

PRÍKLAD 2.3. $\text{var}(1, -2, 2, 0, 0, 3, 4, -5, -2, 0, 3) = 4$.

DEFINÍCIA 2.4. Nech $\mathbf{p} = (p_0, p_1, \dots, p_k)$ je postupnosť polynómov v $\mathbb{R}[x]$ a nech a je reálne číslo. Potom označujeme

$$\text{var}_{\mathbf{p}}(a) = \text{var}(p_0(a), p_1(a), \dots, p_k(a)).$$

VETA 2.5 (**Sturm**). Nech $p \in \mathbb{R}[x]$ a nech $\mathbf{p} = (p_i)_{i=0}^k$ je jeho Sturmova postupnosť. Nech $a, b \in \mathbb{R}$ sú také, že $a < b$, $p(a) \neq 0$, $p(b) \neq 0$. Počet reálnych koreňov polynómu p na intervale (a, b) je rovný číslu

$$\text{var}_{\mathbf{p}}(a) - \text{var}_{\mathbf{p}}(b).$$

Dôkaz Sturmovej vety rozdelíme na viacero tvrdení.

DEFINÍCIA 2.6. Nech $p \in \mathbb{R}[x]$ a nech $\mathbf{p} = (p_i)_{i=0}^k$ je jeho Sturmova postupnosť. Nech $a, b \in \mathbb{R}$, $a < b$. Interval $\langle a, b \rangle$ sa nazýva *fundamentálny interval polynómu* p , ak $p(a) \neq 0, p(b) \neq 0$, a existuje také $\gamma_0 \in (a, b)$, že pre všetky $\gamma \in \langle a, b \rangle$, $\gamma \neq \gamma_0$ platí $p_i(\gamma) \neq 0$ pre všetky p_i , $i = 0, \dots, k$.

Neformálne, $\langle a, b \rangle$ je fundamentálnym intervalom polynómu p , ak každý z polynómov Sturmovej postupnosti má na $\langle a, b \rangle$ najviac jeden koreň, a ten je rovnaký pre všetky členy postupnosti.

TVRDENIE 2.7. *Nech $\langle a, b \rangle$ je fundamentálny interval polynómu $p \in \mathbb{R}[x]$. Ak p nemá reálny koreň na $\langle a, b \rangle$, potom $\text{var}_{\mathbf{p}}(a) = \text{var}_{\mathbf{p}}(b)$.*

Dôkaz. Predpokladajme najprv, že žiaden z polynómov p_i Sturmovej postupnosti nemá koreň na $\langle a, b \rangle$. Potom graf každého polynómu je buď celý nad osou x alebo pod ňou, takže $\text{sgn}(p_i(a)) = \text{sgn}(p_i(b))$, a preto $\text{var}_{\mathbf{p}}(a) = \text{var}_{\mathbf{p}}(b)$.

Nech teraz existuje i také, že $p_i(\gamma_0) = 0$ pre nejaké $\gamma_0 \in (a, b)$. Z (11) vyplýva, že po sebe idúce polynómy v Sturmovej postupnosti nemajú koreň v γ_0 , lebo inak by γ_0 bol koreňom všetkých polynómov postupnosti, aj $p_0 = p$, čo by bol spor s predpokladom. Teda $p_{i-1}(\gamma_0) \neq 0$ a tiež $p_{i+1}(\gamma_0) \neq 0$, a navyše z (11) dostávame aj $p_{i+1}(\gamma_0) = -p_{i-1}(\gamma_0)$. Keďže $\langle a, b \rangle$ je fundamentálny interval, p_{i+1} a p_{i-1} na ňom žiaden koreň nemajú, a tak

$$\text{sgn}(p_{i+1}(a)) = \text{sgn}(p_{i+1}(b)) = -\text{sgn}(p_{i-1}(a)) = -\text{sgn}(p_{i-1}(b)).$$

V oboch postpostupnostiach

$$\begin{aligned} &(p_{i-1}(a), p_i(a), p_{i+1}(a)) \\ &(p_{i-1}(b), p_i(b), p_{i+1}(b)) \end{aligned}$$

tak dochádza presne k jednej znamienkovej zmene. Tým sme ukázali, že počet znamienkových zmien v oboch postupnostiach $(p_i(a))_{i=0}^k$ a $(p_i(b))_{i=0}^k$ je rovnaký. $\square$

TVRDENIE 2.8. *Nech $\langle a, b \rangle$ je fundamentálny interval polynómu $p \in \mathbb{R}[x]$. Ak p má jeden reálny koreň na $\langle a, b \rangle$, potom $\text{var}_{\mathbf{p}}(a) = \text{var}_{\mathbf{p}}(b) + 1$.*

Dôkaz. Nech γ_0 je jednoduchý koreň polynómu p , čiže $p'(\gamma_0) \neq 0$. Tak ako v dôkaze predchádzajúceho tvrdenia potom zistíme, že žiadne dva za sebou idúce polynómy Sturmovej postupnosti nemajú koreň v γ_0 , a tiež odtiaľ analogicky usúdime, že

$$\text{var}(p_1(a), \dots, p_k(a)) = \text{var}(p_1(b), \dots, p_k(b)).$$

Rozdiel medzi počtom znamienkových zmien tak môže nastať iba medzi prvými dvoma členmi postupností.

Keďže $p'(\gamma_0) \neq 0$, p' nemá koreň na $\langle a, b \rangle$ (ide o fundamentálny interval), je p' na celom intervale buď kladný (a p rastie na $\langle a, b \rangle$), alebo záporný (a p klesá). Možnosti pre znamienka prvých dvoch členov postupností tak sú

$p(a)$	$p'(a)$	$p(b)$	$p'(b)$
–	+	+	+
+	–	–	–

a tak v tomto prípade dostávame $\text{var}_{\mathbf{p}}(a) = \text{var}_{\mathbf{p}}(b) + 1$.

Nech $p(\gamma_0) = p'(\gamma_0) = 0$, teda γ_0 je viacnásobný koreň polynómu p . Ak jeho násobnosť v p je r , potom jeho násobnosť v p' je $r - 1$. Navyše p_k je najväčším spoločným deliteľom p a p' (Sturmova postupnosť je až na znamienka postupnosťou zvyškov v euklidovom algoritme), takže γ_0 je $(r - 1)$ -násobný koreň aj v p_k .

Uvažujme postupnosť polynómov

$$\tilde{\mathbf{p}} = (\tilde{p}_0, \tilde{p}_1, \dots, \tilde{p}_k = 1), \quad \text{kde } \tilde{p}_i = \frac{p_i}{p_k}.$$

(Toto už nie je Sturmova postupnosť, lebo $\tilde{p}_1$ nie je deriváciou $\tilde{p}_0$). Pre počty znamienkových zmien platí

$$\text{var}_{\tilde{\mathbf{p}}}(a) = \text{var}_{\mathbf{p}}(a), \quad \text{var}_{\tilde{\mathbf{p}}}(b) = \text{var}_{\mathbf{p}}(b).$$

Taktiež ostáva v platnosti vzťah

$$\tilde{p}_{i-1} = q_{i-1}\tilde{p}_i - \tilde{p}_{i+1}.$$

V postupnosti $\tilde{\mathbf{p}}$ už γ_0 je len jednoduchým koreňom polynómu $\tilde{p}_0$, a nie je koreňom $\tilde{p}_1$. Podobne ako v prvej časti dôkazu tak dostávame, že

$$\text{var}(\tilde{p}_1(a), \dots, \tilde{p}_k(a)) = \text{var}(\tilde{p}_1(b), \dots, \tilde{p}_k(b)),$$

rozdiel v počte znamienkových zmien tak znovu môže nastať len medzi prvými dvoma členmi postupnosti. Keďže $\tilde{p}_1$ nie je deriváciou $\tilde{p}_0$, máme tak viacej možností:

$\tilde{p}_0(a)$	$\tilde{p}_1(a)$	$\tilde{p}_0(b)$	$\tilde{p}_1(b)$
+	+	−	+
+	−	−	−
−	+	+	+
−	−	+	−

Potrebuje ukázať, že možnosti v prvom a štvrtom riadku nenastanú. Pre tento účel potrebujeme rozlíšiť, či násobnosť r koreňa γ_0 v p je párna alebo nepárna.

Ak r je nepárne číslo, potom platí

$$\text{sgn}(p(a)) = -\text{sgn}(p(b))$$

$$\text{sgn}(p'(a)) = \text{sgn}(p'(b))$$

$$\text{sgn}(p_k(a)) = \text{sgn}(p_k(b)).$$

Z tabuľky pre znamienka p a p'

$p(a)$	$p'(a)$	$p(b)$	$p'(b)$
−	+	+	+
+	−	−	−

tak dostávame tabuľku pre $\tilde{p}_0$ a $\tilde{p}_1$:

$\tilde{p}_0(a)$	$\tilde{p}_1(a)$	$\tilde{p}_0(b)$	$\tilde{p}_1(b)$
−	+	+	+
+	−	−	−
+	−	−	−
−	+	+	+

V prípade, keď r je párne, sa postupuje analogicky. □

Dôkaz Vety 2.5. Dve predchádzajúce tvrdenia sú vlasne dôkazom Sturmovej vety v prípade, že interval $\langle a, b \rangle$ je fundamentálny. Ak $\langle a, b \rangle$ nie je fundamentálny, tak ho rozdelíme na fundamentálne intervaly: Nech $\gamma_0 < \gamma_1 < \dots < \gamma_k$ sú všetky korene všetkých polynómov Sturmovej postupnosti na intervale (a, b) . Zvolíme α_i , $i = 1, \dots, k$ tak, že

$$a < \gamma_0 < \alpha_1 < \gamma_1 < \alpha_2 < \dots < \alpha_k < \gamma_k < b.$$

Potom $\langle a_{i-1}, a_i \rangle$ sú fundamentálne intervaly polynómu p a pre počty koreňov tak máme

$$\begin{aligned} \#\text{koreňov na } (a, b) &= \#\text{koreňov na } (a, a_1) + \#\text{koreňov na } (a_1, a_2) + \dots + \#\text{koreňov na } (a_k, b) \\ &= (\text{var}_{\mathbf{p}}(a) - \text{var}_{\mathbf{p}}(a_1)) + (\text{var}_{\mathbf{p}}(a_1) - \text{var}_{\mathbf{p}}(a_2)) + \dots + (\text{var}_{\mathbf{p}}(a_k) - \text{var}_{\mathbf{p}}(b)) \\ &= \text{var}_{\mathbf{p}}(a) - \text{var}_{\mathbf{p}}(b). \end{aligned}$$

□

PRÍKLAD 2.9. Separujeme reálne korene polynómu

$$f = x^3 - 4x^2 + 3x + 1.$$

Pomocou Vety 1.1 najprv nájdeme interval obsahujúci všetky reálne korene polynómu:

$$M = 2 \sum_{i=0}^3 \frac{|f_i|}{|f_3|} = 2(1 + 4 + 3 + 1) = 18,$$

takže všetky reálne korene sú v intervale $(-18, 18)$. Sturmova postupnosť polynómu f je

$$\begin{aligned} p_0 &= f = x^3 - 4x^2 + 3x + 1, \\ p_1 &= f' = 3x^2 - 8x + 3, \\ p_2 &= -\text{rem}(f, f') = \frac{14}{9}x - \frac{7}{3}, \\ p_3 &= -\text{rem}(p_1, p_2) = \frac{9}{4}. \end{aligned}$$

Pomocou Sturmovej vety zistíme počet reálnych koreňov:

$$\begin{aligned} \text{var}_{\mathbf{p}}(-18) &= \text{var}(p_0(-18), p_1(-18), p_2(-18), p_3(-18)) = 3 \\ \text{var}_{\mathbf{p}}(18) &= \text{var}(p_0(18), p_1(18), p_2(18), p_3(18)) = 0 \end{aligned}$$

takže polynóm f má tri reálne korene. Z

$$\text{var}_{\mathbf{p}}(0) = \text{var}\left(1, 3, -\frac{7}{3}, \frac{9}{4}\right) = 2$$

vieme, že dva z koreňov sa nachádzajú v intervale $(0, 18)$ a jeden je v intervale $(-18, 0)$. Nakoniec

$$\text{var}_{\mathbf{p}}(2) = \text{var}\left(-1, -1, \frac{7}{9}, \frac{9}{4}\right) = 1,$$

nám hovorí, že jeden z kladných koreňov je na intervale $(0, 2)$ a druhý na intervale $(2, 18)$. Všetky korene sú jednoduchými koreňmi, takže ich ľahko dohľadáme ľubovoľnou numerickou metódou (napr. aj jednoduchou bisekciou intervalu).