1. Iterační metody pro řešení systémů nelineárních algebraických a transcendentních rovnic

Buďte $f_1, f_2, ..., f_n$ reálné funkce n reálných proměnných. Naším cílem je nalézt řešení soustavy

$$f_1(x_1, x_2, ..., x_n) = 0,\ f_2(x_1, x_2, ..., x_n) = 0,\ \dots,\ f_n(x_1, x_2, ..., x_n) = 0 \tag{1}$$

ve tvaru $\vec{a} = (a_1, a_2, ..., a_n)^T$. Separaci kořenů v tomto případě vůbec neumíme provést, a proto budeme dále předpokládat, že známe konvexní oblast G s vlastností $\vec{a} \in G$.

1.1. **Princip iteračních metod.**

Systém (1) převedeme na systém

$$x_1 = \varphi_1(x_1, x_2, ..., x_n),\ x_2 = \varphi_2(x_1, x_2, ..., x_n),\ ...,\ x_n = \varphi_n(x_1, x_2, ..., x_n) \tag{2}$$

a konstruujeme posloupnost vektorů $(\vec{x}^{(k)})$ takto:

$$\begin{aligned} x_1^{(k+1)} &= \varphi_1(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}),\\ x_2^{(k+1)} &= \varphi_2(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}),\\ &\dots,\\ x_n^{(k+1)} &= \varphi_n(x_1^{(k)}, x_2^{(k)}, ..., x_n^{(k)}). \end{aligned} \tag{3}$$

Tvrzení 1.1. Buďte

(1) řešení $\vec{a}$ systému (2) odseparováno v konvexní oblasti G,
(2) $(\forall i, j \in \hat{n})(\frac{\partial \varphi_i}{\partial x_j}$ spojité na $\overline{G})$,
(3) $(\forall k \in \mathbb{N}_0)(\vec{x}^{(k)} \in G)$,
(4) $\varrho(\mathcal{M}) < 1$, kde $\mathcal{M}$ je matice definovaná vztahem

$$[\mathcal{M}]_{ij} = \max_{\vec{x} \in \overline{G}} \left| \frac{\partial \varphi_i}{\partial x_j}(\vec{x}) \right|.$$

Potom posloupnost $(\vec{x}^{(k)})$ daná předpisem (3) konverguje k řešení $\vec{a}$.

Důkaz. Vyjděme z i-té rovnice v (3). Podle věty o přírůstku reálné funkce více proměnných existuje vektor $\vec{p}_i^{(k)}$ ležící ve „vnitřku" úsečky spojující body $\vec{x}^{(k)}$, $\vec{a}$ tak, že

$$x_i^{(k+1)} - a_i = \varphi_i(\vec{x}^{(k)}) - \varphi_i(\vec{a}) = \sum_{j=1}^{n} \frac{\partial \varphi_i}{\partial x_j}(\vec{p}_i^{(k)})(x_j^{(k)} - a_j). \tag{4}$$

To platí $\forall i \in \hat{n}$, takže máme pro každé $k \in \mathbb{N}$ zaručenu existenci n-tice vektorů $\vec{p}_1^{(k)}, \dots, \vec{p}_n^{(k)}$ splňujících vztah (4). Označíme-li

$$\mathcal{M}_k = \begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1}(\vec{p}_1^{(k)}) & \dots & \frac{\partial \varphi_1}{\partial x_n}(\vec{p}_1^{(k)}) \\ \vdots & & \vdots \\ \frac{\partial \varphi_n}{\partial x_1}(\vec{p}_n^{(k)}) & \dots & \frac{\partial \varphi_n}{\partial x_n}(\vec{p}_n^{(k)}) \end{pmatrix},$$

potom podle (4) zřejmě platí

$$\vec{x}^{(k+1)} - \vec{a} = \mathcal{M}_k(\vec{x}^{(k)} - \vec{a}) = \mathcal{M}_k\mathcal{M}_{k-1}(\vec{x}^{(k-1)} - \vec{a}) = ... = \mathcal{M}_k\mathcal{M}_{k-1}...\mathcal{M}_0(\vec{x}^{(0)} - \vec{a}).$$

Z toho plyne, že nutnou a postačující podmínkou pro konvergenci $\vec{x}^{(k)} \to \vec{a}$ za předpokladů 1–3 je

$$\lim_{k\to\infty} \prod_{i=0}^{k} \mathcal{M}_i = O. \tag{5}$$

Z definice matice $\mathcal{M}$ vyplývá $(\forall k \in \mathbb{N})([\mathcal{M}]_{ij} \geq |[\mathcal{M}_k]_{ij}|)$. Protože oblast G je konvexní, obsahuje i úsečku $\langle \vec{x}^{(k)}, \vec{a} \rangle$. Odtud, z předchozí nerovnosti a z trojúhelníkové nerovnosti plyne další nerovnost $[\mathcal{M}^{k+1}]_{ij} \geq |[\mathcal{M}_k\mathcal{M}_{k-1}...\mathcal{M}_0]_{ij}|$, neboť na obou stranách této nerovnice jsou sumy stejného počtu součinů a vlevo vystupuje místo původního prvku odhad. Z této nerovnosti vyplývá, že konverguje-li posloupnost $(\mathcal{M}^k)$ k nulové matici, potom je splněno i (5). Podle věty **??** tak dostáváme kritérium konvergence $\varrho(\mathcal{M}) < 1$. □

Poznámka (důležitá). Předpoklad 3 v tvrzení 1.1 je podstatný. Následující tvrzení říká, že k jeho splnění je třeba zvolit $\vec{x}^{(0)}$ dostatečně blízko $\vec{a}$, ale nedozvíme se jak.

Tvrzení 1.2. Za předpokladů 1, 2, 4 tvrzení 1.1 existuje takové okolí V bodu $\vec{a}$, že $x^{(0)} \in V \Rightarrow (\forall k \in \mathbb{N})(\vec{x}^{(k)} \in G)$.

Důkaz. Podle předpokladu 4 platí $\mathcal{M}^k \to O$. Z toho podle definice **??** a podle definice normy $\| \ \|_{\mathrm{I}}$ vyplývá existence takového $k \in \mathbb{N}$, že $\left\|\mathcal{M}^{k+1}\right\|_{\mathrm{I}} < 1$. Definujme $V = \{\vec{x} | \, \|\vec{x} - \vec{a}\|_{\mathrm{I}} < R\} = B_{\mathrm{I}}(\vec{a}, R)$, kde R je tak malé, že platí

$$\vec{x}^{(0)} \in V \Rightarrow \vec{x}^{(1)}, \vec{x}^{(2)}, ..., \vec{x}^{(k)} \in G. \tag{6}$$

Vhodné R určitě existuje, protože k je pevné konečné číslo. Buď $\vec{x}^{(0)} \in V$. Potom z vlastností normy $\| \ \|_{\mathrm{I}}$ plyne

$$\left\|\vec{x}^{(k+1)} - \vec{a}\right\|_{\mathrm{I}} = \left\|\mathcal{M}_k\mathcal{M}_{k-1}\dots\mathcal{M}_0(\vec{x}^{(0)} - \vec{a})\right\|_{\mathrm{I}} \leq$$

$$\leq \|\mathcal{M}_k\mathcal{M}_{k-1}\dots\mathcal{M}_0\|_{\mathrm{I}} \left\|\vec{x}^{(0)} - \vec{a}\right\|_{\mathrm{I}} \leq \left\|\mathcal{M}^{k+1}\right\|_{\mathrm{I}} \left\|\vec{x}^{(0)} - \vec{a}\right\|_{\mathrm{I}}.$$

Protože $\left\|\mathcal{M}^{k+1}\right\|_{\mathrm{I}} < 1$ a $\left\|\vec{x}^{(0)} - \vec{a}\right\|_{\mathrm{I}} < R$, je i součin $\left\|\mathcal{M}^{k+1}\right\|_{\mathrm{I}} \left\|\vec{x}^{(0)}\right\|_{\mathrm{I}} < R$, a proto $\vec{x}^{(k+1)} \in V$. To podle (6) znamená, že $\vec{x}^{(k+2)}, \vec{x}^{(k+3)}, \dots, \vec{x}^{(2k+1)} \in G$. Nyní stejný postup aplikujeme na $\vec{x}^{(k+2)}, \vec{x}^{(k+3)}, \dots, \vec{x}^{(2k+1)}$, takže se dozvíme, že $\vec{x}^{(2k+2)} \in V$, atd. □

1.2. **Newtonova metoda.** Přepišme nyní soustavu (1) do vektorové podoby $\vec{f}(\vec{x}) = \vec{o}$, kde $\vec{f}(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_n(x_1, \dots, x_n))^T$. Budiž řešení $\vec{a}$ této rovnice odseparováno v konvexní oblasti G a nechť na G existují $\frac{\partial f_i}{\partial x_j}(\vec{x})$ a jsou spojité. Označme

$$f_x(\vec{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{x}) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\vec{x}) & \dots & \frac{\partial f_n}{\partial x_n}(\vec{x}) \end{pmatrix}$$

a předpokládejme, že matice $f_x(\vec{a})$ je regulární. Potom jsou ovšem regulární i všechny matice $f_x(\vec{x})$ pro $\vec{x} \in V_{\vec{a}}$, kde $V_{\vec{a}}$ je jisté okolí řešení $\vec{a}$. Nyní vyjděme z prvního vztahu ve (**??**) a vytvořme jeho maticovou obdobu

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - f_x^{-1}(\vec{x}^{(k)})\vec{f}(\vec{x}^{(k)}). \tag{7}$$

Vektor $\vec{x}^{(k)}$ převedeme na levou stranu rovnosti a obě její strany vynásobíme zleva maticí $f_x(\vec{x}^{(k)})$, takže dostaneme $f_x(\vec{x}^{(k)})(\vec{x}^{(k+1)} - \vec{x}^{(k)}) = -\vec{f}(\vec{x}^{(k)})$. Označíme-li $\Delta\vec{x}^{(k)} = \vec{x}^{(k+1)} - \vec{x}^{(k)}$, přejde tento vztah ve tvar

$$f_x(\vec{x}^{(k)})\Delta\vec{x}^{(k)} = -\vec{f}(\vec{x}^{(k)}),\ \vec{x}^{(k+1)} = \vec{x}^{(k)} + \Delta\vec{x}^{(k)}.$$

Věta 1.3. (1) Nechť konvexní oblast G obsahuje řešení $\vec{a}$ systému $\vec{f}(\vec{x}) = \vec{o}$.
(2) Buď $f_x(\vec{a})$ regulární.
(3) Buďte $f_1, f_2, \dots, f_n$ spojité na G včetně svých prvních parciálních derivací.

Potom existuje δ-okolí $R = \{\vec{x} | \, \|\vec{x} - \vec{a}\| < \delta\}$ bodu $\vec{a}$ takové, že pro $\vec{x}^{(0)} \in R$ posloupnost $(\vec{x}^{(k)})$ v Newtonově metodě konverguje k $\vec{a}$.

Důkaz. Buďte $\vec{y}, \vec{z} \in G$. Definujme funkci

$$\Phi_i(t) = f_i(\vec{y} + t(\vec{z} - \vec{y})) = f_i(y_1 + t(z_1 - y_1),\ \dots,\ y_n + t(z_n - y_n)).$$

Potom zřejmě $\Phi_i(0) = f_i(\vec{y})$ a $\Phi_i(1) = f_i(\vec{z})$. Dále $\forall i \in \hat{n}$ platí

$$f_i(\vec{z}) - f_i(\vec{y}) = \Phi_i(1) - \Phi_i(0) = \int_0^1 \frac{\mathrm{d}\Phi_i(t)}{\mathrm{d}t}\ \mathrm{d}t = \int_0^1 \frac{\mathrm{d}f_i}{\mathrm{d}t}(\vec{y} + t(\vec{z} - \vec{y}))\ \mathrm{d}t =$$

$$= \int_0^1 \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\vec{y} + t(\vec{z} - \vec{y}))(z_j - y_j)\ \mathrm{d}t = \sum_{j=1}^n \underbrace{\left(\int_0^1 \frac{\partial f_i}{\partial x_j}(\vec{y} + t(\vec{z} - \vec{y}))\ \mathrm{d}t\right)}_{\text{ozn. } \mathcal{F}_{ij}(\vec{y}, \vec{z})}(z_j - y_j).$$

Označíme-li $F(\vec{y}, \vec{z}) = (\mathcal{F}_{ij}(\vec{y}, \vec{z}))$, potom podle předchozí rovnosti $\forall \vec{y}, \vec{z} \in G$ platí

$$\vec{f}(\vec{z}) - \vec{f}(\vec{y}) = F(\vec{y}, \vec{z})(\vec{z} - \vec{y}). \tag{8}$$

Položíme-li $\vec{y} = \vec{z} = \vec{a}$, pak

$$\mathcal{F}_{ij}(\vec{a},\vec{a}) = \frac{\partial f_i}{\partial x_j}(\vec{a}) \Rightarrow F(\vec{a},\vec{a}) = f_x(\vec{a}) \Rightarrow f_x^{-1}(\vec{a})F(\vec{a},\vec{a}) = I.$$

Podle předpokladu 3 proto existuje takové okolí $R \subset G$ řešení $\vec{a}$, že pro každé $\vec{x} \in R$ platí $\left\| I - f_x^{-1}(\vec{x})F(\vec{a},\vec{x}) \right\| \leq K < 1$. Matematickou indukcí dokážeme, že

$$\vec{x}^{(0)} \in R \Rightarrow (\forall k \in \mathbb{N})(\vec{x}^{(k)} \in R).$$

Podle (7) a (8) pro všechna $k \in \mathbb{N}$ taková, že $\vec{x}^{(k)} \in G$, platí

$$\vec{x}^{(k+1)} - \vec{a} = \vec{x}^{(k)} - \vec{a} - f_x^{-1}(\vec{x}^{(k)})(\vec{f}(\vec{x}^{(k)}) - \vec{f}(\vec{a})) =$$
$$= \vec{x}^{(k)} - \vec{a} - f_x^{-1}(\vec{x}^{(k)})F(\vec{a},\vec{x}^{(k)})(\vec{x}^{(k)} - \vec{a}) = (I - f_x^{-1}(\vec{x}^{(k)})F(\vec{a},\vec{x}^{(k)}))(\vec{x}^{(k)} - \vec{a}).$$

(1) Buď $\vec{x}^{(0)} \in R$. Dokážeme, že $\vec{x}^{(1)} \in R$:

$$\left\| \vec{x}^{(1)} - \vec{a} \right\| = \left\| I - f_x^{-1}(\vec{x}^{(0)})F(\vec{a},\vec{x}^{(0)})(\vec{x}^{(0)} - \vec{a}) \right\| \leq$$
$$\leq \left\| I - f_x^{-1}(\vec{x}^{(0)})F(\vec{a},\vec{x}^{(0)}) \right\| \left\| \vec{x}^{(0)} - \vec{a} \right\| < 1.\delta = \delta.$$

(2) Buďte $(\forall i \in \hat{k}_0)(\vec{x}^{(i)} \in R)$. Dokážeme, že $\vec{x}^{(k+1)} \in R$:

$$\left\| \vec{x}^{(k+1)} - \vec{a} \right\| \leq \left\| I - f_x^{-1}(\vec{x}^{(k)})F(\vec{a},\vec{x}^{(k)}) \right\| \left\| \vec{x}^{(k)} - \vec{a} \right\| \leq K \left\| \vec{x}^{(k)} - \vec{a} \right\| \leq$$
$$\leq K^2 \left\| \vec{x}^{(k-1)} - \vec{a} \right\| \leq \ldots \leq K^{k+1} \left\| \vec{x}^{(0)} - \vec{a} \right\|.$$

Protože $K < 1$, dostaneme zlimicením posledního vztahu $\vec{x}^{(k)} \to \vec{a}$. □

Tvrzení 1.4. Charakter konvergence Newtonovy metody je kvadratický, tj.

$$\left\| \vec{x}^{(k+1)} - \vec{a} \right\| \leq L \left\| \vec{x}^{(k)} - \vec{a} \right\|^2.$$

Důkaz. Přesahuje rámec přednášky. („Je dlouhý.“) □