Komentář k oddílu X.3 – Taylorův polynom druhého řádu pro funkce více proměnných

O významu tohoto oddílu a uvedených definic:

- Hlavním cílem tohoto oddílu je zformulovat a dokázat Větu X.9, kterou použijeme v následující kapitole ke studiu lokálních extrémů funkcí více prommměnných.

- Věta X.9 je určitou analogií Věty X.1 (o Taylorově polynomu se zbytkem v Peanově tvaru) pro funkce více proměnných. Pro funkce více proměnných je však situace značně složitější.

- V oddílu X.1 jsme se zabývali Taylorovým polynomem funkcí jedné proměnné. Přitom k tomu, aby byl definován Taylorův polynom k-tého řádu, stačí (a zároveň je potřeba), aby funkce měla vlastní k-tou derivaci v příslušném bodě.

  A jakmile je Taylorův polynom definován, již platí tvrzení Věty X.1 – Taylorův polynom dobře aproximuje funkci v blízkosti zvoleného bodu.

- Taylorův polynom k-tého řádu lze definovat i pro funkce více proměnných.

  Je to ovšem složitější – už vůbec formulovat přesnou analogii předpokladu existence vlastní k-té derivace v daném bodě je obtížné (a dělat to nebudeme).

  Snadno můžeme formulovat silnější podmínku – pokud f je třídy C^k na okolí bodu $\boldsymbol{a}$, pak lze definovat Taylorův polynom řádu k a platí analogie Věty X.1.

  Omezíme se však na případ $k = 1$ nebo $k = 2$, protože tyto případy využijeme a Taylorovy polynomy vyšších řádů jsou ještě složitější na zápis.

- Taylorův polynom prvního řádu již vlastně známe, byť ne pod tímto názvem.

  Pokud f je třídy C^1 na nějakém okolí bodu $\boldsymbol{a} \in \mathbf{R}^n$, pak Taylorův polynom prvního řádu funkce f v bodě $\boldsymbol{a}$ je definován vzorcem

$$T_1^{f,\boldsymbol{a}}(\boldsymbol{x}) = f(\boldsymbol{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{a})(x_i - a_i), \quad \boldsymbol{x} \in \mathbf{R}^n. \tag{$*$}$$

Je to tedy funkce, jejímž grafem je tečná nadrovina ke grafu funkce f v bodě $[\boldsymbol{a}, f(\boldsymbol{a})]$ (viz oddíl V.5).

Podle Věty V.12 pak platí

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} \frac{f(\boldsymbol{x}) - T_1^{f,\boldsymbol{a}}(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{a}\|} = 0,$$

což je vlastně analogie rovnosti z Věty X.1 pro funkce více proměných a $k = 1$.

- Rovnost $(*)$ lze zapsat také v maticovém tvaru. Připomeňme, že gradient funkce f v bodě $\boldsymbol{a}$ je definován jako

  $$\nabla f(\boldsymbol{a}) = \left[\frac{\partial f}{\partial x_1}(\boldsymbol{a}), \dots, \frac{\partial f}{\partial x_n}(\boldsymbol{a})\right].$$

  Můžeme tedy psát

  $$T_1^{f,\boldsymbol{a}}(\boldsymbol{x}) = f(\boldsymbol{a}) + \nabla f(\boldsymbol{a})\cdot(\boldsymbol{x}-\boldsymbol{a}) = f(\boldsymbol{a}) + < \nabla f(\boldsymbol{a}), \boldsymbol{x}-\boldsymbol{a} >, \quad \boldsymbol{x} \in \mathbf{R}^n.$$

  V první verzi je použito maticové násobení – gradient uvažujeme jako řádkový vektor typu $1 \times n$ a $\boldsymbol{x} - \boldsymbol{a}$ uvažujeme jako sloupcový vektor typu $n \times 1$.

  V druhé verzi je použit skalární součin (viz oddíl IX.3).

  Toto značení více připomíná definici Taylorova polynomu pro funkce jedné proměné.

- Nyní se přesuňme k Taylorovu polynomu druhého řádu.

  Předpokládáme, že f je třídy C^2 na okolí bodu $\boldsymbol{a}$. Pak je Taylorův polynom druhého řádu definován vzorcem

  $$\begin{aligned} T_2^{f,\boldsymbol{a}}(\boldsymbol{x}) = f(\boldsymbol{a}) &+ \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{a})(x_i - a_i) \\ &+ \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})(x_i - a_i)(x_j - a_j), \end{aligned} \qquad \boldsymbol{x} \in \mathbf{R}^n. \quad (**)$$

  Opět si ukážeme, že vzorec lze přepsat v maticovém tvaru.

  Část vzorce na prvním řádku je Taylorův polynom prvního řádu.

Část na druhém řádku vyjádříme pomocí Hessovy matice funkce f v bodě a, která je definována vzorcem

$$\nabla^2 f(\boldsymbol{a}) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}) \right)_{\substack{i=1,\dots,n\\ j=1,\dots,n}}.$$

Protože f je třídy C^2 na okolí bodu $\boldsymbol{a}$, z věty o záměně parciálních derivací (Věta V.15) plyne, že matice $\nabla^2 f(\boldsymbol{a})$ je symetrická.

S pomocí maticového násobení můžeme vzorec $(**)$ přepsat ve tvaru

$$\begin{aligned} T_2^{f,\boldsymbol{a}}(\boldsymbol{x}) &= f(\boldsymbol{a}) + \nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x}-\boldsymbol{a}) + \tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^T \cdot \nabla^2 f(\boldsymbol{a}) \cdot (\boldsymbol{x}-\boldsymbol{a}) \\ &= f(\boldsymbol{a}) + < \nabla f(\boldsymbol{a}), \boldsymbol{x}-\boldsymbol{a} > + \tfrac{1}{2} < \nabla^2 f(a)(\boldsymbol{x}-\boldsymbol{a}), \boldsymbol{x}-\boldsymbol{a} >, \end{aligned} \qquad \boldsymbol{x} \in \mathbf{R}^n.$$

Tento vzorec opět formálně připomíná definici Taylorova polynomu druhého řádu pro funkci jedné proměnné.

Ještě si rozeberme tvar jednotlivých členů:

$$T_2^{f,\boldsymbol{a}}(\boldsymbol{x}) = f(\boldsymbol{a}) + \underbrace{\nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x}-\boldsymbol{a})}_{\substack{\text{lineární forma}\\ \text{reprezentovaná maticí } \nabla f(a)\\ \text{v bodě } \boldsymbol{x}-\boldsymbol{a}}} + \frac{1}{2} \underbrace{(\boldsymbol{x}-\boldsymbol{a})^T \cdot \nabla^2 f(\boldsymbol{a}) \cdot (\boldsymbol{x}-\boldsymbol{a})}_{\substack{\text{kvadratická forma}\\ \text{reprezentovaná maticí } \nabla^2 f(a)\\ \text{v bodě } \boldsymbol{x}-\boldsymbol{a}}}$$

- Věta X.9 říká, že za uvedených předpokladů je

$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} \frac{f(\boldsymbol{x}) - T_2^{f,\boldsymbol{a}}(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{a}\|^2} = 0,$$

což je analogie rovnosti z Věty X.1 pro funkce více proměnných a $k = 2$.

Důkaz Věty X.9:

- Důkaz provedeme s použitím Věty X.5. pro vhodnou funkci jedné proměnné.
- Mějme tedy bod $\boldsymbol{a} \in \mathbf{R}^n$, číslo $\Delta > 0$ a funkci f třídy C^2 na $B(\boldsymbol{a}, \Delta)$.

  Naším cílem je dokázat, že

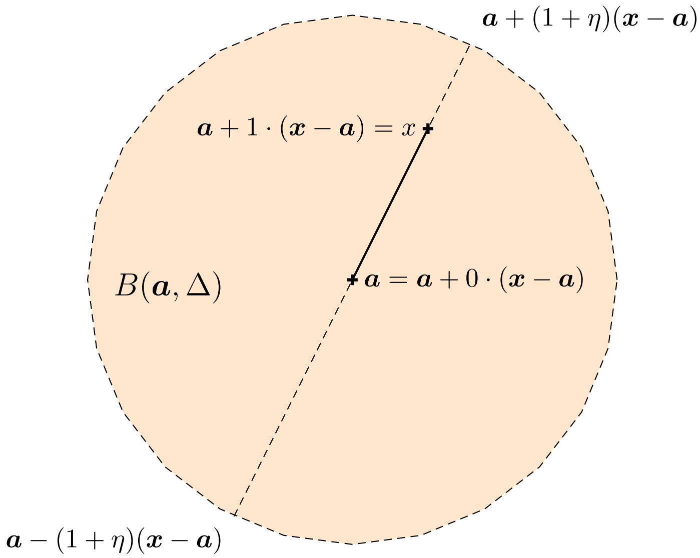
$$\lim_{\boldsymbol{x}\to\boldsymbol{a}} \frac{f(\boldsymbol{x}) - T_2^{f,\boldsymbol{a}}(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{a}\|^2} = 0.$$

  K tomuto účelu nejprve zvolíme pevné $\boldsymbol{x} \in B(\boldsymbol{a}, \Delta) \setminus \{\boldsymbol{a}\}$ a vhodně odhadneme $f(\boldsymbol{x}) - T_2^{f,\boldsymbol{a}}(\boldsymbol{x})$.

- První krok: Definujeme pomocnou funkci φ jedné proměnné vzorcem

  $$\varphi(t) = f(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a})).$$

  Tato funkce je definována na intervalu $(-1-\eta, 1+\eta)$ pro nějaké $\eta > 0$. To je ilustrováno na obrázku:



- Druhý krok: Aplikace Věty X.5 na funkci φ.

  Funkce φ je třídy C^2 na intervalu $(-1-\eta, 1+\eta)$. To plyne z věty o derivaci složené funkce (Věta V.14) – vnější funkce f je třídy C^2, vnitřní funkce jsou funkce tvaru

  $$t \mapsto a_i + t(x_i - a_i),$$

  protože

  $$\varphi(t) = f(a_1 + t(x_1 - a_1), \ldots, a_n + t(x_n - a_n)).$$

  Vnitřní funkce jsou tedy dokonce třídy C^∞.

  Jsou tedy splněny předpoklady Věty X.5 – pro funkci φ na intervalu $(-1-\eta, 1+\eta)$ pro $k = 1$.

  Podle Věty X.5 tedy existuje $\xi \in (0,1)$, pro které platí

  $$\varphi(1) = \varphi(0) + \varphi'(0) \cdot (1-0) + \frac{\varphi''(\xi)}{2} \cdot (1-0)^2 = \varphi(0) + \varphi'(0) + \frac{1}{2}\varphi''(\xi). \quad (\circ)$$

- Třetí krok: Vyjádřeme ($\circ$) pomocí f.

  Nejprve si uvědomme, že
  $$\varphi(0) = f(\boldsymbol{a}) \quad \text{a} \quad \varphi(1) = f(\boldsymbol{x}) \qquad (\triangle)$$
  Dále z definice φ a z Věty V.14 (o derivaci složené funkce) plyne
  $$\varphi'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a_1+t(x_1-a_1),\dots,a_n+t(x_n-a_n))(x_i-a_i), \quad t \in (-1-\eta, 1+\eta).$$
  Dosadíme-li $t = 0$, dostáváme
  $$\varphi'(0) = \frac{\partial f}{\partial x_i}(\boldsymbol{a})(x_i - a_i) = \nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a}). \qquad (\triangle\triangle)$$
  Opětovným použitím věty o derivaci složené funkce (Věta V.14) na vzorec pro φ' dostaneme
  $$\begin{aligned}\varphi''(t) &= \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a_1 + t(x_1 - a_1),\dots,a_n + t(x_n - a_n))(x_j - a_j)(x_i - a_i)\\ &= (\boldsymbol{x} - \boldsymbol{a})^T \nabla^2 f(\boldsymbol{a} + t(\boldsymbol{x} - \boldsymbol{a}))(\boldsymbol{x} - \boldsymbol{a}), \quad t \in (-1 - \eta, 1 + \eta).\end{aligned}$$
  Speciálně dosazením $t = \xi$ dostaneme
  $$\varphi''(\xi) = (\boldsymbol{x} - \boldsymbol{a})^T \nabla^2 f(\boldsymbol{a} + \xi(\boldsymbol{x} - \boldsymbol{a}))(\boldsymbol{x} - \boldsymbol{a}) \qquad (\triangle\triangle\triangle)$$
  Pokud ($\triangle$), ($\triangle\triangle$) a ($\triangle\triangle\triangle$) dosadíme do ($\circ$), dostaneme
  $$f(\boldsymbol{x}) = f(\boldsymbol{a})+\nabla f(\boldsymbol{a})\cdot(\boldsymbol{x}-\boldsymbol{a})+\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^T\nabla^2 f(\boldsymbol{a}+\xi(\boldsymbol{x}-\boldsymbol{a}))(\boldsymbol{x}-\boldsymbol{a}). \quad (\circ\circ)$$

- Čtvrtý krok: Vyjádřeme $f(\boldsymbol{x}) - T_2^{f,\boldsymbol{a}}(\boldsymbol{x})$ s použitím ($\circ\circ$).
  $$\begin{aligned}f(\boldsymbol{x}) - T_2^{f,\boldsymbol{a}}(x) &\overset{(\circ\circ)}{=} f(\boldsymbol{a}) + \nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})^T \nabla^2 f(\boldsymbol{a} + \xi(\boldsymbol{x} - \boldsymbol{a}))(\boldsymbol{x} - \boldsymbol{a})\\ &\qquad - \left(f(\boldsymbol{a}) + \nabla f(\boldsymbol{a}) \cdot (\boldsymbol{x} - \boldsymbol{a}) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})^T \nabla^2 f(\boldsymbol{a})(\boldsymbol{x} - \boldsymbol{a})\right)\\ &= \frac{1}{2}(\boldsymbol{x} - \boldsymbol{a})^T(\nabla^2 f(\boldsymbol{a} + \xi(\boldsymbol{x} - \boldsymbol{a})) - \nabla^2 f(\boldsymbol{a}))(\boldsymbol{x} - \boldsymbol{a}).\end{aligned}$$
  Pokud tuto rovnost vydělíme $\|\boldsymbol{x} - \boldsymbol{a}\|^2$, dostaneme rovnost
  $$\tfrac{f(\boldsymbol{x})-T_2^{f,\boldsymbol{a}}(x)}{\|\boldsymbol{x}-\boldsymbol{a}\|^2} = \frac{1}{2}\tfrac{(\boldsymbol{x}-\boldsymbol{a})^T}{\|\boldsymbol{x}-\boldsymbol{a}\|}(\nabla^2 f(\boldsymbol{a} + \xi(\boldsymbol{x} - \boldsymbol{a})) - \nabla^2 f(\boldsymbol{a}))\tfrac{\boldsymbol{x}-\boldsymbol{a}}{\|\boldsymbol{x}-\boldsymbol{a}\|}. \quad (\circ\circ\circ)$$

- Pátý krok: Závěrečný argument.

  Chceme dokázat, že $\lim\limits_{\boldsymbol{x}\to\boldsymbol{a}} \frac{f(\boldsymbol{x})-T_2^{f,\boldsymbol{a}}(x)}{\|\boldsymbol{x}-\boldsymbol{a}\|^2} = 0$. Jak nám k tomu pomůže $(\circ\circ\circ)$, je intuitivně vyjádřeno zde:

$$\frac{f(\boldsymbol{x})-T_2^{f,\boldsymbol{a}}(x)}{\|\boldsymbol{x}-\boldsymbol{a}\|^2} = \frac{1}{2} \underbrace{\frac{(\boldsymbol{x}-\boldsymbol{a})^T}{\|\boldsymbol{x}-\boldsymbol{a}\|}}_{\text{omezené}} \underbrace{\left(\nabla^2 f(\boldsymbol{a}+\xi(\boldsymbol{x}-\boldsymbol{a})) - \nabla^2 f(\boldsymbol{a})\right)}_{\to 0,\text{ protože } f \text{ je třídy } C^2} \underbrace{\frac{\boldsymbol{x}-\boldsymbol{a}}{\|\boldsymbol{x}-\boldsymbol{a}\|}}_{\text{omezené}} \longrightarrow 0.$$

  Protože se zde ale vyskytuje maticové násobení a navíc číslo ξ závislé na $\boldsymbol{x}$, uveďme přesný argument:

  V prvních čtyřech krocích jsme dokázali, že

$$\forall \boldsymbol{x} \in B(\boldsymbol{a},\Delta)\setminus\{\boldsymbol{a}\}\, \exists \xi \in (0,1)\colon \text{ platí } (\circ\circ\circ). \qquad (\square)$$

  Nechť $\varepsilon > 0$ je libovolné.

  Protože f je třídy C^2 na $B(a,\Delta)$, jsou parciální derivace druhého řádu spojité na $B(a,\Delta)$, speciálně jsou spojité v bodě $\boldsymbol{a}$.

  Tedy existuje $\delta \in (0,\Delta)$, že platí

$$\forall \boldsymbol{x} \in B(\boldsymbol{a},\delta)\, \forall i,j \in \{1,\dots,n\}\colon \left|\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}) - \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})\right| < \frac{\varepsilon}{n^2}. \qquad (\diamond)$$

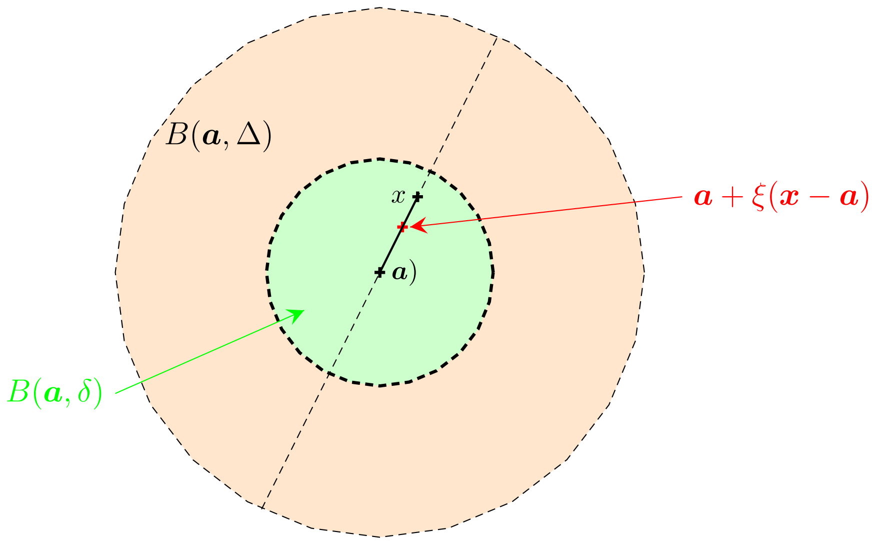
  Nyní zvolme libovolné $\boldsymbol{x} \in B(\boldsymbol{a},\delta)\setminus\{a\}$. Podle $(\square)$ najdeme $\xi \in (0,1)$, aby platilo $(\circ\circ\circ)$.

  Pak platí:

$$\begin{aligned}
\left|\frac{f(\boldsymbol{x})-T_2^{f,\boldsymbol{a}}(x)}{\|\boldsymbol{x}-\boldsymbol{a}\|^2}\right| &= \left|\frac{1}{2}\frac{(\boldsymbol{x}-\boldsymbol{a})^T}{\|\boldsymbol{x}-\boldsymbol{a}\|}\left(\nabla^2 f(\boldsymbol{a}+\xi(\boldsymbol{x}-\boldsymbol{a})) - \nabla^2 f(\boldsymbol{a})\right)\frac{\boldsymbol{x}-\boldsymbol{a}}{\|\boldsymbol{x}-\boldsymbol{a}\|}\right| \\
&= \frac{1}{2}\left|\sum_{i,j=1}^{n}\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}+\xi(\boldsymbol{x}-\boldsymbol{a})) - \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})\right)\cdot\frac{x_i-a_i}{\|x-\boldsymbol{a}\|}\cdot\frac{x_j-a_j}{\|x-\boldsymbol{a}\|}\right| \\
&\le \frac{1}{2}\sum_{i,j=1}^{n}\left|\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}+\xi(\boldsymbol{x}-\boldsymbol{a})) - \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})\right|\cdot\underbrace{\frac{|x_i-a_i|}{\|x-\boldsymbol{a}\|}}_{\le 1}\cdot\underbrace{\frac{|x_j-a_j|}{\|x-\boldsymbol{a}\|}}_{\le 1}, \\
&\le \frac{1}{2}\sum_{i,j=1}^{n}\left|\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}+\xi(\boldsymbol{x}-\boldsymbol{a})) - \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})\right|
\end{aligned}$$

kde jsme použili definice a trojúhelníkovou nerovnost.

Nyní si uvědomme, že $\boldsymbol{a} + \xi(\boldsymbol{x} - \boldsymbol{a})$ leží na úsečce $\boldsymbol{ax}$, tedy patří do $B(\boldsymbol{a}, \delta)$ (viz obrázek).



Proto

$$\left|\frac{f(\boldsymbol{x}) - T_2^{f,\boldsymbol{a}}(x)}{\|\boldsymbol{x}-\boldsymbol{a}\|^2}\right| \leq \frac{1}{2}\sum_{i,j=1}^{n} \underbrace{\left|\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a} + \xi(\boldsymbol{x} - \boldsymbol{a})) - \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})\right|}_{<\frac{\varepsilon}{n^2} \text{ dle } (\square)} < \frac{1}{2} n^2 \cdot \frac{\varepsilon}{n^2} = \frac{\varepsilon}{2} < \varepsilon.$$

Dokázali jsme, že

$$\forall \varepsilon > 0 \, \exists \delta > 0 \, \forall \boldsymbol{x} \in B(\boldsymbol{a}, \delta) \setminus \{\boldsymbol{a}\} : \left|\frac{f(\boldsymbol{x}) - T_2^{f,\boldsymbol{a}}(x)}{\|\boldsymbol{x}-\boldsymbol{a}\|^2}\right| < \varepsilon,$$

tedy

$$\lim_{\boldsymbol{x} \to \boldsymbol{a}} \frac{f(\boldsymbol{x}) - T_2^{f,\boldsymbol{a}}(x)}{\|\boldsymbol{x}-\boldsymbol{a}\|^2} = 0.$$

Tím je důkaz hotov.