Optimalizace bez omezení (unconstraint)

Nederivační (ad hoc) metody

Jednoduché metody

Nelder-Meadova (simplexová) metoda

Derivační metody

První derivace

Metoda největšího spádu + další spádové metody

Metoda konjugovaných gradientů

Druhá derivace

Newtonovské metody

Kvazi-Newtonovské metoda

Metody druhé derivace

Pro stanovení minima funkce f: R^N -> R využívají:

- První derivaci funkce f (gradient).
- Druhou derivaci funkce f (hessián).

  Poskytuje informace o zakřivení funkce v daném bodě.

Newtonovské metody
- *obecně*

Odvození:

Zapišme si funkci f: R -> R pomocí Taylorova polynomu:

$$f(x) = f(x^{(k)}) + (x - x^{(k)}).f'(x^{(k)}) + (x - x^{(k)})^2.f''(x^{(k)})/2 + ...$$

Pro první derivaci funkce f tedy platí:

$$f'(x) = f'(x^{(k)}) + (x - x^{(k)}).f''(x^{(k)}) + ...$$

Pokud předpokládáme, že funkci f lze v okolí bodu x považovat za kvadratickou funkci, můžeme výše uvedenou rovnici zjednodušit na následující tvar:

$$f'(x) = f'(x^{(k)}) + (x - x^{(k)}).f''(x^{(k)})$$

Newtonovské metody

- *obecně II*

V bodě $x^{(k+1)}$ tedy pro f platí:

$$f'(x^{(k+1)}) = f'(x^{(k)}) + (x^{(k+1)} - x^{(k)}).f''(x^{(k)})$$

Pokud se $x^{(k+1)}$ nachází poblíž minima, pak platí:

$$f'(x^{(k+1)}) = 0.$$

Takže výše uvedenou rovnici lze přepsat:

$$f''(x^{(k)}).\delta^{(k)} = -f'(x^{(k)})$$

kde $\delta^{(k)} = x^{(k+1)} - x^{(k)}$

Analogicky pro f: $R^N \to R$ platí:

$$G^{(k)}.\delta^{(k)} = -g^{(k)}$$

Newtonovské metody

- *obecně III*

Jednotlivé typy newtonovských metod se liší v postupu, jakým vyjadřují $G^{(k)}$:

- klasická Newtonova metoda počítá $G^{(k)}$ přesně
- kvazi-newtonovské metody využívají jistou její aproximaci $H^{(k)} \approx G^{(k)}$

Newtonovské metody
- *klasická Newtonova metoda*

Konstruuje $x^{(k+1)}$ přímo pomocí $G^{(k)}$.

V k-ém kroku Newtonovy metody se pak provedou následující operace:

a) Najdi řešení $\delta^{(k)}$ rovnice $G^{(k)}.\delta^{(k)} = -g^{(k)}$

b) Polož $x^{(k+1)} = x^{(k)} + \delta^{(k)}$

Newtonovské metody

- *klasická Newtonova metoda II*

Nejčastějším postupem, jak získat $\delta^{(k)}$ je tento:

$$\delta^{(k)} = - (G^{(k)})^{-1} . g^{(k)}$$

Pro $x^{(k+1)}$ tedy platí:

$$x^{(k+1)} = x^{(k)} - (G^{(k)})^{-1} . g^{(k)}$$

Newtonovské metody
- *klasická Newtonova metoda III*

Příklad:

funkce: $f(x, y) = x^2 + 2y^2$

bod: $x^{(0)} = (9, 9)^T$

Newtonovské metody
- *klasická Newtonova metoda III*

Příklad:

funkce: $f(x, y) = x^2 + 2y^2$

bod: $x^{(0)} = (9, 9)^T$

gradient v bodě $x^{(0)}$:

Newtonovské metody

- *klasická Newtonova metoda III*

Příklad:

funkce: $f(x, y) = x^2 + 2y^2$

bod: $x^{(0)} = (9, 9)^T$

gradient v bodě $x^{(0)}$: $g^{(0)} = (18, 36)^T$ $\qquad g = (2x, 4y)$

Newtonovské metody
- *klasická Newtonova metoda III*

Příklad:

funkce: $f(x, y) = x^2 + 2y^2$

bod: $x^{(0)} = (9, 9)^T$

gradient v bodě $x^{(0)}$: $g^{(0)} = (18, 36)^T$ $\qquad$ $g = (2x, 4y)$

hessian v bodě $x^{(0)}$:

Newtonovské metody
- *klasická Newtonova metoda III*

Příklad:

funkce: $f(x, y) = x^2 + 2y^2$

bod: $x^{(0)} = (9, 9)^T$

gradient v bodě $x^{(0)}$: $g^{(0)} = (18, 36)^T$ $\quad$ $g = (2x, 4y)$

hessian v bodě $x^{(0)}$: $h^{(0)} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$

Newtonovské metody
- *klasická Newtonova metoda III*

Příklad:

funkce: $f(x, y) = x^2 + 2y^2$

bod: $x^{(0)} = (9, 9)^T$

gradient v bodě $x^{(0)}$: $g^{(0)} = (18, 36)^T$ $\qquad$ $g = (2x, 4y)$

hessian v bodě $x^{(0)}$: $h^{(0)} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$

Newtonovské metody
- *klasická Newtonova metoda III*

Příklad:

funkce: $f(x, y) = x^2 + 2y^2$

bod: $x^{(0)} = (9, 9)^T$

gradient v bodě $x^{(0)}$: $g^{(0)} = (18, 36)^T$ $\quad$ $g = (2x, 4y)$

hessian v bodě $x^{(0)}$: $h^{(0)} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$

inverze hessianu:

$$h^{(0)^{-1}} = \begin{pmatrix} {h_{22}}/{D} & -{h_{21}}/{D} \\ -{h_{12}}/{D} & {h_{11}}/{D} \end{pmatrix}$$

Newtonovské metody
- *klasická Newtonova metoda III*

Příklad:

funkce: $f(x, y) = x^2 + 2y^2$

bod: $x^{(0)} = (9, 9)^T$

gradient v bodě $x^{(0)}$: $g^{(0)} = (18, 36)^T$ $g = (2x, 4y)$

hessian v bodě $x^{(0)}$: $h^{(0)} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$

inverze hessianu: $h^{(0)^{-1}} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/4 \end{pmatrix}$ $h^{(0)^{-1}} = \begin{pmatrix} h_{22}/D & -h_{21}/D \\ -h_{12}/D & h_{11}/D \end{pmatrix}$

Newtonovské metody
- klasická Newtonova metoda III

Příklad:

funkce: $f(x, y) = x^2 + 2y^2$

bod: $x^{(0)} = (9, 9)^T$

gradient v bodě $x^{(0)}$: $g^{(0)} = (18, 36)^T$ $\qquad$ $g = (2x, 4y)$

hessian v bodě $x^{(0)}$: $h^{(0)} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$

inverze hessianu: $h^{(0)^{-1}} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/4 \end{pmatrix}$ $\quad h^{(0)^{-1}} = \begin{pmatrix} h_{22}/D & -h_{21}/D \\ -h_{12}/D & h_{11}/D \end{pmatrix}$

bod $x^{(1)}$: $x^{(1)} =$

$$x^{(1)} = x^{(0)} - h^{(0)-1} . g^{(0)}$$

Newtonovské metody
- *klasická Newtonova metoda III*

Příklad:

funkce: $f(x, y) = x^2 + 2y^2$

bod: $x^{(0)} = (9, 9)^T$

gradient v bodě $x^{(0)}$: $g^{(0)} = (18, 36)^T$ $g = (2x, 4y)$

hessian v bodě $x^{(0)}$: $h^{(0)} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$

inverze hessianu: $h^{(0)^{-1}} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/4 \end{pmatrix}$

bod $x^{(1)}$: $x^{(1)} = \begin{pmatrix} 9 \\ 9 \end{pmatrix} - \begin{pmatrix} 1/2 & 0 \\ 0 & 1/4 \end{pmatrix} . \begin{pmatrix} 18 \\ 36 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$$x^{(1)} = x^{(0)} - h^{(0)-1} . g^{(0)}$$

Newtonovské metody
- *klasická Newtonova metoda IV*

Není-li k dispozici procedura pro výpočet hessiánu funkce f, lze jej aproximovat pomocí konečných diferencí z gradientu. i-tý sloupec γ^i hessiánu $G^{(k)}$ je pak odhadnut jako:

$$\gamma_i \approx \frac{g(x^{(k)} + h_i e_i) - g^{(k)}}{h_i}$$

kde e_i je i-tý jednotkový vektor a $h_i > 0$ je vhodně zvolené malé číslo.

Newtonovské metody
- *klasická Newtonova metoda V*

Problémy, spojené s použitím Newtonovy metody:

- Pokud matice $G^{(k)}$ není kladně definitní, pak metoda nemusí konvergovat k minimu funkce f
- Určení $(G^{(k)})^{-1}$ (tedy invertování matice $G^{(k)}$), má velkou časovou složitost, konkrétně: $O(N^3)$.
- Uživatel musí mít nejen algoritmus pro výpočet $f(x^{(k)})$ a $g^{(k)}$, ale i pro výpočet $G^{(k)}$.

Všechny uvedené problémy jsou vyřešeny v kvazi-newtonovských metodách.

Newtonovské metody
- *klasická Newtonova metoda VI*

Newtonova metoda konverguje kvadraticky v okolí minima, jak ukazuje věta:

Předpokládejme, že funkce f má spojité druhé derivace a její hessián je lipschitzovsky spojitý v nějakém okolí lokálního minima x^*. Pokud $x^{(k)}$ je dostatečně blízko x^* pro některé k a pokud G^* je kladně definitní, pak Newtonova metoda pro funkci f konverguje kvadraticky k x^*.

V praxi to znamená, že v každé iteraci se počet platných cifer odhadu řešení zhruba zdvojnásobuje.

Newtonovské metody
- *klasická Newtonova metoda VII*

Newtonovské metody lze využít i pro řešení soustav nelineárních rovnic.

Pokud řešíme rovnici $F(x) = 0$, kde $F: R^N \rightarrow R^N$, pak můžeme F rozvinout do Taylorovy řady jako:

$$F(x^{(k+1)}) \approx F(x^{(k)}) + \nabla F(x^{(k)})^T . \delta^{(k)}$$

kde $\delta^{(k)} = x^{(k+1)} - x^{(k)}$

pokud zanedbáme členy druhého a vyššího stupně.

Newtonovské metody

- *klasická Newtonova metoda VIII*

Pokud požadujeme, aby $F(x^{(k+1)}) = 0$, dostáváme z předchozí rovnice vztah, popisující Newtonovu metodu pro soustavu nelineárních rovnic:

$$- F(x^{(k)}) = \nabla F(x^{(k)})^T \, . \, \delta^{(k)T}$$

Newtonova metoda je zde opět dána postupem:

a) Najdi řešení $\delta^{(k)}$ rovnice $G^{(k)}.\delta^{(k)} = -g^{(k)}$

b) Polož $x^{(k+1)} = x^{(k)} + \delta^{(k)}$

Kde $G^{(k)} = \nabla F(x^{(k)})^T$ a vektor $g^{(k)}$ je nahrazen funkční hodnotou $F(x^{(k)})$.

Newtonovské metody
- *klasická Newtonova metoda IX*

V jednodimenzionárním případě dostáváme z rovnice: - $F(x^{(k)}) = \nabla F(x^{(k)})^T \cdot \delta^{(k)T}$ známou Newton-Raphsonovu metodu (někdy je také označována *metoda tečen*):

$$x^{(k+1)} = x^{(k)} - \frac{F(x^{(k)})}{F'(x^{(k)})}$$

Kvazi-newtonovské metody

V rámci Newtonovy metody je místem s největší složitostí výpočet inverzního hessiánu, nutného pro řešení rovnice:

$G^{(k)}.\delta^{(k)} = -g^{(k)}$

Tento krok lze obejít a místo inverzního hessiánu použít sérii matic, které se mu limitně blíží:

$$\lim_{k \to \infty} H^{(k)} = (G^{(k)})^{-1}$$

Kvazi-newtonovské metody

Algoritmus kvazi-newtonovských metod je modifikací algoritmu Newtonovy metody, doplněnou o lineární vyhledávání a rekurentní výpočet matice $H^{(k)}$ (která je nyní pouze aproximací hessiánu):

a) Najdi řešení $s^{(k)}$ rovnice $H^{(k)}.s^{(k)} = -g^{(k)}$

b) Najdi vhodné $\alpha^{(k)}$ a polož:

$$x^{(k+1)} = x^{(k)} + \alpha^{(k)} s^{(k)}$$

c) Vypočti $H^{(k+1)}$ na základě $H^{(k)}$ tak, aby platilo:

$$H^{(k+1)} s^{(k)} = y^{(k)}, \text{ kde } y^{(k)} = g(x^{(k+1)}) - g(x^{(k)})$$

Kvazi-newtonovské metody

Inicializace:

Matice H(0) může být libovolná pozitivně definitní matice.

Není-li známo, že minimalizační úloha má nějakou speciální strukturu, volí se obvykle $H^{(0)} = I$.

Kvazi-newtonovské metody

Kvazi-Newtonovské metody jsou obecně konstruovány tak, aby $H^{(k+1)}$ byla vždy symetrická a kladně definitní.

Díky tomuto faktu a díky své efektivitě jsou kvazi-newtonovské metody nejpoužívanějšími minimalizačními metodami založenými na Taylorově větě.

Kvazi-newtonovské metody

Pokud chceme zkonstruovat efektivní metodu pro generování matic $H^{(k)}$, měli bychom hledat rekurentní vyjádření $H^{(k+1)}$ z $H^{(k)}$. Výpočet $H^{(k+1)}$ podle této rekurentní formule by se měl dát provést pomocí relativně malého počtu operací, což v praxi znamená časovou složitost $O(N^2)$. Navíc by matice $H^{(k+1)}$ měla být symetrická, pokud tuto vlastnost má $H^{(k)}$.

Kvazi-newtonovské metody

Nejjednodušší formulí, splňující požadavky uvedené na předchozím slidu, je:

$$\mathbf{H}^{(k+1)} = \mathbf{H}^{(k)} + \alpha\, \mathbf{u}\mathbf{u}^T$$

kde **u** je jistý vektor (obecně závislý na $\mathbf{H}^{(k)}$, $\mathbf{s}^{(k)}$ a $\mathbf{y}^{(k)}$). Matice $\mathbf{u}\mathbf{u}^T$ má řád 1 a její výpočet vyžaduje jen $n^2 + n$ násobení.

Z rovnice $\mathbf{H}^{(k+1)}\mathbf{s}^{(k)} = \mathbf{y}^{(k)}$ plyne:

$$\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{s}^{(k)} = \alpha\, \mathbf{u}\mathbf{u}^T\, \mathbf{s}^{(k)}$$

Kvazi-newtonovské metody

Odtud je vidět, že vektor $\mathbf{u}$ je skalárním násobkem vektoru $\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{s}^{(k)}$, neboť $\alpha\mathbf{u}\mathbf{u}^T\,\mathbf{s}^{(k)} = \mathbf{u}(\alpha\mathbf{u}^T\,\mathbf{s}^{(k)})$ a $(\alpha\mathbf{u}^T\,\mathbf{s}^{(k)}) \in R$. Bez újmy na obecnosti (z hlediska vektoru $\mathbf{u}$) můžeme proto volit $\mathbf{u} = \mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{s}^{(k)}$.

Nyní můžeme rovnici $\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{s}^{(k)} = \alpha\,\mathbf{u}\mathbf{u}^T\,\mathbf{s}^{(k)}$ upravit následovně:

$$\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{s}^{(k)} = (\mathbf{y}^{(k)} - \mathbf{H}^{(k)}\mathbf{s}^{(k)})\alpha\mathbf{u}^T\,\mathbf{s}^{(k)}$$

Kvazi-newtonovské metody

Předchozí rovnici násobíme zleva vektorem:

$$\left(\left\|y^{(k)} - H^{(k)}.\delta^{(k)}\right\|_2^2\right)^{-1}.\left(y^{(k)} - H^{(k)}.\delta^{(k)}\right)^T$$

Získáme rovnici $1 = \alpha\mathbf{u}^{\mathrm{T}}\,\mathbf{s}^{(\mathrm{k})}$, ze které po úpravě dostaneme vztah pro výpočet α:

$$\alpha = \frac{1}{\left(y^{(k)} - H^{(k)}.\delta^{(k)}\right)^T.\delta^{(k)}}$$

Kvazi-newtonovské metody

Dosadíme-li vztahy pro **u** a α do rovnice:

$\mathbf{H}^{(k+1)} = \mathbf{H}^{(k)} + \alpha\, \mathbf{uu}^{T}$

Získáme vztah, který lze využít pro výpočet $\mathbf{H}^{(k+1)}$:

$$H^{(k+1)} = H^{(k)} + \frac{\left(y^{(k)} - H^{(k)}.\delta^{(k)}\right).\left(y^{(k)} - H^{(k)}.\delta^{(k)}\right)^{T}}{\left(y^{(k)} - H^{(k)}.\delta^{(k)}\right)^{T}.\delta^{(k)}}$$

Lze ukázat, že tato formule obecně nezachovává pozitivní definitnost matice H.

Kvazi-newtonovské metody

Je proto vhodné analyzovat formule druhého řádu:

$$\mathbf{H}^{(k+1)} = \mathbf{H}^{(k)} + \alpha\, \mathbf{u}\mathbf{u}^{T} + \beta\, \mathbf{v}\mathbf{v}^{T}$$

Zde již nejsou vektory u a v určeny jednoznačně.

Po dosazení do vztahu:

$$\mathbf{H}^{(k+1)}\mathbf{s}^{(k)} = \mathbf{y}^{(k)}$$

Dostaneme:

$$\mathbf{y}^{(k)} = \mathbf{H}^{(k)}\,\mathbf{s}^{(k)} + \alpha\, \mathbf{u}\mathbf{u}^{T}\,\mathbf{s}^{(k)} + \beta\, \mathbf{v}\mathbf{v}^{T}\,\mathbf{s}^{(k)}$$

Kvazi-newtonovské metody

Odtud analogicky jako v předchozím případě odvodíme vztah pro výpočet $H^{(k+1)}$:

$$H_{BFGS}^{(k+1)} = H + \frac{yy^T}{y^T s} - \frac{Hss^T H}{s^T Hs}$$

horní indexy (k) na pravé straně vynecháváme.

Tuto metodu navrhli v roce 1970 nezávisle čtyři autoři: Broyden, Fletcher, Goldfard a Shanno. Proto je označována jako metoda BFGS.

Kvazi-newtonovské metody

Jinou kvazi-newtonovskou metodu dostaneme, když budeme hledat řešení vyhovující rekurentní formuli druhého řádu, ale pro inverzi matice H(k). Položíme-li $B^{(k)} = (H^{(k)})^{-1}$, jde o formuli:

$$\mathbf{B}^{(k+1)} = \mathbf{B}^{(k)} + \alpha\, \mathbf{u}\mathbf{u}^T + \beta\, \mathbf{v}\mathbf{v}^T$$

Pokud nyní volíme $\mathbf{u} = \mathbf{s}^{(k)}$ a $\mathbf{v} = \mathbf{B}^{(k)}\mathbf{y}^{(k)}$, dostáváme obdobným postupem jako u metody BFGS vztah:

$$B_{DFP}^{(k+1)} = B + \frac{ss^T}{s^T y} - \frac{Byy^T B}{y^T By}$$

Kvazi-newtonovské metody

Tato metoda byla v jisté podobě obsažena v práci Davidona z roku 1959 a později explicitně popsána Fletcherem a Powellem v roce 1963. Proto bývá označována zkratkou DFP.

Cvičení

- *klasická Newtonova metoda*

funkce: $f(\mathbf{x}) = 3x_1^2 + 4x_2^2 – 5x_1x_2 – 2x_1$

bod: $x^{(0)} = (5, 5)^T$

Cvičení

- *klasická Newtonova metoda*

funkce: $f(\mathbf{x}) = 3x_1^2 + 4x_2^2 - 5x_1x_2 - 2x_1$

bod: $x^{(0)} = (5, 5)^T$

gradient v bodě $x^{(0)}$:

Cvičení

- *klasická Newtonova metoda*

funkce: $f(\mathbf{x}) = 3x_1^2 + 4x_2^2 - 5x_1x_2 - 2x_1$

bod: $x^{(0)} = (5, 5)^T$

gradient v bodě $x^{(0)}$:

$$g^{(0)} = \left(6x_1 - 5x_2 - 2, \quad 8x_2 - 5x_1\right)^T =$$

Cvičení

- *klasická Newtonova metoda*

funkce: $f(\mathbf{x}) = 3x_1^2 + 4x_2^2 - 5x_1x_2 - 2x_1$

bod: $x^{(0)} = (5, 5)^T$

gradient v bodě $x^{(0)}$:

$$g^{(0)} = \left(6x_1 - 5x_2 - 2,\quad 8x_2 - 5x_1\right)^T = \left(3,\ 15\right)^T$$

Cvičení

- *klasická Newtonova metoda*

funkce: $f(\mathbf{x}) = 3x_1^2 + 4x_2^2 – 5x_1x_2 – 2x_1$

bod: $x^{(0)} = (5, 5)^T$

gradient v bodě $x^{(0)}$:

$$g^{(0)} = \left(6x_1 - 5x_2 - 2,\ \ 8x_2 - 5x_1\right)^T = \left(3,\ 15\right)^T$$

hessian v bodě $x^{(0)}$:

Cvičení
- *klasická Newtonova metoda*

funkce: $f(\mathbf{x}) = 3x_1^2 + 4x_2^2 - 5x_1x_2 - 2x_1$

bod: $x^{(0)} = (5, 5)^T$

gradient v bodě $x^{(0)}$:

$$g^{(0)} = \left(6x_1 - 5x_2 - 2, \quad 8x_2 - 5x_1\right)^T = \left(3,\ 15\right)^T$$

hessian v bodě $x^{(0)}$:

$$h^{(0)} = \begin{pmatrix} 6 & -5 \\ -5 & 8 \end{pmatrix}$$

Cvičení

- *klasická Newtonova metoda*

funkce: $f(\mathbf{x}) = 3x_1^2 + 4x_2^2 – 5x_1x_2 – 2x_1$

bod: $x^{(0)} = (5, 5)^T$

gradient v bodě $x^{(0)}$:

$$g^{(0)} = \left(6x_1 - 5x_2 - 2,\quad 8x_2 - 5x_1\right)^T = \left(3,\ 15\right)^T$$

hessian v bodě $x^{(0)}$:

$$h^{(0)} = \begin{pmatrix} 6 & -5 \\ -5 & 8 \end{pmatrix}$$

inverze hessianu:

$$h^{(0)^{-1}} = \begin{pmatrix} h_{22}/D & -h_{21}/D \\ -h_{12}/D & h_{11}/D \end{pmatrix}$$

Cvičení

- *klasická Newtonova metoda*

funkce: $f(\mathbf{x}) = 3x_1^2 + 4x_2^2 – 5x_1x_2 – 2x_1$

bod: $x^{(0)} = (5, 5)^T$

gradient v bodě $x^{(0)}$:

$$g^{(0)} = \left(6x_1 - 5x_2 - 2,\quad 8x_2 - 5x_1\right)^T = \left(3,\ 15\right)^T$$

hessian v bodě $x^{(0)}$:

$$h^{(0)} = \begin{pmatrix} 6 & -5 \\ -5 & 8 \end{pmatrix}$$

inverze hessianu:

$$h^{(0)^{-1}} = \begin{pmatrix} {h_{22}}/{D} & -{h_{21}}/{D} \\ -{h_{12}}/{D} & {h_{11}}/{D} \end{pmatrix}$$

Cvičení

- *klasická Newtonova metoda*

funkce: $f(\mathbf{x}) = 3x_1^2 + 4x_2^2 - 5x_1x_2 - 2x_1$

bod: $x^{(0)} = (5, 5)^T$

gradient v bodě $x^{(0)}$:

$$g^{(0)} = \left(6x_1 - 5x_2 - 2,\quad 8x_2 - 5x_1\right)^T = \left(3,\ 15\right)^T$$

hessian v bodě $x^{(0)}$:

$$h^{(0)} = \begin{pmatrix} 6 & -5 \\ -5 & 8 \end{pmatrix}$$

inverze hessianu:

$$h^{(0)^{-1}} = \begin{pmatrix} h_{22}/D & -h_{21}/D \\ -h_{12}/D & h_{11}/D \end{pmatrix} = \begin{pmatrix} 8/23 & 5/23 \\ 5/23 & 6/23 \end{pmatrix}$$

bod $x^{(1)}$:

$$x^{(1)} = \begin{pmatrix} 5 \\ 5 \end{pmatrix} - \begin{pmatrix} 8/23 & 5/23 \\ 5/23 & 6/23 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 15 \end{pmatrix} = \begin{pmatrix} 5 \\ 5 \end{pmatrix} - \begin{pmatrix} 99/23 \\ 105/23 \end{pmatrix} = \begin{pmatrix} 16/23 \\ 10/23 \end{pmatrix}$$

Cvičení
- *klasická Newtonova metoda*

Rosenbrockova funkce:

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

Výchozí bod:

$x_0 = (-2, 2)$

Naprogramujte klasickou Newtonovu metodu.

Vyzkoušejte pro Rosenbrockovu funkci.