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*
  - Newton-Raphsonova metoda
  - Quasi-Newtonova metoda

Spádové metody
-obecně

Algoritmus:

- zvolíme výchozí bod $x^{(0)}$
- k-tá iterace:

  bod $x^{(k+1)}$ vypočítáme z bodu $x^{(k)}$ pomocí vztahu:

$$x^{(k+1)} = x^{(k)} + \alpha . s^{(k)}, \text{ kde:}$$

$s^{(k)}$	**směr přesunu** z bodu $x^{(k)}$
α	koeficient, popisující **délku daného přesunu**

Metoda největšího spádu
-*obecně*

$s^{(k)} = -g^{(k)}$

Algoritmus:

- zvolíme výchozí bod $x^{(0)}$
- k-tá iterace:

bod $x^{(k+1)}$ vypočítáme z bodu $x^{(k)}$ pomocí vztahu:

$$x^{(k+1)} = x^{(k)} - \alpha . g^{(k)}, \text{ kde:}$$

$-g^{(k)}$ zjednodušený zápis $-\nabla f(x^{(k)})$, určuje **směr přesunu** z bodu $x^{(k)}$

α koeficient, popisující **délku daného přesunu**

Metoda největšího spádu

- *volba α v metodě největšího spádu*

Z bodu $x^{(k)}$ se přesunujeme po polopřímce:

$x(\alpha) = x^{(k)} + \alpha.x^{(k)}$, kde $\alpha > 0$

Hodnotu funkce f na této polopřímce popisuje funkce $\phi(\alpha)$: $\phi(\alpha) = f(x(\alpha))$

Je zřejmé, že musíme zvolit takové α^{OK}, aby platilo:

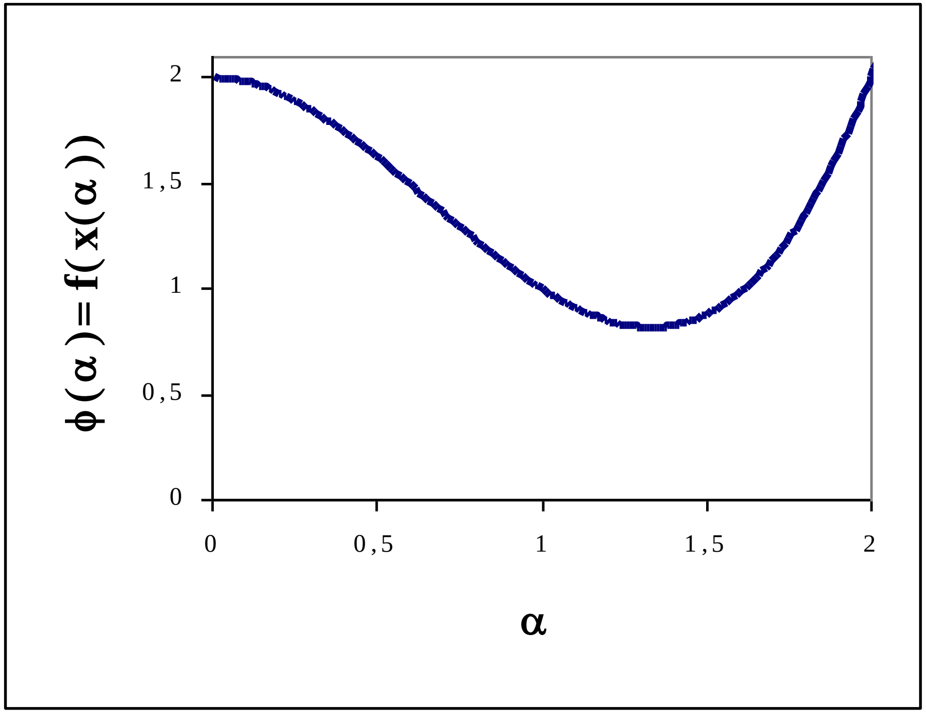
$f(x^{(k)}) > f(x^{(k+1)})$, kde $x^{(k+1)} = x(\alpha^{OK})$ pro dostatečný počet iterací.

Poznámka: „Dostatečný počet“ = dostačuje k tomu, aby metoda konvergovala k minimu ($\lim_{k\to\infty} f(x^{(k)}) = 0$).

Metoda největšího spádu

- *volba a v metodě největšího spádu*

Funkce $\phi(\alpha)$ má následující tvar:



Metoda největšího spádu

- *volba a v metodě největšího spádu*

Metoda největšího spádu volí pro každý krok stejnou hodnotu α.

Konkrétně velmi malou hodnotu α.

Poznámka: Hodnoty α musí být dostatečně malá, aby metoda konvergovala.

Metoda největšího spádu
zhodnocení

Výhody:

- Implementačně jednoduché
- Nízká prostorová složitost

Nevýhody:

- Velmi pomalá konvergence (speciálně v oblastech malého spádu => nízkých hodnot gradientu).
- Chyby, způsobené zaokrouhlením. Mohou vést i k tomu, že se výpočet vůbec nedostane rozumně blízko k minimu. Ale při (ideální) přesné aritmetice metoda konverguje vždy k nějakému lokálnímu minimu.

Funkce: $f(x_1, x_2) = x_1^2 + 2x_2^2$

Výchozi bod: $x^0 = (2, 1)$

Parametry: $\alpha = 0.25$

Funkce: $f(x_1, x_2) = x_1^2 + 2x_2^2$

Výchozi bod: $x^0 = (2, 1)$

Parametry: $\alpha = 0.25$

Gradient funkce: $\nabla f(x_1, x_2) = (2x_1, 4x_2)$

Funkce: $f(x_1, x_2) = x_1^2 + 2x_2^2$
Výchozi bod: $x^0 = (2, 1)$
Parametry: $\alpha = 0.25$
Gradient funkce: $\nabla f(x_1, x_2) = (2x_1, 4x_2)$

Směr funkce: $s^k = -g^k$

Následující bod: $x^{k+1} = x^k + \alpha . s^k$

Gradient v bodě x^0: $g^0 =$

Funkce: $f(x_1, x_2) = x_1^2 + 2x_2^2$

Výchozí bod: $x^0 = (2, 1)$

Parametry: $\alpha = 0.25$

Gradient funkce: $\nabla f(x_1, x_2) = (2x_1, 4x_2)$

Směr funkce: $s^k = -g^k$

Následující bod: $x^{k+1} = x^k + \alpha . s^k$

Gradient v bodě x^0: $g^0 = (4,4)$

Směr v bodě x^0: $s^0 =$

Funkce: $f(x_1, x_2) = x_1^2 + 2x_2^2$
Výchozi bod: $x^0 = (2, 1)$
Parametry: $\alpha = 0.25$
Gradient funkce: $\nabla f(x_1, x_2) = (2x_1, 4x_2)$

Směr funkce: $s^k = -g^k$

Následujicí bod: $x^{k+1} = x^k + \alpha.s^k$

Gradient v bodě x^0: $g^0 = (4,4)$
Směr v bodě x^0: $s^0 = (-4,-4)$
Bod x^1: x^1 =

Funkce: $f(x_1, x_2) = x_1^2 + 2x_2^2$
Výchozi bod: $x^0 = (2, 1)$
Parametry: $\alpha = 0.25$
Gradient funkce: $\nabla f(x_1, x_2) = (2x_1, 4x_2)$

Směr funkce: $s^k = -g^k$

Následujici bod: $x^{k+1} = x^k + \alpha.s^k$

Gradient v bodě x^0: $g^0 = (4,4)$
Směr v bodě x^0: $s^0 = (-4,-4)$
Bod x^1: $x^1 = (2,1) + 0{,}25.(-4,-4) = (1,0)$

Gradient v bodě x^1: $g^1 =$

Funkce: $f(x_1, x_2) = x_1^2 + 2x_2^2$

Výchozí bod: $x^0 = (2, 1)$

Parametry: $\alpha = 0.25$

Gradient funkce: $\nabla f(x_1, x_2) = (2x_1, 4x_2)$

Směr funkce: $s^k = -g^k$

Následující bod: $x^{k+1} = x^k + \alpha . s^k$

Gradient v bodě x^0: $g^0 = (4,4)$

Směr v bodě x^0: $s^0 = (-4,-4)$

Bod x^1: $x^1 = (2,1) + 0,25.(-4,-4) = (1,0)$

Gradient v bodě x^1: $g^1 = (2,0)$

Směr v bodě x^1: $s^1 =$

Funkce: $f(x_1, x_2) = x_1^2 + 2x_2^2$
Výchozí bod: $x^0 = (2, 1)$
Parametry: $\alpha = 0.25$
Gradient funkce: $\nabla f(x_1, x_2) = (2x_1, 4x_2)$

Směr funkce: $s^k = -g^k$

Následující bod: $x^{k+1} = x^k + \alpha . s^k$

Gradient v bodě x^0: $g^0 = (4,4)$
Směr v bodě x^0: $s^0 = (-4,-4)$
Bod x^1: $x^1 = (2,1) + 0,25.(-4,-4) = (1,0)$

Gradient v bodě x^1: $g^1 = (2,0)$

Směr v bodě x^1: $s^1 = (-2,0)$
Bod x^2: $x^2 =$

Funkce: $f(x_1, x_2) = x_1^2 + 2x_2^2$
Výchozi bod: $x^0 = (2, 1)$
Parametry: $\alpha = 0.25$
Gradient funkce: $\nabla f(x_1, x_2) = (2x_1, 4x_2)$

Směr funkce: $s^k = -g^k$

Následujici bod: $x^{k+1} = x^k + \alpha . s^k$

Gradient v bodě x^0: $g^0 = (4,4)$
Směr v bodě x^0: $s^0 = (-4,-4)$
Bod x^1: $x^1 = (2,1) + 0,25.(-4,-4) = (1,0)$

Gradient v bodě x^1: $g^1 = (2,0)$

Směr v bodě x^1: $s^1 = (-2,0)$
Bod x^2: $x^2 = (1,0) + 0,25.(-2,0) = (0.5,0)$

Metoda největšího spádu
- *příklad*

Rosenbrockova funkce:

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

Gradient Rosenbrockovy funkce:

$$\nabla f(x) = \left(-400x_1(x_2 - x_1^2) - 2(1 - x_1), \quad 200(x_2 - x_1^2)\right)^T$$

Výchozí bod:

$x_0 = (-2, 2)$

Parametry:

$\alpha = 0{,}001$

Vyzkoušejte doma :-)