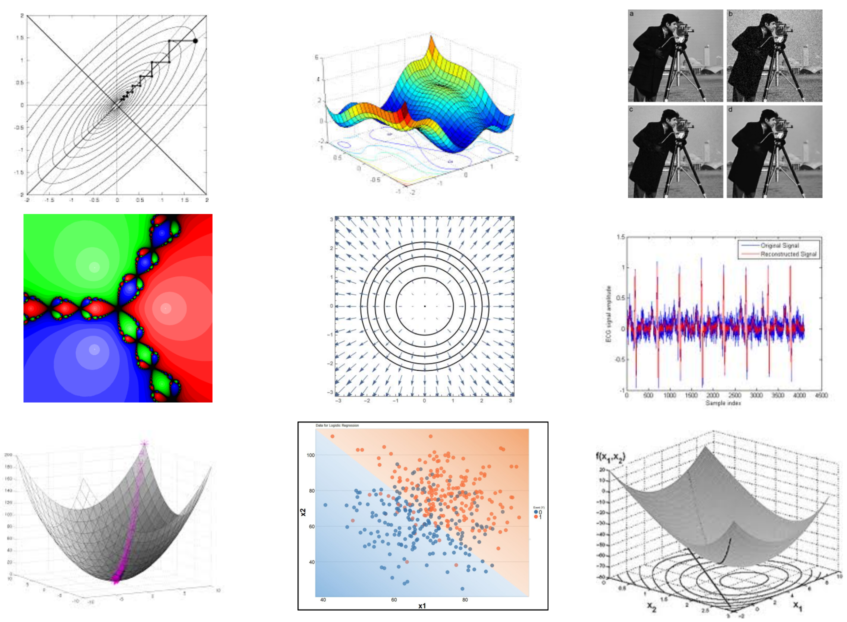
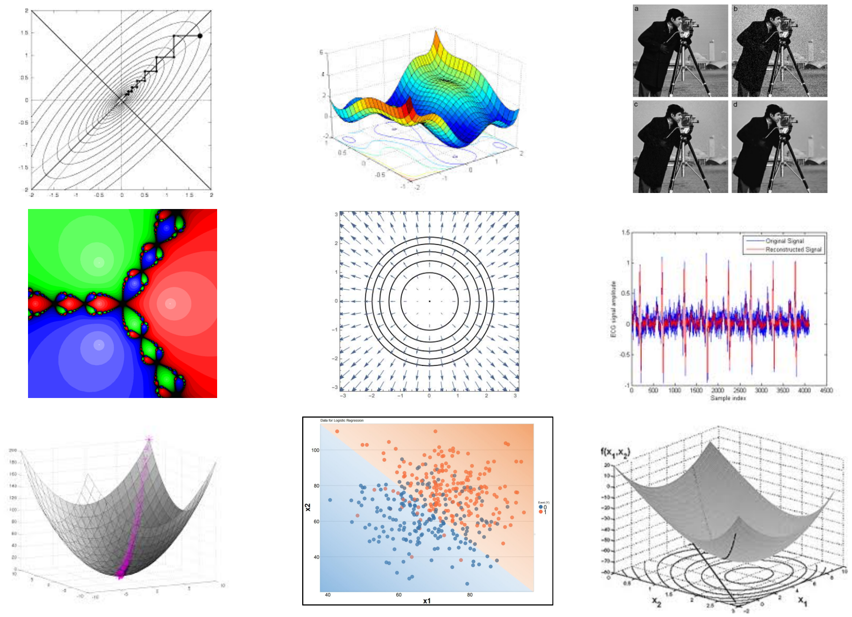


Metódy voľnej optimalizácie - letný semester 2018

M. Trnovská, KAMŠ, FMFI UK



Metóda cyklickej súradnicovej redukcie

Metóda cyklickej súradnicovej redukcie

V tejto metóde sa za smery s_k volia **jednotkové vektory** a to postupne $e_1, e_2, \ldots, e_n$ a potom odznova cyklicky. Teda

$$s_k = e_i, \quad i \equiv 1 + (k \mod n), \quad k = 0, 1, 2, \ldots.$$

Dĺžku kroku volíme **optimálnu**, t. j.

$$\lambda_k = \arg\min_{\lambda} f(x^k + \lambda s_k).$$

Zrejme takéto smery vo všeobecnosti **nie sú spádové** - optimálna dĺžka kroku môže byť záporné číslo.

Z hľadiska praktickej realizácie je takáto metóda menej efektívna ako gradientné metódy.

Metóda cyklickej súradnicovej redukcie

Výnimka: separovateľné funkcie:

$$f(x) = \sum_{i=1}^{n} f_i(x_i)$$

Metóda CRS konverguje za n krokov.

Paradox: Konvergenčná veta vyžaduje diferencovateľnosť funkcie f.

Konvergenčná veta: Nech $f(x)$ je konvexná na $\mathbb{R}^n$ a má spojité prvé parciálne derivácie. Nech štartovací bod x^0 je taký, že množina

$$K_0 = \{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\}$$

je ohraničená. Potom metóda cyklickej súradnicovej redukcie konverguje k optimálnemu riešeniu úlohy $(U1)$.

Metóda cyklickej súradnicovej redukcie

Príklad.

Daná je funkcia dvoch premenných

$$f(x_1, x_2) = \frac{1}{2}(x_1^2 + x_2^2) - x_1 - x_2 + |x_1 - x_2| =$$

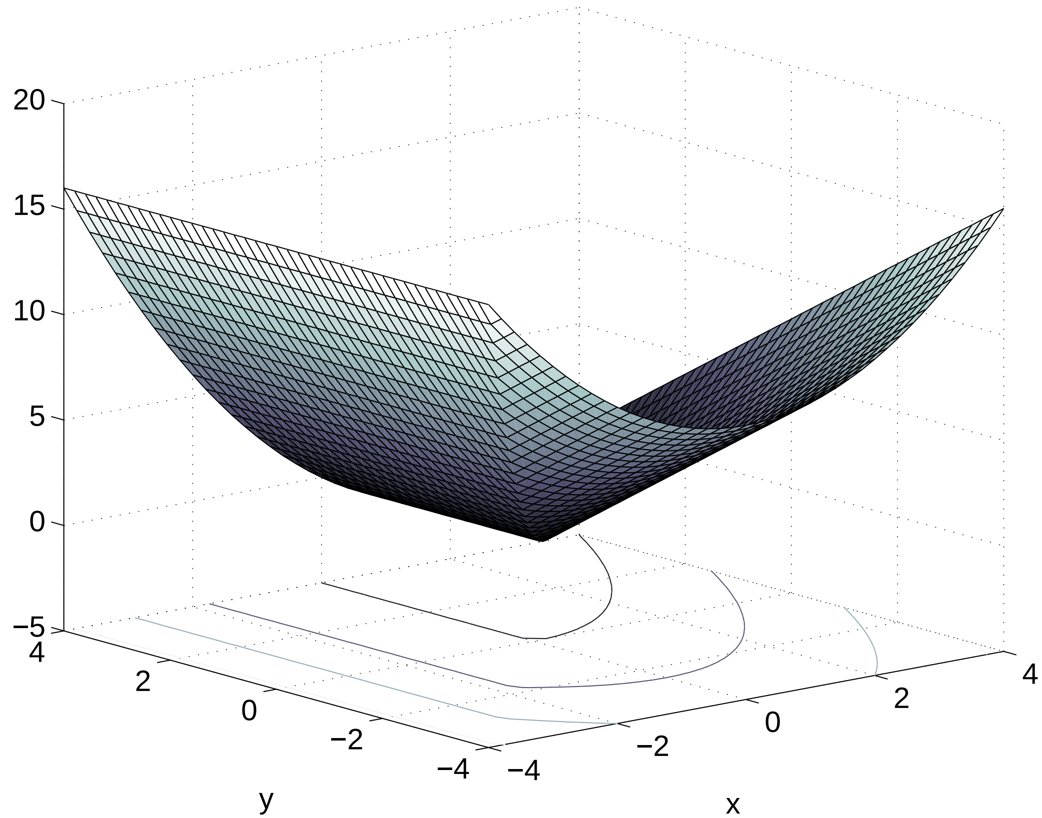
$$= \begin{cases} \frac{1}{2}(x_1^2 + x_2^2) - 2x_1 & x_1 \leq x_2 \\ \frac{1}{2}(x_1^2 + x_2^2) - 2x_2 & x_1 \geq x_2 \end{cases}$$

Optimálne riešenie je jediné (je to rýdzo konvexná funkcia) - je to bod $\hat{x} = [1, 1]^T$.

Dá sa ukázať, že metóda CSR so štartovacím bodom $x^0 = [0, 0]^T$ zlyhá.

Domáca úloha: Čo sa stane ak budeme štartovať z bodov $x^0 = [0, 1]^T$, resp. $x^0 = [0, 2]^T$?

Metóda cyklickej súradnicovej redukcie



Metóda cyklickej súradnicovej redukcie

Zhrnutie

Konvergencia metódy cyklickej súradnicovej redukcie je v porovnaní s gradientnými metódami pomalá, výnimkou sú separovateľné funkcie, pre ktoré metóda konverguje za n krokov.

Paradoxne metóda nemusí konvergovať ano v prípade konvexných funkcí, ak nie sú diferencovateľné.

Hlavnou výhodou je to, že metóda nepotrebuje počítať smery.