

# METODY OBLICZENIOWE OPTYMALIZACJI

## 4

Arkadiusz Tomczyk

Institut Informatyki Politechniki Łódzkiej

10 kwietnia 2011

## Funkcja celu

$$f : \mathbb{R}^n \rightarrow \mathbb{R}$$

## Metody bezgradientowe

- **Metoda Hooke-Jeevesa.**
- **Metod Rosenbrocka.**
- **Simpleksu Nelder-Meada.**
- Metoda Gaussa-Seidla.
- Metoda DSC.
- Metoda Powella
- Metoda Zangwilla.

## Założenia

Funkcja celu jest funkcją wypukłą.

## Metoda Hooke'a-Jeevesa

- Niech  $\mathbf{x} \in \mathbb{R}^n$  oznacza początkowe rozwiązanie oraz niech  $\mathbf{e}_i \in \mathbb{R}^n$  dla  $i = 1, \dots, n$  oznacza bazę wzajemnie ortogonalnych wektorów w przestrzeni  $\mathbb{R}^n$ . Ponadto niech  $0 < \beta < 1$  oznacza współczynnik korekcyjny zmniejszający długość kroków oraz niech  $\lambda \in \mathbb{R}$  oznacza początkową długość kroku.
- Zapamiętaj aktualne rozwiązanie jako  $\mathbf{x}_0 = \mathbf{x}$ . Jeśli jest to pierwsza iteracja oznacz również  $\mathbf{x}_b = \mathbf{x}$ .
- Dla każdego dla  $j = 1, \dots, n$  wykonaj:
  - Wyznacz  $\mathbf{x}_j = \mathbf{x}_{j-1} + \lambda \mathbf{e}_j$ .
  - Jeśli  $f(\mathbf{x}_j) \geq f(\mathbf{x}_{j-1})$  wyznacz  $\mathbf{x}_j = \mathbf{x}_{j-1} - \lambda \mathbf{e}_j$ .
  - Jeśli wciąż  $f(\mathbf{x}_j) \geq f(\mathbf{x}_{j-1})$  oznacz  $\mathbf{x}_j = \mathbf{x}_{j-1}$ .
- Jeśli wzdłuż żadnego kierunku nie udało się uzyskać poprawy przypisz  $\mathbf{x} = \mathbf{x}_b$  i zmniejsz długość kroku zgodnie ze wzorem  $\lambda = \beta\lambda$ . Jeśli udało się uzyskać poprawę wyznacz:
$$\mathbf{x} = \mathbf{x}_n + (\mathbf{x}_n - \mathbf{x}_b) = 2\mathbf{x}_n - \mathbf{x}_b$$
oraz przypisz  $\mathbf{x}_b = \mathbf{x}_n$ .
- Jeśli nie jest spełniony warunek stopu wróć do punktu drugiego.

## Ortogonalizacja Grama-Schmidta

Niech  $k \leq n$  oraz niech dany będzie zestaw liniowo niezależnych wektorów  $\mathbf{v}_i \in \mathbb{R}^n$  dla  $i = 1, \dots, k$ . W celu wyznaczenia zestawu ortogonalnych wektorów jednostkowych  $\mathbf{e}_i \in \mathbb{R}^n$  dla  $i = 1, \dots, k$  rozpinających tę samą podprzestrzeń  $\mathbb{R}^n$  należy zastosować następującą procedurę:

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{v}_1 \\ \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\ &\dots \\ \mathbf{u}_i &= \mathbf{v}_i - \sum_{j=1}^{i-1} \langle \mathbf{v}_i, \mathbf{e}_j \rangle \mathbf{e}_j \\ \mathbf{e}_i &= \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|} \end{aligned}$$

dla  $i = 2, \dots, k$ .

## Obrót współrzędnych

Niech  $\mathbf{e}_i \in \mathbb{R}^n$  dla  $i = 1, \dots, n$  oznacza bazę wzajemnie ortogonalnych wektorów oraz niech  $s_i \neq 0$  dla  $i = 1, \dots, n$  oznaczają długości kroków wzdłuż odpowiednich kierunków. Aby dokonać obrotu współrzędnych wyznacz wektory  $\mathbf{v}_i \in \mathbb{R}^n$  dla  $i = 1, \dots, n$ :

$$\mathbf{v}_i = \sum_{j=i}^n s_j \mathbf{e}_j$$

oraz zastosuj dla nich ortogonalizację Grama-Schmidta.

## Metoda Rosenbrocka

- Niech  $\mathbf{x} \in \mathbb{R}^n$  oznacza początkowe rozwiązanie oraz niech  $\mathbf{e}_i \in \mathbb{R}^n$  dla  $i = 1, \dots, n$  oznacza początkową bazę wzajemnie ortogonalnych wektorów w przestrzeni  $\mathbb{R}^n$ . Ponadto niech  $\alpha > 0$  i  $0 < \beta < 1$  oznaczają współczynniki korekcyjne, odpowiednio zwiększający i zmniejszający, długość kroków oraz niech  $\lambda_i \in \mathbb{R}$  dla  $i = 1, \dots, n$  oznaczają początkowe długości kroków dla odpowiednich kierunków bazy.
- Zapamiętaj aktualne rozwiązanie jako  $\mathbf{x}_0 = \mathbf{x}$ .
- Dla każdego dla  $j = 1, \dots, n$  wykonaj:
  - Wyznacz  $\mathbf{x}_j = \mathbf{x}_{j-1} + \lambda_j \mathbf{e}_j$ .
  - Jeśli  $f(\mathbf{x}_j) < f(\mathbf{x}_{j-1})$  zwiększ długość kroku  $\lambda_j = \alpha \lambda_j$ .
  - W przeciwnym razie przypisz  $\mathbf{x}_j = \mathbf{x}_{j-1}$ , zmień kierunek na  $\lambda_j = -\lambda_j$  oraz zmniejsz długość kroku  $\lambda_j = \beta \lambda_j$ .
- Powyższe kroki powtarzaj tak długo jak udaje się uzyskać lepsze rozwiązanie dla któregoś z kierunków bazy.
- Oblicz sumaryczne długości kroków  $s_i$  dla  $i = 1, \dots, n$  jakie wykonane zostały w każdym z kierunków  $\mathbf{e}_i$  od czasu ostatniej zmiany bazy wektorów.
- Jeśli nie jest spełniony warunek stopu dokonaj obrotu współrzędnych w celu wyznaczenia nowej bazy wektorów, a następnie wróć do punktu drugiego.

## Simpleks

Simpleksem  $n$ -wymiarowym o  $n + 1$  wierzchołkach nazywamy zbiór wszystkich punktów  $\mathbf{p} \in \mathbb{R}^n$  opisanych przez wektory wierzchołków  $\mathbf{p}_i \in \mathbb{R}^n$  dla  $i = 1, \dots, n + 1$  takich, że:

$$\mathbf{p} = \sum_{i=1}^{n+1} \theta_i \mathbf{p}_i$$

przy czym  $\theta_i \geq 0$  oraz

$$\sum_{i=1}^{n+1} \theta_i = 1;$$

## Oznaczenia

- Wierzchołek simpleksu, dla którego funkcja celu osiąga wartość największą:

$$\mathbf{p}_h \in \mathbb{R}^n \text{ dla } h \in \mathbb{N}$$

- Wierzchołek simpleksu, dla którego funkcja celu osiąga wartość najmniejszą:

$$\mathbf{p}_l \in \mathbb{R}^n \text{ dla } l \in \mathbb{N}$$

- Środek symetrii simpleksu po wyłączeniu wierzchołka  $\mathbf{p}_h$ :

$$\mathbf{p}^s = \frac{1}{n} \sum_{i=1}^{n+1} \mathbf{p}_i \text{ dla } i \neq h$$



### Odbicie

$$\mathbf{p}^o = (1 + \alpha)\mathbf{p}^s - \alpha\mathbf{p}_h \text{ dla } \alpha > 0$$

### Ekspansja

$$\mathbf{p}^e = (1 - \gamma)\mathbf{p}^o - \gamma\mathbf{p}^s \text{ dla } \gamma > 1$$

### Kontrakcja

$$\mathbf{p}^k = \beta\mathbf{p}_h + (1 - \beta)\mathbf{p}^s \text{ dla } 0 < \beta < 1$$

### Redukcja

$$\mathbf{p}_i = \frac{\mathbf{p}_i + \mathbf{p}_l}{2} \text{ dla } i = 1, \dots, n + 1$$

## Simpleks Nelder-Meada

- Niech  $\mathbf{p}_1, \dots, \mathbf{p}_{n+1}$  gdzie  $\mathbf{p}_i \in \mathbb{R}^n$  dla  $i = 1, \dots, n+1$  oznaczają początkowe wierzchołki simpleksu.
- Wyznacz  $f_h = f(\mathbf{p}_h)$  oraz  $f_l = f(\mathbf{p}_l)$ .
- Wyznacz środek symetrii simpleksu  $\mathbf{p}^s \in \mathbb{R}^n$  i przyjmij  $f^s = f(\mathbf{p}^s)$ , a następnie wykonaj odbicie znajdując  $\mathbf{p}^o \in \mathbb{R}^n$  i przyjmij  $f^o = f(\mathbf{p}^o)$ .
- Jeśli  $f^o < f_l$  to:
  - Dokonaj ekspansji znajdując  $\mathbf{p}^e \in \mathbb{R}^n$  i przyjmij  $f^e = f(\mathbf{p}^e)$ .
  - Jeśli  $f^e < f_h$  to przyjmij  $\mathbf{p}_h = \mathbf{p}^e$ , w przeciwnym razie przyjmij  $\mathbf{p}_h = \mathbf{p}^o$ .
  - Jeśli nie jest spełniony warunek stopu to wróć do punktu drugiego.
- Jeśli  $f^o \geq f_l$  to:
  - Jeśli  $f^o \geq f_h$  to dokonaj kontrakcji znajdując  $\mathbf{p}^k \in \mathbb{R}^n$  i przyjmij  $f^k = f(\mathbf{p}^k)$ . Jeśli  $f^k \geq f_h$  to dokonaj redukcji simpleksu, w przeciwnym razie przyjmij  $\mathbf{p}_h = \mathbf{p}^k$ .
  - W przeciwnym razie przyjmij  $\mathbf{p}_h = \mathbf{p}^o$ .
  - Jeśli nie jest spełniony warunek stopu to wróć do punktu drugiego.

## Kryteria stopu

- Zadana liczba iteracji.
- W metodzie Hooke'a-Jeevesa można zakończyć algorytm gdy długość kroku spadnie poniżej zadanej wartości.
- W metodzie Rosenbrocka kryterium zakończenia algorytmu może być brak postępów w kolejnych kierunkach aktualnej bazy.
- W metodzie simpleksu Nelder-Meada można zakończyć algorytm gdy odległość pomiędzy wierzchołkami simpleksu spadnie poniżej zadanej wartości.

## Rozwiązanie

W metodzie simpleksu Nelder-Meada za rozwiązanie można przyjąć środek ciężkości całego simpleksu.

## Brak postępów

Jeśli wybór rozwiązania początkowego w metodzie Hooke'a-Jeevesa lub metodzie Rosenbrocka nie daje poprawy wyniku po pierwszej iteracji można rozpocząć algorytm z nowego punktu początkowego.