METODY OBLICZENIOWE OPTYMALIZACJI

5

Arkadiusz Tomczyk

Instytut Informatyki Politechniki Łódzkiej

6 kwietnia 2011

Funkcja celu

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

Metody bezgradientowe

- Metoda Hooka-Jeevesa.
- Metod Rosenbrocka.
- Simpleksu Neldera-Meada.
- **Metoda Gaussa-Seidla**.
- **Metoda DSC**.
- **Metoda Powella**
- **Metoda Zangwilla**.

Założenia

Funkcja celu jest funkcją wypukłą, ograniczoną od dołu, posiadającą ciągłe drugie pochodne, a ponadto w pobliżu minimum może być dobrze aproksymowana formą kwadratową o postaci:

$$\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}\mathbf{x} + c$$

Metoda Gaussa-Seidla

- Niech $\mathbf{x} \in \mathbb{R}^n$ oznacza początkowe rozwiązanie oraz niech $\mathbf{e}_i \in \mathbb{R}^n$ dla $i = 1, \ldots, n$ oznacza bazę wzajemnie ortogonalnych wektorów w przestrzeni $\mathbb{R}^n$.
- Dla każdego kierunku $\mathbf{e}_i$ dla $i = 1, \ldots, n$ po kolei wykonaj optymalizację kierunkową funkcji $f(\mathbf{x} + \lambda \mathbf{e}_i)$ wyznaczając odpowiednie λ_i i przypisz $\mathbf{x} = \mathbf{x} + \lambda_i \mathbf{e}_i$.
- Jeśli nie spełnione zostało kryterium stopu wróć do punktu drugiego.

Uwagi

- Mała efektywaność przy długich wąskich dolinach zorientowanych wzdłuż kierunków innych niż kierunki wektorów należących do bazy.

Metoda DSC

- Niech $\mathbf{x} \in \mathbb{R}^n$ oznacza początkowe rozwiązanie oraz niech $\mathbf{e}_i \in \mathbb{R}^n$ dla $i = 1, \ldots, n$ oznacza początkową bazę wzajemnie ortogonalnych wektorów w przestrzeni $\mathbb{R}^n$. Ponadto $0 < \beta < 1$ oznaczają współczynnik korekcyjny zmniejszający długość krokół oraz niech $\eta_i \in \mathbb{R}$ dla $i = 1, \ldots, n$ oznaczają początkowe minimalne długości kroków dla odpowiednich kierunków bazy.
- Dla każdego kierunku $\mathbf{e}_i$ dla $i = 1, \ldots, n$ po kolei wykonaj optymalizację kierunkową funkcji $f(\mathbf{x} + \lambda \mathbf{e}_i)$ wyznaczając odpowiednie λ_i i przypisz $\mathbf{x} = \mathbf{x} + \lambda_i \mathbf{e}_i$.
- Oblicz sumaryczne długości kroków s_i dla $i = 1, \ldots, n$ jakie wykonane zostały w każdym z kierunków $\mathbf{e}_i$ od czasu ostatniej zmiany bazy wktorów.
- Jeśli przy zadanej bazie wktorów zachodzi $\lambda_i < \eta_i$ dla każdego $i = 1, \ldots, n$ i nie jest spełniony warunek stopu dokonaj obrotu współrzędnych w celu wyznaczenia nowej bazy wektorów, przypisz $\eta_i = \beta \eta_i$, a następnie wróć do punktu drugiego.

Uwagi

- Nazwa pochodzi od nazwisk Daviesa, Swanna i Campeya.
- Przy obrocie współrzędnych stosowane jest podejście znana z metody Rosenbrocka.

Kierunki sprzężone

Dwa kierunki $\mathbf{d}_1 \in \mathbb{R}^n$ oraz $\mathbf{d}_2 \in \mathbb{R}^n$ są wzajemnie sprzężone względem dodanio określonej macierzy $\mathbf{A}$ jeśli:

$$\mathbf{d}_1^T \mathbf{A} \mathbf{d}_2 = 0$$

Kierunki sprzężone

- Można wykazać, że kierunki wzajemnie sprzężone są liniowo niezależne.
- Jeśli kierunki $\mathbf{d}_i \in \mathbb{R}^n$ dla $i = 1, \ldots, n$ są sprzężone względem macierzy $\mathbf{A}$ definiującej pewną formę kwadratową to minimum dla tej formy kwadratowej może być wyznaczone poprzez minimalizację wzdłuż każdego kierunku tylko raz.
- Jeśli punkty $\mathbf{x}_1 \in \mathbb{R}^n$ oraz $\mathbf{x}_2 \in \mathbb{R}^n$ stanowią minimum formy kwadratowej wzdłuż tego samego kierunku $\mathbf{d} \in \mathbb{R}^n$ to kierunek $\mathbf{x}_2 - \mathbf{x}_1$ jest sprzężony z kierunkiem $\mathbf{d}$.

Metoda Powella 1

- Niech $\mathbf{x} \in \mathbb{R}^n$ oznacza początkowe rozwiązanie oraz niech $\mathbf{e}_i \in \mathbb{R}^n$ dla $i = 1, \ldots, n$ oznacza początkową bazę wzajemnie ortogonalnych wektorów w przestrzeni $\mathbb{R}^n$.
- Zapamiętaj aktualne rozwiązanie jako $\mathbf{x}_0 = \mathbf{x}$.
- Dla każdego kierunku $\mathbf{e}_i$ dla $i = 1, \ldots, n$ po kolei wykonaj optymalizację kierunkową funkcji $f(\mathbf{x}_{i-1} + \lambda \mathbf{e}_i)$ wyznaczając odpowiednie λ_i i przypisz $\mathbf{x}_i = \mathbf{x}_{i-1} + \lambda_i \mathbf{e}_i$.
- Wyznacz nowy kierunek sprzężony jako:

$$\mathbf{e} = \frac{\mathbf{x}_n - \mathbf{x}_0}{\|\mathbf{x}_n - \mathbf{x}_0\|}$$

- Wykonaj optymalizację kierunkową funkcji $f(\mathbf{x}_n + \lambda \mathbf{e})$ wyznaczając odpowiednie λ i przypisz $\mathbf{x} = \mathbf{x}_n + \lambda \mathbf{e}$.
- Dokonaj modyfikacji kierunków bazy w ten sposób, że $\mathbf{e}_i = \mathbf{e}_{i+1}$ dla $i = 1, \ldots, n-1$ oraz $\mathbf{e}_n = \mathbf{e}$.
- Jeśli nie spełnione zostało kryterium stopu wróć do punktu drugiego.

Metoda Powella 2

- Niech $\mathbf{x} \in \mathbb{R}^n$ oznacza początkowe rozwiązanie oraz niech $\mathbf{e}_i \in \mathbb{R}^n$ dla $i = 1, \ldots, n$ oznacza początkową bazę wzajemnie ortogonalnych wektorów w przestrzeni $\mathbb{R}^n$.
- Zapamiętaj aktualne rozwiązanie jako $\mathbf{x}_0 = \mathbf{x}$.
- Dla każdego kierunku $\mathbf{e}_i$ dla $i = 1, \ldots, n$ po kolei wykonaj optymalizację kierunkową funkcji $f(\mathbf{x}_{i-1} + \lambda \mathbf{e}_i)$ wyznaczając odpowiednie λ_i i przypisz $\mathbf{x}_i = \mathbf{x}_{i-1} + \lambda_i \mathbf{e}_i$.
- Wyznacz nowy kierunek sprzężony jako:
$$\mathbf{e} = \frac{\mathbf{x}_n - \mathbf{x}_0}{\|\mathbf{x}_n - \mathbf{x}_0\|}$$
- Wykonaj optymalizację kierunkową funkcji $f(\mathbf{x}_n + \lambda \mathbf{e})$ wyznaczając odpowiednie λ i przypisz $\mathbf{x} = \mathbf{x}_n + \lambda \mathbf{e}$.
- Znajdź:
$$m = \arg \max_{i=1,\ldots,n} \lambda_i$$
i jeśli spełniony jest odpowiedni warunek zamień wektor $\mathbf{e}_m$ na wektor $\mathbf{e}$.
- Jeśli nie spełnione zostało kryterium stopu wróć do punktu drugiego.

Warunek

$$\frac{\lambda_m \det \mathbf{E}}{\alpha} \geqslant 0.8$$

- Macierz $\mathbf{E}$ to macierz złożona z wektorów bazy.
- Wartość α to długość kroku po uwzględnieniu minimalizacji we wszystkich kierunkach bazy.

Uwagi

W literaturze można znaleźć inne warunki określające moment zmiany bazy kierunków.

Metoda Zangwilla

- Niech $\mathbf{x} \in \mathbb{R}^n$ oznacza początkowe rozwiązanie oraz niech $\mathbf{e}_i \in \mathbb{R}^n$ dla $i = 1, \ldots, n$ oznacza początkową bazę wzajemnie ortogonalnych wektorów w przestrzeni $\mathbb{R}^n$. Niech dodatkowo $\mathbf{f}_i \in \mathbb{R}^n$ dla $i = 1, \ldots, n$ oznacza początkową bazę wzajemnie ortogonalnych wektorów w przestrzeni $\mathbb{R}^n$, która nie będzie podlegać zmianom w trkacie działania algorytmu. Ustaw $j = 1$.
- Wykonaj optymalizację kierunkową funkcji $f(\mathbf{x} + \lambda \mathbf{f}_j)$ wyznaczając odpowiednie λ. Jeśli $\lambda = 0$ przypisz cyklicznie $j = j + 1$ i powtórz ten krok. Jeśli sprawdzone zostały wszystkie kierunki bez znalezienia lepszego rozwiązania zakończ algorytm.
- Wykonaj kroki od drugiego do szóstego metody Powella 1.
- Jeśli nie spełnione zostało kryterium stopu wróć do punktu drugiego.

Kryteria stopu

- Zadana liczba iteracji.
- Brak znaczącej zmiany w każdym z kierunków bazy.
- Brak znaczących postępów algorytmu w kolejnych przebiegach.

Uwagi

- Podczas optymalizacji kierunkowej warto wpierw metodą próbkowania wyznaczyć przedział unimodalności.
- Motywacją dla stworzenia metody Powella 2 i metody Zangwilla jest możliwość powstania kierunków liniowo zależnych w metodzie Powella 1 co prowadzi do braku zbieżności tej metody.