METODY OBLICZENIOWE OPTYMALIZACJI

6

Arkadiusz Tomczyk

Instytut Informatyki Politechniki Łódzkiej

14 kwietnia 2011

Funkcja celu

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

Metody gradientowe

- **Metoda najszybszego spadku.**
- **Metoda gradientu sprzężonego Fletchera-Reevesa.**
- **Metoda Newtona-Raphsona.**
- **Metoda quasi-Newtona.**
- **Metoda DFP.**
- **Metoda BFGS.**
- **Metoda Broydena.**
- **Metoda Pearsona.**

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:

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

Wspólny schemat obliczeń

- Niech $\mathbf{x} \in \mathbb{R}^n$ oznacza początkowe rozwiązanie.
- Przypisz $\mathbf{x}_0 = \mathbf{x}$, $i = 1$ oraz $j = 1$.
- **Wyznacz kierunek poszukiwań** $\mathbf{d}_i \in \mathbb{R}^n$.
- **Wyznacz długość kroku** λ_i **dla funkcji kierunkowej** $f(\mathbf{x}_{i-1} + \lambda\mathbf{d}_i)$ **i przypisz** $\mathbf{x}_i = \mathbf{x}_{i-1} + \lambda_i\mathbf{d}_i$.
- Przypisz $i = i + 1$ oraz jeśli $j = n$ przypisz $j = 1$, a w przeciwnym przypadku $j = j + 1$.
- Jeśli nie spełnione zostało kryterium stopu wróć do punktu trzeciego.

Metoda najszybszego spadku

- Wyznacz kierunek poszukiwań jako:

$$\mathbf{d}_i = -\nabla f(\mathbf{x}_{i-1})$$

Uwagi

- Można stosować stały krok co prowadzi do metody gradientu prostego.

Metoda gradientu sprzężonego Fletchera-Reevesa

- Jeśli $j = 1$ wyznacz kierunek poszukiwań jako:

$$\mathbf{d}_i = -\nabla f(\mathbf{x}_{i-1})$$

- W przeciwnym przypadku wyznacz kierunek poszukiwań jako:

$$\beta = \frac{\nabla f(\mathbf{x}_{i-1})^T \nabla f(\mathbf{x}_{i-1})}{\nabla f(\mathbf{x}_{i-2})^T \nabla f(\mathbf{x}_{i-2})}$$

$$\mathbf{d}_i = -\nabla f(\mathbf{x}_{i-1}) + \beta \mathbf{d}_{i-1}$$

Uwagi

- Kolejne wyznaczane kierunki są kierunkami sprzężonymi.

Metoda Newtona-Raphsona

- Jeśli $j = 1$ przyjmij jako $\mathbf{H}_i$ macierz jednostkową.
- W przeciwnym przypadku wyznacz macierz $\mathbf{H}_i$ jako:

$$\mathbf{H}_i = (Hf)^{-1}(\mathbf{x}_{i-1})$$

- Wyznacz kierunek poszukiwań jako:

$$\mathbf{d}_i = -\mathbf{H}_i \nabla f(\mathbf{x}_{i-1})$$

Uwagi

- Metoda ta zwana jest często po prostu metodą Newtona.
- Metoda może być wyprowadzona z rozwinięcia gradientu funkcji w szereg Taylora.
- Trudnością jest wyznaczanie hesjanu oraz jego odwrotności.

Metoda quasi-Newtona

- Metoda ta polega na przbliżaniu odwrotności hesjanu $(Hf)^{-1}$ w punkcie $\mathbf{x}_{i-1}$ przez macierz $\mathbf{H}_i$ w oparciu o kolejne wartości gradientu ∇f.
- Istnieją różne odmiany tej metody różniące się sposobami wyznaczania kolejnych wartości macierzy $\mathbf{H}_i$.
- Zazwyczaj zakłada się, że $\mathbf{H}_0$ jest macierzą jednostkową.
- Metoda ta zwana jest także metodą zmiennej metryki.

Oznaczenia

$$\alpha = \mathbf{x}_{i-1} - \mathbf{x}_{i-2}$$

$$\gamma = \nabla f(\mathbf{x}_{i-1}) - \nabla f(\mathbf{x}_{i-2})$$

Metoda DFP

$$\mathbf{H}_i = \mathbf{H}_{i-1} + \frac{\alpha\alpha^T}{\alpha^T\gamma} - \frac{\mathbf{H}_{i-1}\gamma\gamma^T\mathbf{H}_{i-1}}{\gamma^T\mathbf{H}_{i-1}\gamma}$$

Uwagi

- Nazwa metody pochodzi od nazwisk: Davidon, Fletcher, Powell.
- Kolejne wyznaczane kierunki są kierunkami sprzężonymi.
- W metodzie tej często wykorzystuje się kryterium Wolfe'a w celu wyznaczenia przybliżonej wartości długości kroku co gwarantuje dodatnią określoność przybliżanego hesjanu.

Metoda BFGS

$$\mathbf{H}_i = (\mathbf{I} - \frac{\gamma\alpha^T}{\gamma^T\alpha})^T\mathbf{H}_{i-1}(\mathbf{I} - \frac{\gamma\alpha^T}{\gamma^T\alpha}) + \frac{\alpha\alpha^T}{\gamma^T\alpha}$$

Uwagi

- Nazwa metody pochodzi od nazwisk: Broyden, Fletcher, Goldfarb, Shanno
- W metodzie tej często wykorzystuje się kryterium Wolfe'a w celu wyznaczenia przybliżonej wartości długości kroku co gwarantuje dodatnią określoność przybliżanego hesjanu.

Metoda Broydena

$$\mathbf{H}_i = \mathbf{H}_{i-1} + \frac{(\alpha - \mathbf{H}_{i-1}\gamma)\gamma^T \mathbf{H}_{i-1}}{\gamma^T \mathbf{H}_{i-1}\alpha}$$

Uwagi

- Istnieje również cała rodzina metod Broydena stanowiąca kombinację metod DFP i BFGS.

Metoda Pearsona 1

$$\mathbf{H}_i = \mathbf{H}_{i-1} - \frac{(\mathbf{H}_{i-1}\gamma)(\mathbf{H}_{i-1}\gamma)^T}{\gamma^T \mathbf{H}_{i-1}\gamma}$$

Metoda Pearsona 2

$$\mathbf{H}_i = \mathbf{H}_{i-1} + \frac{(\alpha - \mathbf{H}_{i-1}\gamma)\alpha^T}{\gamma^T \alpha}$$

Metoda Pearsona 3

$$\mathbf{H}_i = \mathbf{H}_{i-1} + \frac{(\alpha - \mathbf{H}_{i-1}\gamma)(\mathbf{H}_{i-1}^T\gamma)^T}{\gamma^T \mathbf{H}_{i-1}\gamma}$$

Kryteria stopu

- Zadana liczba iteracji.
- Odpowiednio mała długość wektora gradientu.
- Brak znaczących postępów algorytmu w kolejnych przebiegach.

Wyznaczanie długości kroku

- Wartość stała.
- Poszukiwanie dokładne (optymalizacja kierunkowa).
- Poszukiwanie niedokładne (kryteria Armijo, Wolfe'a, Goldsteina).

Uzupełnienia

Metoda Levenberga-Marquardta łączy cechy metody Newtona-Raphsona oraz metody najszybszego spadku (a właściwie metody gradientu prostego) w celu zwiększenia szybkości zbieżności.