METODY OBLICZENIOWE OPTYMALIZACJI

3

Arkadiusz Tomczyk

Instytut Informatyki Politechniki Łódzkiej

30 marca 2011

Funkcja kwadratowa

$$w(x) = \gamma x^2 + \beta x + \alpha$$

$$w'(x^*) = 0 \Rightarrow x^* = -\frac{\beta}{2\gamma}$$

$$w''(x^*) > 0 \Rightarrow \gamma > 0 \Rightarrow \text{ minimum}$$

Uwagi

- Interpolacja w oparciu o wartości funkcji f w trzech punktach.
- Metoda interpolacji kwadratowej jest metodą zerowego rzędu.
- Można wybierać inne strategie przyjmowania wartości początkowej x w kolejnych iteracjach niż te opisane poniżej.

Interpolacja kwadratowa Powella

- Ustal wartość początkową $x = 0$ oraz długość kroku $\Delta x > 0$.
- Wyznacz $f_0 = f(x)$ oraz $f_1 = f(x + \Delta x)$.
- Jeśli $f_1 < f_0$ wyznacz $f_2 = f(x + 2\Delta x)$ oraz oblicz:

$$\gamma = \frac{f_2 + f_0 - 2f_1}{2\Delta x^2}, \ \beta = \frac{4f_1 - 3f_0 - f_2}{2\Delta x}, \ \alpha = f_0$$

- W przeciwnym wypadku wyznacz $f_2 = f(x - \Delta x)$ oraz oblicz:

$$\gamma = \frac{f_1 - 2f_0 + f_2}{2\Delta x^2}, \ \beta = \frac{f_1 - f_2}{2\Delta x}, \ \alpha = f_0$$

- Jeśli x^* nie jest miniumum lub jest bardzo oddalone od x za x^* przyjmij punkt oddalony od x o pewną założoną maksymalną długość kroku w kierunku malejących wartości funkcji f.
- Jeśli nie jest spełniony warunek stopu przyjmij $x = x^*$ i powtórz powyższe kroki zaczynając od punktu drugiego.

Wielomian trzeciego stopnia

$$w(x) = \delta x^3 + \gamma x^2 + \beta x + \alpha$$

$$w'(x^*) = 0$$

$$w''(x^*) > 0 \Rightarrow \text{ minimum}$$

Uwagi

- Interpolacja w oparciu o wartości funkcji f i pochodnej f' w dwóch punktach.
- Metoda interpolacji sześciennej jest metodą pierwszego rzędu.

Interpolacja sześcienna Davidona

- Ustal wartość początkową $x = 0$ oraz długość kroku $\Delta x > 0$.
- Wyznacz $f_0 = f(x)$, $f_0' = f'(x)$, $f_1 = f(x + \Delta x)$ oraz $f_1' = f'(x + \Delta x)$.
- Jeśli $f_1' > 0$ lub $f_1 > f_0$ oblicz:

$$x^* = \beta \left(\frac{f_0' + \eta + \theta}{f_0' + f_1' + 2\eta} \right)$$

gdzie:

$$\eta = \frac{3}{\beta}(f_0 - f_1) + f_0' + f_1', \;\; \theta = \sqrt{\eta^2 - 2 f_0' f_1'}$$

- W przeciwnym przypadku przyjmij $\Delta x = 2\Delta x$ i wróc do punktu drugiego.
- Jeśli $f'(x^*) \geqslant 0$ przyjmij $\Delta x = x^* - x$.
- W przeciwnym przypadku przyjmij $\Delta x = x + \Delta x - x^*$ oraz $x = x^*$.
- Jeśli nie jest spełniony warunek stopu powtórz powyższe kroki zaczynając od punktu drugiego.

Uwagi

Zawsze $f_0' < 0$.

Optymalizacja kierunkowa

Funkcja jednowymiarowa argumentu $\lambda \in \mathbb{R}$ wyznaczana dla punktu $\mathbf{x} \in \mathbb{R}^n$ w kierunku $\mathbf{d} \in \mathbb{R}^n$:

$$\phi(\lambda) = f(\mathbf{x} + \lambda \mathbf{d})$$

Kierunek spadku wartości

$$\langle \nabla f(\mathbf{x}), \mathbf{d} \rangle < 0 \Rightarrow \phi'(0) < 0$$

Uwagi

- Wybór kroku w danym kierunku nie musi oznaczać znajdowania dokładnej wartości minium w danym kierunku.
- Wybierany krok nie może być ani za długi ani za krótki.
- Sposób wybierania długości kroków powinien pozwalać na zmniejszanie wartości funkcji celu.
- Kryteria wyboru długości powinny łatwo sprawdzalne.

Kryterium Armijo

$$f(\mathbf{x} + \lambda \mathbf{d}) \leqslant f(\mathbf{x}) + c_1 \lambda \langle \nabla f(\mathbf{x}), \mathbf{d} \rangle$$

$$c_1 \in (0, 1)$$

Backtracking

- Wybierze początkową wartość kroku λ oraz $\tau \in (0, 1)$.
- Tak długo jak nie jest spełnione kryterium Armijo wyznacz $\lambda = \tau\lambda$.

Uwagi

- Bez uwzględnienia c_1 może nie istnieć wartość λ spełniająca powyższy warunek.
- Powyższe kryterium zwane jest też kryterium dostatecznego spadku.

Kryteria krzywizny

- Słabe kryterium krzywizny:

$$\langle \nabla f(\mathbf{x} + \lambda \mathbf{d}), \mathbf{d} \rangle \geqslant c_2 \langle \nabla f(\mathbf{x}), \mathbf{d} \rangle$$

$$c_2 \geqslant c_1, c_2 \in (0, 1)$$

- Silne kryterium krzywizny:

$$|\langle \nabla f(\mathbf{x} + \lambda \mathbf{d}), \mathbf{d} \rangle| \leqslant c_2 |\langle \nabla f(\mathbf{x}), \mathbf{d} \rangle|$$

$$c_2 \geqslant c_1, c_2 \in (0, 1)$$

Kryteria Wolfe'a

- Słabe kryterium Wolfe'a to kryterium Armijo oraz słabe kryterium krzywizny.
- Silne kryterium Wolfe'a to kryterium Armijo oraz silne kryterium krzywizny.

Uwagi

Bez uwzględnienia $c_2 \geqslant c_1$ może nie istnieć wartość λ spełniająca powyższe warunki.

Kryterium Goldsteina

$$f(\mathbf{x}) + (1 - c_3)\lambda \langle \nabla f(\mathbf{x}), \mathbf{d} \rangle \leqslant f(\mathbf{x} + \lambda \mathbf{d}) \leqslant f(\mathbf{x}) + c_3 \lambda \langle \nabla f(\mathbf{x}), \mathbf{d} \rangle$$

$$c_3 \in (0, 0.5)$$

Uwagi

Zarówno w przypadku kryteriów Wolfe'a oraz kryterium Goldsteina można przyjąć podobną strategię wyznaczania wartości λ. W obu przypadkach wystarczy znaleźć wartość λ, dla której nie jest spełnione górne kryterium, a następnie zastosować metodę bisekcji w celu znaleznia wartości spełniającej oba warunki.

Uzupełnienia

Dla konkretnych metod optymalizacji wielowymiarowej gdzie dokonywany jest wybór kierunku $\mathbf{d}$ przy dodatkowych założeniach dotyczących funkcji celu można wykazać, że wybór niektórych z powyższych kryteriów gwarantuje globalną zbieżność.

Uzupełnienia

$$\phi'(\lambda) = \frac{\langle \nabla f(\mathbf{x} + \lambda \mathbf{d}), \mathbf{d} \rangle}{\|\mathbf{d}\|}$$