ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 1983

Seria : AUTOMATYKA z. 68 Nr kol. 767

Andrzej DĄBROWSKI
Instytut Meteorologii i Gospodarki Wodnej
Oddział Wrocław

NIEGRADIENTOWA METODA OPTYMALIZACJI FUNKCJI CZASU

Streszczenie. Omówiono metodę optymalizacji na podstawie badania różnic uporządkowanego zbioru realizacji funkcji spełniających warunek Lipschitza. Dla metody opracowano i zbadano programy obliczeniowe.

Zagadnienia związane z optymalizacją funkcji występują w teorii sterowania głównie w związku z dwoma problemami : wyborem optymalnej strategii oraz z dopasowaniem nieznanych parametrów, opisujących system. W wielu sytuacjach, szczególnie w sterowaniu z adaptacją parametrów [1] , należy oba te zagadnienia rozwiązywać łącznie.

Klasyczne, gradientowe metody optymalizacji nie mają częstokroć zastosowania. Wynika to bądź ze zbyt skomplikowanej postaci funkcji celu, bądź wręcz z nieistnienia pochodnych funkcji.

Przedstawię tu metodę pochodzącą od Strongina [2] , dla której zestawy programów opracowano w IMGW we Wrocławiu.

Niech $f(x)$ będzie funkcją, określoną na N-wymiarowej kostce $D = \{x \in R^N : a_i \leqslant x_i \leqslant b_i , i = 1,2, \dots , N\}$. Należy zaprojektować na podstawie par obserwacji (x^i , z^i), gdzie $i = 1,2, \dots , k$, $z^i = f(x^i) + e^i$ oraz e^i jest błędem i-tej obserwacji, następną obserwację w punkcie x^{k+1} tak, by ciąg $\{x^n\}$ był zbieżny do lokalnego minimum funkcji f .

Reguła wyboru x^{k+1} zależna jest na ogół od własności funkcji f i metody gradientowe bazują na znajomości pochodnych tej funkcji. Inne, niegradientowe metody, na przykład metoda Fibonacciego, bazuje na założeniu jednomodalności funkcji.

Zaletą opisywanej metody jest duża ogólność klasy funkcji celu f , dla których można dowieść zbieżność ciągu $\{x^n\}$ do minimum. Jest to prawdą dla funkcji spełniających warunek Lipschitza.

Opis metody

Metoda Strongina bazuje na teorii decyzji statystycznych. Zakłada się, że zbiór par $\{(x^i , z^i): i = 1,2, \dots , k\}$ stanowi próbkę trajektorii pewnego procesu stochastycznego w punktach x^i. Punkt x^{k+1} znajdowany jest z warunku, że prawdopodobieństwo, że w punkcie x^{k+1} trajektoria ma lokalne minimum, przy warunku (x^i, z^i) $i = 1,2, \dots , k$, osiąga maksimum.

Algorytm dla $N = 1$.

1°. Uporządkować punkty $x^0, x^1, \dots, x^k$, gdzie $x^0 = a$, $x^1 = b$, rosnąco. Uporządkowane liczby oznaczymy indeksami u dołu :

$$a = x_0 < x_1 < \dots < x_k = b .$$

Niech $z_i = f(x_i)$.

2°. Znaleźć :

$$M = \max_{1 \leqslant i \leqslant k} \frac{|z_i - z_{i-1}|}{x_i - x_{i-1}} .$$

3°. Dla zadanego z góry (parametr metody) $r > 1$ obliczyć :

$$m = \begin{cases} 1 & \text{gdy } M=0 \\ rM & \text{gdy } M>0 \end{cases}$$

4°. Dla $i = 1,2, \dots , k$ obliczyć :

$$R(i) = m(x_i - x_{i-1}) + \frac{(z_i - z_{i-1})^2}{m(x_i - x_{i-1})} - 2(z_i + z_{i-1}).$$

5°. Znaleźć wskaźnik t, spełniający warunek :

$$R(t) = \max_{1 \leqslant i \leqslant k} R(i).$$

6°. $$x^{k+1} = \frac{x_t + x_{t-1}}{2} + \frac{z_t - z_{t-1}}{2m} .$$

Operacje te powtarzamy dopóty, dopóki $x_t - x_{t-1} < \varepsilon$, gdzie ε jest zadaną dokładnością. Wtedy minimum osiągnięte jest w punkcie

$x_{\bar{x}} = \arg\min_{1 \leqslant i \leqslant k} f(x_i)$ i wynosi ono $f_{\bar{x}} = \min_{1 \leqslant i \leqslant k} z_i$

Prawdziwe jest następujące twierdzenie :

Niech punkt $\bar{x}$ będzie punktem skupienia ciągu $\{x^k\}$, generowanego za pomocą wzorów 1° - 6° dla funkcji f, spełniającej warunek Lipschitza ze stałą K, na odcinku $[a,b]$. Wtedy :

1/ jeśli f ma w $[a,b]$ skończoną liczbę lokalnych ekstremów, to $\bar{x}$ jest jednym z nich,

2/ jeśli ciąg $\{x^k\}$ ma inny punkt skupienia $\hat{x}$ to $f(\hat{x}) = f(\bar{x})$,

3/ dla każdego k $f(x^k) \geqslant f(\bar{x})$,

4/ jeżeli dla pewnego k liczba m wyznaczona w punkcie 3° spełnia nierówność $m > 2K$, to zbiór punktów skupienia ciągu $\{x^k\}$ pokrywa się ze zbiorem punktów, w których f osiąga absolutne minimum w $[a,b]$.

Algorytm dla $N>1$.

Algorytm dla $N=1$ można wzbogacić o procedurę, generującą krzywą Peano, to znaczy krzywą ciągłą, odwzorowującą odcinek $[-0,5\ ,\ 0.5]$. na kostkę $D=\{x\in R^N : a_i\leq x_i\leq b_i \quad i=1,2,\ldots,N\}$. Jeżeli odwzorowanie to oznaczymy przez d, to mamy przekształcenie $d: [-0,5,0,5]\to D$ które sprowadza przypadek $N>1$ do sytuacji $N=1$. Mamy bowiem:

$$f_x = \min_{d\in D} f(d) = \min_{x\in[-0,5,0.5]} f(d(x)).$$

Przeprowadzono eksperymenty porównujące szybkość znajdowania ekstremum dla różnych metod niegradientowych. Opisana metoda szybciej niż inne lokalizowała punkt ekstremum przy tych samych warunkach, określających dokładność.

Literatura

[1] G.N.SARIDIS - Self-organizing control of stochastic systems, New York and Basel 1977.

[2] R.G. STRONGIN - Cislennyje metody v mnogoekstremalnych zadacach. Moskva 1978.

NONGRADIENT OPTIMIZATION TECHNIQUE FOR TIME FUNCTIONS

Summary. An optimization technique based on examination of differences of ordered set of functions realizations satisfying Lipschitz condition is described. Computer programs has been developed and tested to verify the method.

БЕЗГРАДИЕНТНЫЙ МЕТОД ОПТИМИЗАЦИИ ВРЕМЕННОЙ ФУНКЦИИ

/ Резюме /

В работе оговорен метод оптимизации на основе исследований упорядоченных разниц множества реализации функций, выполняющих условие Липшица. Для метода разработано и исследовано соответствующие программы.