

Zbieżność globalna stochastycznych algorytmów optymalizacji

13 maj 2011

Algorytmy stochastyczne stały się powszechnie używanym narzędziem optymalizacji funkcji. Większość z nich jest objaśniana i używana heurystycznie. Celem referatu jest prezentacja twierdzenia, oraz jego dowodu i zastosowań, traktującego o globalnej zbieżności stochastycznych algorytmów.

Streszczenie:

1 Algorytm stochastyczny:

Założenia

- (Ω, Σ, P) – przestrzeń probabilistyczna,
- (A, d) – przestrzeń metryczna zwarta, (B, d_B) – przestrzeń metryczna ośrodkowa,
- $X_0: \Omega \rightarrow A$ – mierzalne
- odwzorowania $Y_t: \Omega \rightarrow B, t = 0, 1, \dots$ są mierzalne,
- $X_0, Y_0, Y_1, \dots$ są niezależne,
- $T_t: A \times B \rightarrow A, t = 0, 1, \dots$ – mierzalne

Definicja algorytmu

$$X_{t+1} := T_t(X_t, Y_t)$$

2 Twierdzenie

2.1 Oznaczenia:

- (M, d_M) – przestrzeń miar probabilistycznych na A z metryką Prohorova d_M
- (N, d_N) – przestrzeń miar probabilistycznych na (B, d_B) z metryką Prohorova
- (Γ, d_Γ) – przestrzeń odwzorowań mierzalnych $T: A \times B \rightarrow A$ z metryką supremum d_T
- $A \times B \supset D_T$ – zbiór punktów nieciągłości odwzorowania T

2.2 Twierdzenie

Założenia:

Niech $U \subset \Gamma \times N$ będzie zbiorem punktów $(T, \nu) \in \Gamma \times N$ spełniających

(A) Dla każdego $x \in A^*$ oraz $y \in B$, $T(x, y) \in A^*$

(B) Dla każdego $x \in A$

$$\int_B f(T(x, y)) \nu(dy) \leq f(x). \quad (1)$$

Niech $U_0 \subset U$ będzie zwarty oraz taki, że dla każdego $(T, \nu) \in U_0$

(C) Dla każdego $x \in A \setminus A^*$

$$\int_B f(T(x, y)) \nu(dy) < f(x). \quad (2)$$

(D) Dla każdego $m \in M$ $(m \times \nu)(W_T) = 0$.

Niech $\Lambda \subset U^\mathbb{N}$ spełnia

$$\Lambda := \{(u_0, u_1, u_2, \dots) \in U^\mathbb{N} : \forall t \exists t_1 > t \ u_{t_1} \in U_0\}$$

oraz niech $u = \{(\Gamma_i, \nu_i)\}_{i=0}^\infty \in \Lambda$.

Teza:

$$\forall \epsilon > 0 \ P(d(X_t, A^*) < \epsilon) \rightarrow 1, \text{ gdzie } X_{t+1} = T_t(X_t, Y_t).$$