

Библиографические ссылки

1. *Аргучинцев А. В., Дыхта В. А., Срочко В. А.* Оптимальное управление: нелокальные условия, вычислительные методы и вариационный принцип максимума // Известия вузов. Математика. 2009. № 1. С. 3–43.
2. *Срочко В. А.* Итерационные методы решения задач оптимального управления. М.: Физматлит, 2000.
3. *Антипина Н. В., Дыхта В. А.* Линейные функции Ляпунова-Кротова и достаточные условия оптимальности в форме принципа максимума // Известия вузов. Математика. 2002. № 12. С. 11–22.
4. *Никольский М. С.* О достаточности принципа максимума Понтрягина в некоторых оптимизационных задачах // Вестник Моск. ун-та. Сер. 15. Вычислит. матем. и киберн. 2005. № 1. С. 35–43.
5. *Аксенюшкина Е. В., Срочко В. А.* Достаточные условия оптимальности для одного класса невыпуклых задач управления // Журн. вычисл. матем. и матем. физ. 2015. Т. 55. № 10. С. 1670–1680.
6. *Антоник В. Г., Срочко В. А.* Условия оптимальности типа принципа максимума в билинейных задачах управления // Журн. вычисл. матем. и матем. физ. 2016. Т. 56. № 12. С. 2054–2064.
7. *Srochko V., Antonik V., Aksenyushkina E.* Sufficient optimality conditions for extremal controls based on functional increments // Numerical Algebra, Control and Optimization. 2017. Vol. 7. № 2. С. 191–199.

ОБ УСКОРЕНИИ СУБГРАДИЕНТНОГО МЕТОДА С ШАГОМ ПОЛЯКА

П.И. Стецюк

Институт кибернетики им. В.М. Глушкова НАН Украины
пр. Глушкова, 40, 03187 Киев, Украина
stetsyukp@gmail.com

Пусть $f(x)$ – выпуклая функция, $x \in \mathbb{R}^n$. Ее субградиент $\partial f(x)$ удовлетворяет условию:

$$(x - x^*, \partial f(x)) \geq f(x) - f^*, \quad \forall x \in \mathbb{R}^n. \quad (1)$$

Здесь (x, y) – скалярное произведение векторов $x \in \mathbb{R}^n$ и $y \in \mathbb{R}^n$; x^* – точка минимума функции $f(x)$; f^* – минимальное значение функции $f(x)$: $f^* = f(x^*)$.

Пусть известно f^* . Для нахождения точки $x^* \in \mathbb{R}^n$ можно использовать субградиентный метод Поляка [1]:

$$x_{k+1} = x_k - h_k \frac{\partial f(x_k)}{\|\partial f(x_k)\|}, \quad h_k = \frac{f(x_k) - f^*}{\|\partial f(x_k)\|}, \quad k = 0, 1, 2, \dots \quad (2)$$

Для выпуклой функции $f(x)$ шаг h_k задает величину максимального сдвига в направлении нормированного антисубградиента, при котором условие (1) гарантирует, что угол между антисубградиентом и направлением из точки x_{k+1} на точку минимума будет тупым.

Теорема 1. *[[2], стр. 300]. Последовательность $\{x_k\}_{k=0}^{\infty}$, генерируемая методом (1), удовлетворяет неравенствам*

$$\|x_{k+1} - x^*\|^2 \leq \|x_k - x^*\|^2 - \frac{(f(x_k) - f^*)^2}{\|\partial f(x_k)\|^2}, \quad k = 0, 1, 2, \dots$$

Недостатком метода (2) является его медленная сходимость для овражных функций. Так, например, для негладкой функции двух переменных $f(x_1, x_2) = |x_1| + 10|x_2|$ скорость сходимости метода (2) определяется геометрической прогрессией со знаменателем $\sqrt{1 - 1/10^2}$.

Ускорить метод (2) можно, если пространство переменных преобразовать так, чтобы тупой угол между двумя последовательными субградиентами уменьшался. Субградиентный метод с шагом Поляка в преобразованном пространстве, где тупой угол между двумя субградиентами преобразуется в прямой, имеет следующий вид [2]:

$$x_{k+1} = x_k - h_k B_k \frac{B_k^T \partial f(x_k)}{\|B_k^T \partial f(x_k)\|}, \quad h_k = \frac{f(x_k) - f^*}{\|B_k^T \partial f(x_k)\|}, \quad k = 0, 1, 2, \dots, \quad (3)$$

где матрица $B_0 = I$, а матрица B_{k+1} пересчитывается по правилу:

$$B_{k+1} = \begin{cases} B_k + (B_k \eta) \xi_{k+1}^T, & \text{если } \mu_k < 0, \\ B_k, & \text{иначе,} \end{cases}$$

где

$$\mu_k = (\xi_k, \xi_{k+1}), \quad \xi_k = \frac{B_k^T g_f(x_k)}{\|B_k^T g_f(x_k)\|}, \quad \xi_{k+1} = \frac{B_k^T g_f(x_{k+1})}{\|B_k^T g_f(x_{k+1})\|},$$

$$\eta = \left(\frac{1}{\sqrt{1 - \mu_k^2}} - 1 \right) \xi_{k+1} - \frac{\mu_k}{\sqrt{1 - \mu_k^2}} \xi_k.$$

Теорема 2. *[[2], стр. 203]. Пусть $A_k = B_k^{-1}$, $A_{k+1} = B_{k+1}^{-1}$. Последовательность $\{x_k\}_{k=0}^{\infty}$, генерируемая методом (3), удовлетворяет неравенствам*

$$\|A_{k+1}(x_{k+1} - x^*)\|^2 \leq \|A_k(x_k - x^*)\|^2 - \frac{(f(x_k) - f^*)^2}{\|B_k^T \partial f(x_k)\|^2}, \quad k = 0, 1, \dots \quad (4)$$

Неравенства (4) означают, что в методе (3) преобразование пространства таково, что в каждом очередном преобразованном пространстве переменных гарантируется уменьшение расстояния до множества точек минимума. Антиовражная техника в методе (3) направлена на уменьшение степени овражности поверхностей уровня выпуклых функций подобно тому, как это сделано в r -алгоритмах Н.З. Шора [3]. Для овражных функций это обеспечивает ускоренную сходимость метода (3) по отношению к методу (2) при произвольной начальной стартовой точке x_0 . Так, например, для овражной функции двух переменных $f(x_1, x_2) = |x_1| + t|x_2|$ метод (3) будет находить точку x^* не более чем за три итерации независимо от значения $t > 0$.

Библиографические ссылки

1. Поляк Б.Т. Введение в оптимизацию. М.: Наука, 1983.
2. Стецюк П.И. Методы эллипсоидов и r -алгоритмы. Кишинэу: Эврика, 2014.
3. Шор Н.З. Методы минимизации недифференцируемых функций и их приложения. Киев: Наукова думка, 1979.

ВЕКТОРНОЕ ЭНТРОПИЙНОЕ УПРАВЛЕНИЕ ГАУССОВСКИМИ СТОХАСТИЧЕСКИМИ СИСТЕМАМИ

А.Н. Тырсин, Г.Г. Геворгян

Уральский федеральный университет, Мира 19, 620002 Екатеринбург, Россия
a.n.tyrsin@urfu.ru, garnik.ggg@gmail.com

В докладе на примере гауссовских систем рассмотрен подход к управлению многомерными стохастическими системами, основанный на использовании векторной формы дифференциальной энтропии.

Пусть имеем гауссовскую стохастическую систему S , описываемую в виде нормально распределенного случайного вектора $\mathbf{Y} = (Y_1, Y_2, \dots, Y_m)$, имеющего ковариационную матрицу $\Sigma = \{\sigma_{ij}\}_{m \times m}$, $\sigma_{ij} = \sigma_i^2$. В [1] предложено векторное представление дифференциальной энтропии $H(\mathbf{Y})$ как

$$\mathbf{h}(\mathbf{Y}) = (h_V; h_R) = (H(\mathbf{Y})_V; H(\mathbf{Y})_R), \quad (1)$$

где $H(\mathbf{Y})_V$ – энтропия хаотичности, равная сумме энтропий случайных величин Y_i ; $H(\mathbf{Y})_R$ – энтропия самоорганизации, характеризующая вклад в энтропию тесноты корреляционных связей между ком-