Семинар 11. Методы оптимизации. МФТИ. Осень 2017. Тренин С.А.

Субградиентный метод.

Пусть $f: \mathbb{R}^n \to \mathbb{R}$ - выпуклая функция, вектор $c \in \mathbb{R}^n$ и для него выполнено неравенство: $f(x+y) \geq f(x) + c^T y$ для любого $y \in \mathbb{R}^n$, тогда c – называется субградиентом функции $f(x)$ в точке x и обозначается $\partial f(x)$. (Такой вектор в точке x может быть не единственным, в зависимости от контекста субградиентом в точке можно называть как все множество векторов, для которых выполнено характеристическое неравенство, так и один конкретный вектор из этого множества).

Истины следующие утверждения (правила субдифференциального исчисления):

Правило 1. Пусть $f_1(x)$ и $f_2(x)$ – выпуклые функции, а $\partial f_1(x)$ и $\partial f_2(x)$ – соответствующие множества субградиентов. Тогда, для $f(x) = \alpha f_1(x) + \beta f_2(x)$, $\partial f(x) = \alpha \partial f(x) + \beta \partial g(x)$. (Знак + для множеств следует понимать в следующем смысле: если A, B и C – множества в $\mathbb{R}^n$, то $C = \alpha A + \beta B$ означает, что $C = \{c = \alpha a + \beta b, a \in A, b \in B\}$.)

Правило 2. Пусть $f(x) = \max\limits_{0 \leq i \leq m} f_i(x)$, где $f_i(x)$ – выпуклые функции. Тогда

$$\partial f(x) = Conv \bigcup_{i \in I(x)} \partial f_i(x), \qquad I(x) = \{i: f_i(x) = f(x)\}$$

Правило 3. Пусть A – матрица $m \times n$, $\varphi(x)$ – выпуклая функция на $\mathbb{R}^m$, $f(x) = \varphi(Ax)$ $x \in \mathbb{R}^n$. Тогда $\partial f(x) = A^T \partial \varphi(Ax)$.

Докажите следующие утверждения:

1. Если $f(x)$ - дифференцируема, в точке x, то субградиент определен однозначно и совпадает с градиентом $\partial f(x) = \nabla f(x)$.
2. Выпуклая функция $f(x)$ имеет в произвольной точке x одностороннюю производную по любому направлению, равномерно ограниченную по нарпавлениям.
3. Необходимым и достаточным условием минимума выпуклой функции $f(x)$ в точке x^* является то, что $0 \in \partial f(x)$.

Ответьте на следующие вопросы:

4. Всегда ли множество субградиентов в каждой точке выпуклой функции является a) непустым? b) выпуклым? c) ограниченым?

Точкой острого экстремума выпуклой функции называется такая точка x^*, что для всех x $f(x) \geq f(x^*) + \lambda \|x\|_2, \lambda > 0$.

5. Докажите, что, если для выпуклой функции $f(x)$, x^*- точка острого экстремума, и дана некоторая выпуклая функция $g(x)$, то найдется такое $\varepsilon_0 > 0$, что для любого $\varepsilon \in [0, \varepsilon_0)$ точка минимума функции $f(x) + \varepsilon g(x)$ единственна и совпадает с x^*.

6. Определите значения субградиента по x для всех точек области определения функции:
   a. $|x|,\ x \in \mathbb{R}$
   b. $\|x\|_2,\ x \in \mathbb{R}^n$
   c. $\|Ax - b\|_1,\ x \in \mathbb{R}^n, b \in \mathbb{R}^m$
   d. $\|Ax - b\|_\infty,\ x \in \mathbb{R}^n, b \in \mathbb{R}^m$
   e. $0.5\|x - b\|^2 + \lambda\|x\|_1, x \in \mathbb{R}^n, b \in \mathbb{R}^n$

Субградиентным методом решения задачи безусловной минимизации $f(x) \to min$, называется итеративная процедура построения последовательности $\{x_k\}$, в которой x_{k+1} выбирается по правилу $x_{k+1} = x_k - t_k \frac{\partial f(x_k)}{\|\partial f(x_k)\|}$, где $\partial f(x_k)$ – один из субградиентов функции в точке x_k, а $t_k > 0$. Величина t_k (*размер шага*) может выбираться по различным правилам, например так, чтобы $t_k \to 0$, но $\sum_{k=0}^{\infty} t_k = \infty$: $t_k = \frac{t_0}{k+c}$, $t_k = \frac{t_0}{k^q}, 0 < q \leq 1$, $t_k = \frac{t_0}{k \ln k}$.

7. Пусть решается задача оптимизации функции $f(x) = |x^1 - x^2| + 0.2|x^1 + x^2|$. Определите значение всех возможных субградиентов этой функции в точке (1,1). Приведите пример, для которого направление направление субградиетного метода не является направлением убывания.
8. Пусть $\varphi_k = \min\limits_{0 \leq i \leq k} f(x_i)$ – минимальное значение выпуклой функции f среди всех значений на траектории субградиентного метода. Докажите, что $\varphi_k \to f^* = \inf\limits_{x \in \mathbb{R}^n} f(x)$.