Методы оптимизации. Задание 2

Сергей Володин, 374 гр.

задано 2016.03.29

Задача 1

Доказать: Пусть f — β-гладкая. Тогда $\forall x, y \hookrightarrow f(x) \leqslant f(y) + \nabla^T f(y)(x-y) + \frac{\beta}{2}||x-y||^2$.

1. Обозначим $\mu(t) = f(y+t(x-y))\colon [0,1] \to \mathbb{R}$. Поскольку f дифференцируема, μ также дифференцируема как композиция дифференцируемых функций. Тогда $\mu(1) = \mu(0) + \int\limits_0^1 \mu'(t)dt$. Подставим определение μ, получим формулу Ньютона-Лейбница для f на отрезке $[y, x] \in \mathbb{R}^n$: $f(x) = f(y) + \int\limits_0^1 dt \nabla^T f(y+t(x-y))(x-y)$.
2. Рассмотрим величину $\alpha \overset{\text{def}}{=} f(x) - f(y) - \nabla^T f(y)(x-y)$ и докажем, что $\alpha \leqslant 0$:
3. $\alpha = \int\limits_0^1 dt \nabla^T f(y+t(x-y))(x-y) - \nabla^T f(y)$. Внесем второе слагаемое под интеграл, получим
$$\alpha = \int\limits_0^1 dt \left(\nabla^T f(y+t(x-y)) - \nabla^T f(y)\right)(x-y)$$
4. $|\alpha| \leqslant \int\limits_0^1 dt |\underbrace{\left(\nabla^T f(y+t(x-y)) - \nabla^T f(y)\right)}_{A}(x-y)|$.
5. A— оператор, действующий на $x-y$. Поскольку f — β-гладкая, т.е.
$$\forall x, y \hookrightarrow ||\nabla f(x) - \nabla f(y)||_* \leqslant \beta||x-y||,$$
Получаем $||A||_* = \sup\limits_{x \in \mathbb{R}} \frac{|Ax|}{||x||} \leqslant \beta||y - t(x-y) - y|| = \beta t||x-y||$, откуда $|A(x-y)| \leqslant ||A||_*||x-y|| \leqslant \beta t||x-y||^2$
6. Получаем $|\alpha| \leqslant \int\limits_0^1 \beta t||x-y||^2 dt = \frac{\beta}{2}||x-y||^2$ ■

Задача 2

Определим $\delta_t \overset{\text{def}}{=} f(x_t) - f(x^*)$, где x_t — t-я точка в алгоритме Frank-Wolfe. Получена оценка $\delta_{t+1} \leqslant \frac{\beta R^2}{2}(\prod\limits_{k=1}^t (1-\gamma_k) + \sum\limits_{k-1}^n \gamma_{t-k}^2 \prod\limits_{j=t-k}^t (1-\gamma_j)$. Оценить выражение как функцию γ_t и выбрать γ_t как минимум этой функции.

Задача 3

$E = (\mathbb{R}^n, ||\cdot||)$. Определим $f^*(p) \overset{\text{def}}{=} \sup\limits_{x \in \mathbb{R}^n} (p^T x - f(x))$, $p \in E^*$. Найти субдифференциал $\partial f^*(p)$.

Задача 4

Задача про метод проекции субградиента из задания 1 (какая скорость сходимости, если делать проекцию каждые k шагов)