

ЗАДАНИЕ

по курсу “**Методы глобальной оптимизации в задачах теории управления и машинном обучении**”

Реализовать решение задачи аппроксимации

$$Q(\theta) = \frac{1}{l} \sum_{i=1}^l \underbrace{[f(\theta, x_i) - y_i]^2}_{L(\theta, x_i, y_i)} = \frac{1}{l} \sum_{i=1}^l L(\theta, x_i, y_i) \rightarrow \min_{\theta \in R^n},$$

где

$\theta \in R^n$ – вектор подбираемых параметров;

$f(\theta, x) = \theta_0 + \theta_1 x + \theta_2 x^2$,

$(x_i, y_i)_{i=1}^l$ – обучающая выборка.

Положить $l = 10$.

Язык – произвольный.

Использовать метод, определяемый номером по списку.

Сравнить с методом под номером 15.

1. Stochastic Gradient Descent (SGD) /метод стохастического градиентного спуска/.
2. Mini-batch Gradient Descent /минипакетный метод градиентного спуска/.
3. Stochastic Average Gradient (SAG) /Метод среднего стохастического градиента/.
4. Адаптивный быстрый градиентный метод (Нестеров).
5. Nesterov Accelerated Gradient (NAG) /Ускоренный градиентный метод Нестерова/;
6. Метод адаптивного градиента (Adaptive Gradient, AdaGrad),
7. Метод скользящего среднего (Root Mean Square Propagation, RMSProp),
8. Метод адаптивной оценки моментов (Adaptive Moment Estimation, Adam).
9. Модификация метода Adam (Adamax).
10. Ускоренный по Нестерову метод адаптивной оценки моментов (Nesterov–accelerated Adaptive Moment Estimation, Nadam).
11. Метод скользящего среднего с адаптацией шага (AdaDelta).
12. Метод Ньютона-Рафсона Метод Ньютона-Гаусса
13. Метод Давидона-Флетчера-Пауэлла
14. Метод Бройдена-Флетчера-Гольдфарба-Шенно
15. Batch Gradient Descent (Vanilla Gradient Descent) /пакетный метод градиентного спуска/.

Программа реализует 2 метода:

- ввод обучающей выборки
- получение параметров
- вычисление критерия

Подсчет значений целевой функции для тестовой выборки.

График

- аппроксимирующей функции
- точки обучающей выборки.

Отчет содержит

- титульный лист
- краткое описание методов
- аналитическое нахождение градиентов
- код
- результаты, анализ влияния параметров
- выводы