

Гробов А.В.
Кафедра физики
элементарных частиц
МИФИ

Машинное обучение

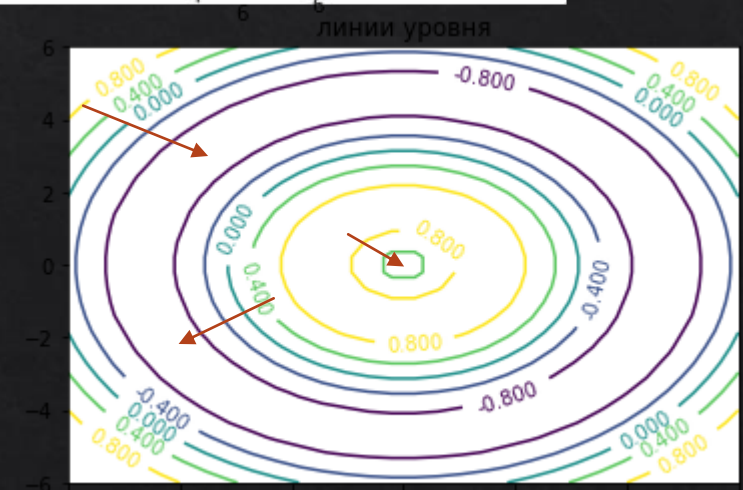
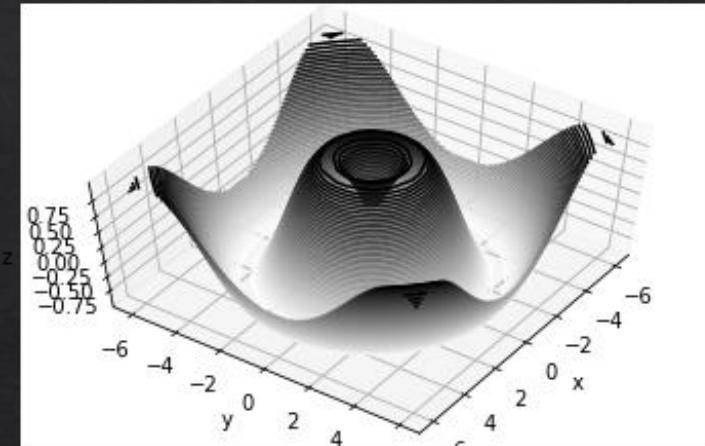
Лекция 2

- ◆ Градиентный спуск.
- ◆ Генетический алгоритм
- ◆ Сбор и обработка данных.
- ◆ Корреляции и зависимости в данных.
- ◆ Обучение с учителем.
- ◆ Задачи классификации и регрессии.



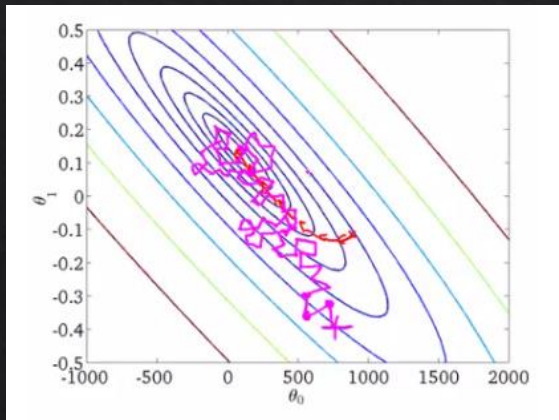
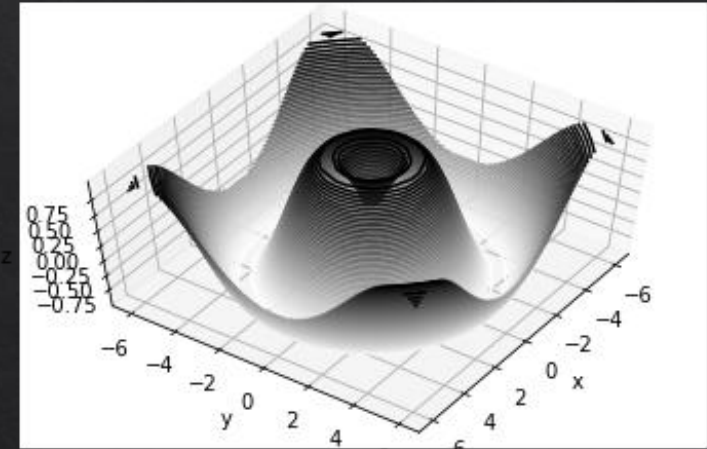
Градиентный спуск

- ✧ Рассмотрим ситуацию, когда нам надо решить задачу оптимизации – т.е. найти минимум некоторой функции $f(x_1, x_2, \dots, x_n)$. Пример – оптимизация весов нейронной сети при помощи минимизации функционала ошибки.
 - ✧ Градиент функции $\text{grad}(f(x)) = 0$ – условие минимума
 - ✧ Зачастую функция сложная и имеет много переменных
 - ✧ Поэтому искать минимум надо численно
 - ✧ Проблема в том, что длительность и аккуратность метода зависят от числа шагов и величины шага, а так же начального значения x_0 .
 - ✧ Локальные минимумы – это ловушка.
-
- ✧ $x_{n+1} = x_n - \text{step} \cdot \nabla f(x_n)$
 - ✧ Остановка: $x_{n+1} - x_n < \varepsilon$ или ограничение на количество шагов



Стохастический градиентный спуск

- ◆ Чтобы выбраться из локального минимума, мы можем двигаться в некотором случайном направлении – случайный поиск $\frac{f(x) - f(x + \eta \cdot l)}{\eta}$, где η – шаг, l – единичный вектор случайного направления
- ◆ SGD (Stochastic Gradient Descent)
- ◆ Рассмотрим функционал $SE = f(\omega_i, x_i, y_i)$
- ◆ Оптимизация весов: $\omega_i^1 = \omega_i^0 - \sum_x \eta \frac{\partial SE}{\partial \omega_i} \Big|_0$



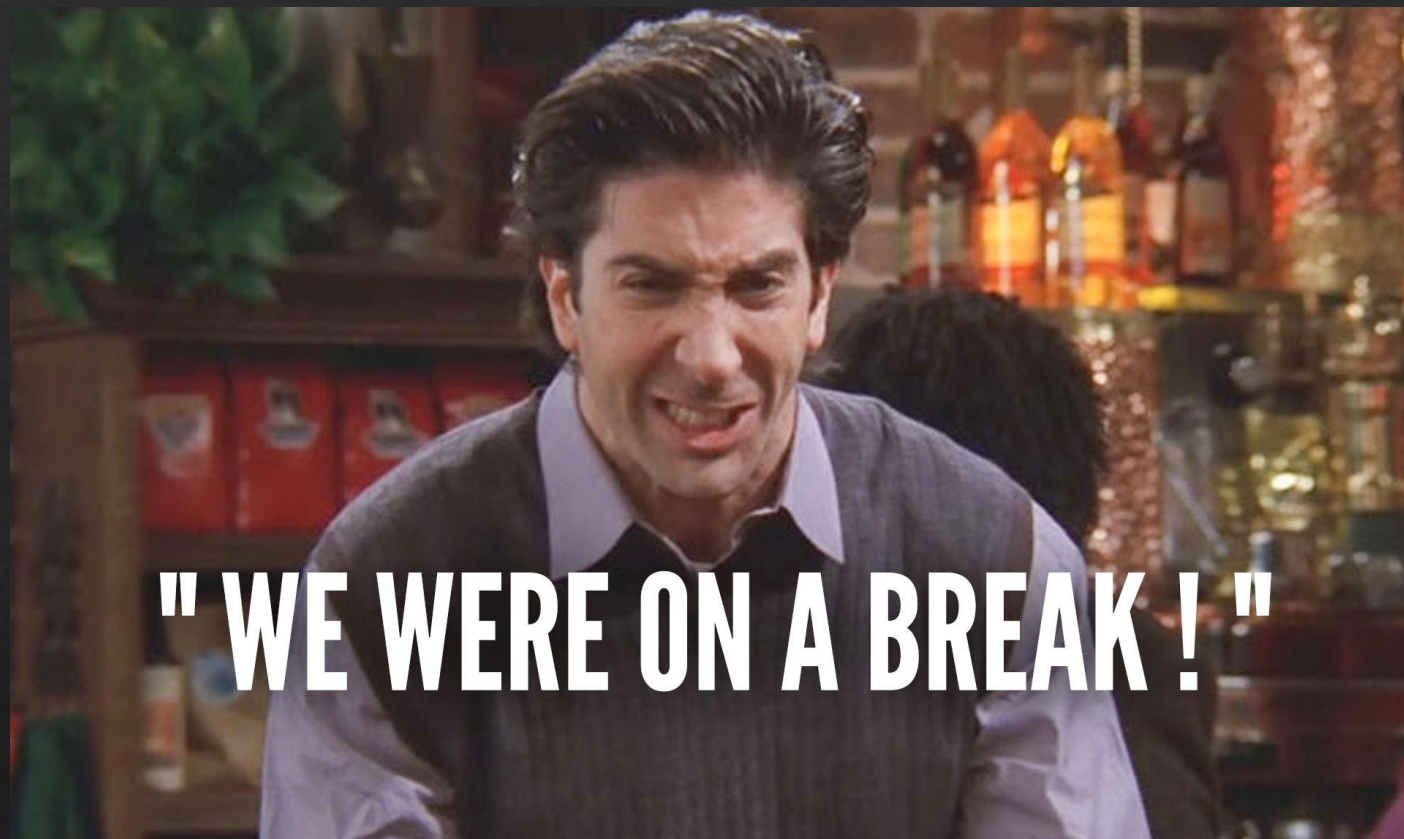
Основная идея – избавиться от суммы, т.е. проводить оптимизацию по случайным объектам, аналогично случайному поиску для функционалов

Генетический алгоритм – дифференциальная ЭВОЛЮЦИЯ

◆ Основные шаги:

- Выбор популяции: выбор некоторого числа объектов - N векторов из пространства R^n
- Мутация: - изменение одного из объектов - $C' = C - F * (A - B)$
- Скрещивание: создание нового объекта T , который наследует от мутировавшего C' с некоторой вероятностью P , от немутировавшего родителя с вероятностью $1-P$
- Отбор: согласно критериям для решения задачи –после формирования потомка T , сравниваем например $MSE(T)$ и $MSE(X)$, кто лучше, тот попадает в следующую популяцию.

Перерыв



" WE WERE ON A BREAK ! "

Домашнее задание

- ◇ <http://www.deeplearningbook.org/contents/numerical.html>
- ◇ Baldi, P., P. Sadowski, and D. Whiteson. “Searching for Exotic Particles in High-energy Physics with Deep Learning.” Nature Communications 5 (July 2, 2014).
- ◇ Реализовать градиентный спуск и стохастический градиентный спуск для среднеквадратичной ошибки в Jupyter Notebook.

Вопросы?

