Национальный исследовательский университет ИТМО
Факультет информационных технологий и программирования
Прикладная математика и информатика

# Методы оптимизации

Отчет по лабораторной работе №3

⟨Собрано 5 июня 2023 г.⟩

**Работу выполнили:**
Бактурин Савелий Филиппович M32331
Вереня Андрей Тарасович M32331
Сотников Максим Владимирович M32331

**Преподаватель:**
Ким Станислав Евгеньевич

# Решение задачи нелинейной регрессии

Часто решая задачу создания регрессионной модели мы сталкиваемся с тем, что по жизни очень немногие рассматриваемые функции оказываются не представимы в виде обобщенной линейной зависимости или полиномиальной некоторой конечной степени $k$. Такая же ситуация часто случается и с некоторым набором данных, который нужно как-то обобщить. Именно в таких случаях к нам на помощь приходит более частный случай регрессионного анализа – *нелинейная регрессия*.

Идея построения нелинейной регрессии как и в случае с полиномиальной заключается в том, чтобы найти математическую функцию, которая максимально точно описывает зависимость между независимой переменной и зависимой от нее. Например, для построения нелинейной регрессии можно использовать функции типа полинома, логарифмической или экспоненциальной зависимости.

В целом весь процесс нахождения нелинейной регрессионной модели можно поделить на два этапа:

- ▷ Определить регрессионную модель $f(w, x)$, которая зависит от параметров $w = (w_1, \ \dots, \ w_W)$ и свободной переменной $x$.
- ▷ Решить задачу по нахождению минимума сумма квадратов регрессионных остатков:
$$S = \sum_{i=1}^{m} r_i^2, \ r_i = y_i - f(w, x_i)$$

Однако, решая в лоб такую задачу, мы сталкиваемся с оптимизационной задачи нахождения параметров нелинейной регрессионной модели. Тут к нам и приходят на помощь различные методы нахождения, в том числе и рассматриваемые ниже: *Gauss-Newton* и *Powell Dog Leg*.

## Gauss-Newton

Напомним, что мы решаем следующую задачу: дана нелинейная модель $f(w, x)$, где $w \in \mathbb{R}^n$, тогда сумма квадратов регрессионных остатков высчитывается как

$$S = \sum_{i=1}^{\texttt{sizeof}\ X} (f(w, x_i) - y_i)^2 \to \min$$

Итак, теперь введем некоторые новые объекты для решения задачи, пусть $w^0 = (w_0^0, \ w_1^0, \ \dots, \ w_p^t)$ – начальное приближение, и

$$\daleth = \left( \frac{\partial f}{\partial w_j}(w^t, x_i) \right)_{(\texttt{sizeof}\ X) \times p} \ - \ \text{Якобиан, или матрица первых производных}$$
$$\vec{f}_t = \big( f(w^t, x_i) \big)_{(\texttt{sizeof}\ X) \times 1} \ - \ \text{вектор значений } f$$
$$\eth_t = \texttt{const} \ - \ \text{размера шага}$$

Тогда, формула $t$-й итерации рассматриваемого метода будет высчитываться как

$$w^{t+1} \leftarrow w^t - \eth_t \cdot \underbrace{\left(\daleth_t^{\mathrm{T}} \daleth_t\right)^{-1} \daleth_t^{\mathrm{T}}}_{\Psi} (\vec{f}_t - y),$$

где $\Psi$ – это псевдообратная матрица, или решение некоторой задачи многомерной линейной регрессии, где мы ищем такой вектор $\Psi$, что

$$\left\| \beth_t \Psi - (\vec{f}_t - y) \right\|^2 \to \min,$$

где $y$ – вектор правильных/настоящих ответов нашей модели. Получается, для решения задачи, мы, так называемую, *невязку* пытаемся приблизить линейной комбинацией вектора из матрицы Якобиана так, что при следующем шаге итерации получить такой $w^{t+1}$, который бы сократил нам расстояние невязки. Причем, заметим, что на каждом шаге, задача будет новой, так как $\beth_t$ зависит от текущего приближения, чтобы решить задачу многомерной регрессии.

### Исследования

## Powell Dog Leg

### Исследования

# BFGS

## Исследования

## L-BFGS

### Исследования