Задача 1.

Рассмотрим аппроксимирующую функцию

$$f(x) = \alpha x^{\gamma} \tag{1}$$

которая используется для нахождения приближения дискретных данных $\{(x_i, y_i)\}_{i=1}^{n}$. Требуется записать формулировку оптимизационной задачи для метода наименьших квадратов для нахождения коэффициентов α, γ и найти ее аналитическое решение.

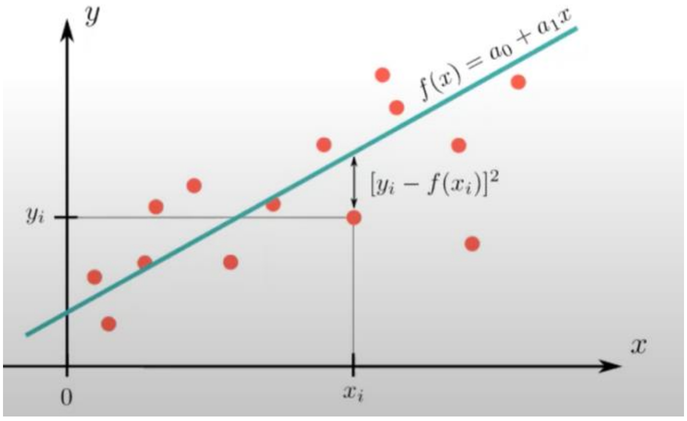
Решение:

Рассмотрим:

- Набор дискретных данных $D = \{(x_i, y_i)\}_{i=1}^{n}$
- И аппроксимирующую функцию $f(x) = a_0 + a_1 \cdot x$

Требуется найти такие a_0 и a_1, что $f(x)$ приближается к данным D наилучшим образом с помощью минимизации суммы квадратов отклонений:

$$\min_{a_0, a_1} E_2(a_0, a_1) = \min_{a_0, a_1} \sum_{i=1}^{n} (y_i - a_0 - a_1 \cdot x)^2$$



Прологарифмируем функцию f(x), чтобы линеаризовать ее:

$$\ln\big(f(x)\big) = \ln(\alpha x^{\gamma}) = \ln(\alpha) + \gamma \ln(x)$$

$$\min_{a_0, a_1} E_2(a_0, a_1) = \min_{a_0, a_1} \sum_{i=1}^{n} (\ln(y_i) - \ln(\alpha) - \gamma \ln(x))^2$$

Чтобы решить задачу минимизации, мы должны взять частные производные от функции E_2 по $a_0 = \ln(\alpha)$ и по $a_1 = \gamma$ и приравнять их к нулю.

$$\frac{\partial E_2}{\partial(\ln(\alpha))} = 0 \Rightarrow -2\sum_{i=1}^{n}(\ln(y_i) - \ln(\alpha) - \gamma\ln(x_i)) = 0$$

$$\frac{\partial E_2}{\partial(\gamma)} = 0 \Rightarrow -2\sum_{i=1}^{n}\ln(x_i)\cdot(\ln(y_i) - \ln(\alpha) - \gamma\ln(x_i)) = 0$$

$$\begin{cases} \sum_{i=1}^{n}\ln(y_i) - \ln(\alpha)\cdot n - \gamma\cdot\sum_{i=1}^{n}\ln(x_i) = 0 \\ \sum_{i=1}^{n}\ln(y_i)\cdot\ln(x_i) - \ln(\alpha)\sum_{i=1}^{n}\ln(x_i) - \gamma\cdot\sum_{i=1}^{n}(\ln(x_i))^2 = 0 \end{cases}$$

Поиск неизвестных коэффициентов осуществляется по МНК в соответствии с системой уравнений:

$$\begin{cases} \ln(\alpha)\cdot n + \gamma\cdot\sum_{i=1}^{n}\ln(x_i) = \sum_{i=1}^{n}\ln(y_i) \\ \ln(\alpha)\cdot\sum_{i=1}^{n}\ln(x_i) + \gamma\cdot\sum_{i=1}^{n}(\ln(x_i))^2 = \sum_{i=1}^{n}\ln(y_i)\cdot\ln(x_i) \end{cases}$$

Решаем полученную выше систему линейных уравнений.

Коэффициенты аппроксимирующей функции в аналитическом виде определяются следующим образом:

$$\ln(\alpha)^{(opt)} = \frac{\sum_{i=1}^{n}\ln(y_i)\cdot\sum_{i=1}^{n}(\ln(x_i))^2 - \sum_{i=1}^{n}(\ln(y_i)\cdot\ln(x_i))\cdot\sum_{i=1}^{n}\ln(x_i)}{n\cdot\sum_{i=1}^{n}(\ln(x_i))^2 - (\sum_{i=1}^{n}\ln(x_i))^2}$$

$$\gamma^{(opt)} = \frac{n\cdot\sum_{i=1}^{n}(\ln(y_i)\cdot\ln(x_i)) - \sum_{i=1}^{n}\ln(y_i)\cdot\sum_{i=1}^{n}\ln(x_i)}{n\cdot\sum_{i=1}^{n}(\ln(x_i))^2 - (\sum_{i=1}^{n}\ln(x_i))^2}$$

Оптимальные коэффициенты $\ln(\alpha)^{(opt)}$ и $\gamma^{(opt)}$ для задачи линейной регрессии в одномерном случае, оптимального приближения прямой к заданным дискретным узлам $D = \{(x_i, y_i)\}_{i=1}^{n}$.