

Задание 1.

I. Сгенерируйте выборки для заданных функций с моделированием случайной ошибки измерения.

1). Сформируйте выборку $\{(x^{(i)}, y^{(i)})\}_{i=1}^N$, $y^{(i)} = f(x^{(i)}) + \varepsilon^{(i)}$, где $\varepsilon^{(i)}$ — ошибка, генерируемая случайно, значение которой принадлежит заданному интервалу $[-\varepsilon_0, +\varepsilon_0]$, $\varepsilon_0 > 0$. Аргументы $x^{(i)}$ генерируются случайно в соответствии с равномерным распределением на интервале $[-1, 1]$.

Варианты распределения ошибки на интервале $[-\varepsilon_0, +\varepsilon_0]$:

- а) ошибка ε распределена равномерно,
- б) ошибка ε распределена нормально.

Варианты функции f :

- а) $f = ax^3 + bx^2 + cx + d$, коэффициенты a, b, c, d сгенерировать случайно из интервала $[-3, 3]$;
- б) $f = x \sin(2\pi x)$.

2). Отобразите на координатной плоскости график функции f и выборку, полученную по ней. Выберите разные варианты значения ε_0 и разные параметры нормального распределения.

II. Восстановление функциональной зависимости с помощью полиномиальной регрессии.

Реализуйте с помощью полиномиальной регрессии для полученной выборки восстановление функциональной зависимости, по которой получена выборка. Приведите примеры выборок и степеней полиномов, при которых

- а) происходит недообучение,
- б) происходит переобучение,
- в) полученная функциональная зависимость пригодна для прогнозирования значения восстанавливаемой функциональной зависимости в x , не содержащемся в выборке.

Разбейте сгенерированную выборку на обучающую и тестовую. Приведите значения функции ошибки на обучающей и тестовой выборке для указанных выше ситуаций.

Полиномиальная регрессия — стр. 47–52 учебного пособия, прилагаемого к заданию (файл `mtgo.pdf`). Формулы для составления системы линейных алгебраических уравнений — задача 2 на стр. 64.

Используйте библиотеки Python `numpy`, `math`, `scipy`, модуль `matplotlib.pyplot`.