

Практика 7. Анализ регрессионных остатков и погрешности регрессионных моделей характеристик насосных агрегатов

Часть 1. Анализ регрессионные остатков

1. Сформулировать задачу анализа регрессионных остатков критерием серий (Вальда-Вилкоксона). Применить функцию `runstest` для анализа остатков моделей всех порядков из предыдущего ДЗ по напорной характеристике и характеристике мощности.
2. Выбрать модель минимального порядка, для которой гипотеза о наличии закономерности в данных отклоняется.

Часть 2. Скользящий контроль

1. Построить оценки погрешности модели на скользящем контроле LOO (leave-one-out):
 - a. Коэффициент детерминации R^2 , усредненный по экзаменационным выборкам
 - b. Коэффициент детерминации R^2 , усредненный по обучающим выборкам
 - c. Оценка дисперсия шума, построенная по экзаменационным выборкам (как в лекциях).
 - d. Несмещенная оценка дисперсии шума, построенная по обучающим выборкам (для сравнения с предыдущей).
 - e. Проявился ли эффект оптимистической смещенности
2. Для каждого разбиения выборки на скользящем контроле построить 95%-доверительный интервал предсказания нового значения. Вывести количество попаданий прогноза в 95%-доверительный интервал.

Вопросы к защите (20.05.2017)

1. Статистический критерий серий.
 - a. Случайный ряд, ряд с трендом, ряд с периодичностью. Исходя из определения серии, указать, для какого типа ряда их количество N_S больше или меньше.
 - b. Формулировки основной и альтернативной гипотезы.
 - c. Распределения статистики N_S и стандартизированной статистики z в условиях альтернативной гипотезы (точная форма распределения).
 - d. Распределения статистик N_S и z в условиях альтернативной гипотезы (эскиз).
2. Анализ регрессионных остатков
 - a. Типичный вид графика регрессионных остатков при заниженном порядке аппроксимации.
 - b. Статистические свойства регрессионных остатков.
 - c. Какие из статистических свойств нарушаются при занижении порядка аппроксимации?
 - d. Применение критерия серия к анализу коррелированности регрессионных остатков. Какой тип ряда с точки зрения критерия серий формируется при занижении порядка аппроксимации?
 - e. Какую модель следует выбрать, исходя из анализа регрессионных остатков.
3. Скользящий контроль
 - a. Что такое эмпирическая модель? Какие еще бывают модели, которые нельзя назвать эмпирическими?
 - b. Что такое обучающая и экзаменационная выборка?
 - c. Какая проблема возникает при оценивании погрешности эмпирической модели по обучающей выборке?
 - d. Что такое проблема переобучения?
 - e. Контроль по одному элементу и контроль по блокам (LOO, q-fold).