

Практика 5. Регрессия. МНК и свойства МНК-оценок

Часть 1. Критерий МНК как функция коэффициентов регрессии в трехмерном пространстве

1. Сгенерировать выборку объема n по модели из своего варианта на сайте. Вариант выбирается по списку на сайте. Для модели заданы истинные коэффициенты регрессии и дисперсия шума. Объем выборки n задать от 10 до 20.
2. Рассчитать МНК-оценки коэффициентов регрессии на основе аналитического решения задачи минимизации функционала качества. Расчет сделать функцией regress и по формуле:
$$\hat{\beta} = (X^T X)^{-1} X^T Y$$
3. Построить трехмерный график (surf) и график изолиний (contour) функционала качества МНК $Q(\beta)$. Отобразить точкой аналитическое решение (plot3)

Часть 2. Статистические свойства МНК-оценок

1. Методом статистического моделирования построить распределения оценок коэффициентов регрессии как случайных величин. По результатам численного моделирования определить 95% доверительный интервал коэффициентов.
2. Методом статистического моделирования построить полигон относительных частот прогноза $\hat{y}(x_n)$ при $x=200$ и $x = 240$.
3. Методом статистического моделирования построить полигон относительных частот ошибки предсказания истинного значения $e_n = \hat{y}(x) - y(x)$ при $x=200$ и $x = 300$.
4. Методом статистического моделирования построить полигон относительных частот распределения ошибки предсказания нового измерения $e_n = \hat{y}(x) - y_n$ при $x=200$ и $x = 300$.
5. Построить теоретические плотности распределения ошибки предсказания нового измерения e_n при $x=200$ и $x = 300$. Отметить 95%-доверительный диапазон. График совместить с графиками из предыдущего пункта.

Вопросы к защите

1. Объяснение невязок между моделью и данными как реализаций случайного шума.
2. Вероятностная модель выборки до измерений (векторный вид, скалярный вид).
3. Как сформировать вектор факторов
 - a. для задания линейной зависимости
 - b. для задания полиномиальной зависимости
4. Формальная постановка задачи поиска оценок коэффициентов регрессии как задачи минимизации функционала (в векторном виде).
5. Что такое система нормальных уравнений? Как она выводится?
6. Оценки коэффициентов регрессии $\hat{\beta}$ как случайные величины.
 - a. как проявляется их случайность?
 - b. доказательство линейности связи $\hat{\beta}$ и вектора шума
 - c. доказательство несмещенности оценок $\hat{\beta}$
 - d. вывод ковариационной матрицы оценок $cov(\hat{\beta})$
 - e. какое теоретическое распределение имеет коэффициент $\hat{\beta}_i$?
7. Прогноз регрессионной модели как случайная величина
 - a. ошибки прогноза нового значения и истинного значения
 - b. доказательство несмещенности прогноза
 - c. доверительный интервал прогноза при известной дисперсии
 - i. прогноз истинного значения
 - ii. прогноз нового значения