Задачник по эконометрике-1

(с шахматами и поэтэссами)

Дмитрий Борзых, Борис Демешев

5 октября 2012 г.

1. Регрессионная модель задана в матричном виде при помощи уравнения $y = X\beta + \varepsilon$, где $\beta = (\beta_1, \beta_2, \beta_3)'$. Известно, что $\mathbb{E}(\varepsilon) = 0$ и $\mathrm{Var}(\varepsilon) = \sigma^2 \cdot I$. Известно также, что $y =$, $X =$. Для удобства расчетов ниже приведены матрицы $X'X =$ и $(X'X)^{-1} =$.
   (a) Укажите число наблюдений.
   (b) Укажите число регрессоров с учетом свободного члена.
   (c) Рассчитайте $TSS = \sum(y_i - \bar{y})^2$, $RSS = \sum(y_i - \hat{y}_i)^2$ и $ESS = \sum(\hat{y}_i - \bar{y})^2$.
   (d) Рассчитайте при помощи метода наименьших квадратов $\hat{\beta}$, оценку для вектора неизвестных коэффициентов.
   (e) Чему равен $\hat{\varepsilon}_5$, МНК-остаток регрессии, соответствующий 5-ому наблюдению?
   (f) Чему равен R^2 в модели? Прокомментируйте полученное значение с точки зрения качества оцененного уравнения регрессии.
   (g) Используя приведенные выше данные, рассчитайте несмещенную оценку для неизвестного параметра σ^2 регрессионной модели.
   (h) Рассчитайте $\widehat{\mathrm{Cov}}(\hat{\beta})$, оценку для ковариационной матрицы вектора МНК-коэффициентов $\hat{\beta}$.
   (i) Найдите $\widehat{\mathrm{Var}}(\hat{\beta}_1)$, несмещенную оценку дисперсии МНК-коэффициента $\hat{\beta}_1$.
   (j) Найдите $\widehat{\mathrm{Var}}(\hat{\beta}_2)$, несмещенную оценку дисперсии МНК-коэффициента $\hat{\beta}_2$.
   (k) Найдите $\widehat{\mathrm{Cov}}(\hat{\beta}_1, \hat{\beta}_2)$, несмещенную оценку ковариации МНК-коэффициентов $\hat{\beta}_1$ и $\hat{\beta}_2$.
   (l) Найдите $\widehat{\mathrm{Var}}(\hat{\beta}_1 + \hat{\beta}_2)$, $\widehat{\mathrm{Var}}(\hat{\beta}_1 - \hat{\beta}_2)$, $\widehat{\mathrm{Var}}(\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3)$, $\widehat{\mathrm{Var}}(\hat{\beta}_1 + \hat{\beta}_2 - 2\hat{\beta}_3)$
   (m) Найдите $\mathrm{Corr}(\hat{\beta}_1, \hat{\beta}_2)$, оценку коэффициента корреляции МНК-коэффициентов $\hat{\beta}_1$ и $\hat{\beta}_2$.
   (n) Найдите $s_{\hat{\beta}_1}$, стандартную ошибку МНК-коэффициента $\hat{\beta}_1$.
2. Априори известно, что парная регрессия должна проходить через точку (x_0, y_0).
   (a) Выведите формулы МНК оценок;
   (b) В предположениях теоремы Гаусса-Маркова найдите дисперсии и средние оценок

   Вроде бы равносильно переносу начала координат и применению результата для регрессии без свободного члена. Должна остаться несмещенность.
3. Слитки-вариант. Перед нами два золотых слитка и весы, производящие взвешивания с ошибками. Взвесив первый слиток, мы получили результат 300 грамм, взвесив второй слиток — 200 грамм, взвесив оба слитка — 400 грамм. Предположим, что ошибки взвешивания — независимые одинаково распределенные случайные величины с нулевым средним.
   (a) Найдите несмещеную оценку веса первого шара, обладающую наименьшей дисперсией.

(b) Как можно проинтерпретировать нулевое математическое ожидание ошибки взвешивания?

Как отсутствие систематической ошибки.

4. Вася считает, что $\mathrm{sCov}(y,\hat{y}) = \frac{\sum(y_i-\bar{y})(\hat{y}_i-\bar{y})}{\sqrt{\sum(y_i-\bar{y})^2\sum(\hat{y}_i-\bar{y})^2}}$ это неплохая оценка для $\mathrm{Cov}(y_i,\hat{y}_i)$. Прав ли он? Не прав. Ковариация $\mathrm{Cov}(y_i,\hat{y}_i)$ зависит от i, это не одно неизвестное число, для которого можно предложить одну оценку.

1 МНК без матриц и вероятностей

1. Даны n пар чисел: (x_1,y_1), ..., (x_n,y_n). Мы прогнозируем y_i по формуле $\hat{y}_i = \hat{\beta}x_i$. Найдите $\hat{\beta}$ методом наименьших квадратов. $\hat{\beta} = \sum x_i y_i / \sum x_i^2$
2. Даны n чисел: y_1, ..., y_n. Мы прогнозируем y_i по формуле $\hat{y}_i = \hat{\beta}$. Найдите $\hat{\beta}$ методом наименьших квадратов. $\hat{\beta} = \bar{y}$
3. Даны n пар чисел: (x_1,y_1), ..., (x_n,y_n). Мы прогнозируем y_i по формуле $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i$. Найдите $\hat{\beta}_1$ и $\hat{\beta}_2$ методом наименьших квадратов. $\hat{\beta}_2 = \sum(x_i-\bar{x})(y_i-\bar{y})/\sum(x_i-\bar{x})^2$, $\hat{\beta}_1 = \bar{y} - \hat{\beta}_2\bar{x}$
4. Даны n пар чисел: (x_1,y_1), ..., (x_n,y_n). Мы прогнозируем y_i по формуле $\hat{y}_i = 1 + \hat{\beta}x_i$. Найдите $\hat{\beta}$ методом наименьших квадратов. $\hat{\beta} = \sum x_i(y_i-1)/\sum x_i^2$
5. Перед нами два золотых слитка и весы, производящие взвешивания с ошибками. Взвесив первый слиток, мы получили результат 300 грамм, взвесив второй слиток — 200 грамм, взвесив оба слитка — 400 грамм. Оцените вес каждого слитка методом наименьших квадратов. $(300-\hat{\beta}_1)^2 + (200-\hat{\beta}_2)^2 + (400-\hat{\beta}_1-\hat{\beta}_2)^2 \to \min$
6. Аня и Настя утверждают, что лектор опоздал на 10 минут. Таня считает, что лектор опоздал на 3 минуты. С помощью мнк оцените на сколько опоздал лектор. $2\cdot(10-\hat{\beta})^2 + (3-\hat{\beta})^2 \to \min$
7. Регрессия на дамми-переменную...
8. Функция $f(x)$ дифференциируема на отрезке $[0;1]$. Найдите аналог МНК-оценок для регрессии без свободного члена в непрерывном случае. Более подробно: найдите минимум по $\hat{\beta}$ для функции
$$Q(\hat{\beta}) = \int_0^1 (f(x)-\hat{\beta}x)^2\,dx \qquad (1)$$

2 Проекция, Картинка

1. Найдите на Картинке четыре прямоугольных треугольника. Сформулируйте четыре теоремы Пифагора. $\sum y_i^2 = \sum \hat{y}_i^2 + \sum \hat{\varepsilon}_i^2$, $TSS = ESS + RSS$,
2. Покажите на Картинке TSS, ESS, RSS, R^2, $\mathrm{sCov}(\hat{y},y)$
3. Предложите аналог R^2 для случая, когда константа среди регрессоров отсутствует. Аналог должен быть всегда в диапазоне $[0;1]$, совпадать с обычным R^2, когда среди регрессоров есть константа, равняться единице в случае нулевого $\hat{\varepsilon}$. Спроецируем единичный столбец на «плоскость», обозначим его $1'$. Делаем проекцию y на «плоскость» и на $1'$. Далее аналогично.
4. Вася оценил регрессию y на константу, x и z. А затем, делать ему нечего, регрессию y на константу и полученный $\hat{y}$. Какие оценки коэффициентов у него получатся? Чему будет равна оценка дисперсии коэффицента при $\hat{y}$? Почему оценка коэффициента неслучайна, а оценка её дисперсии положительна? Проекция y на $\hat{y}$ это $\hat{y}$, поэтому оценки коэффициентов будут 0 и 1. Оценка дисперсии $\frac{RSS}{(n-2)ESS}$. Нарушены предпосылки теоремы Гаусса-Маркова, например, ошибки новой модели в сумме дают 0, значит коррелированы.

3 Голая линейная алгебра

Здесь будет собран минимум задач по линейной алгебре.

1. Приведите пример таких A и B, что $\det(AB) \neq \det(BA)$. Например, $A = (1, 2, 3)$, $B = (1, 0, 1)'$