1. Методы получения оценок

Методы получения оценок: метод максимального правдоподобия, метод моментов, метод наименьших квадратов.

2. Свойства оценок

Свойства оценок: несмещённость, состоятельность, эффективность в классе.

3. Асимптотические методы

Центральная предельная теорема. Лемма Слуцкого. Дельта-метод. Построение асимптотических доверительных интервалов.

4. Святая троица тестов

Три классических теста: LM, LR, Wald.

Чёрный трек: тесты в матричной форме для вектора параметров?

5. ШБ — МНК

МНК в скалярной и матричной форме без статистических свойств. Строгая мультиколлинеарность.

Рассказать, что коэффициенты при стандартизации всех переменных называют частными корреляциями.

Коммент: Здесь первый раз говорим слова "строгая мультиколлинеарность".

Чёрный трэк: нелинейный мнк численно?

Задачи для доски:

МНК и R2 руками на доске

Задачи для колаба:

МНК и R2

Рост R2 с ростом числа регрессоров

Рост RSS с ростом числа наблюдений

6. Предпосылки о математическом ожидании и дисперсии

МНК со статистическими предпосылками на ожидание и дисперсию. Теорема Гаусса-Маркова.

6.1. Ожидание и ковариационная матрица

Пусть r — случайный вектор размерности $n \times 1$, s — случайный вектор размерности $k \times 1$. A и b — неслучайные матрица и вектор соответственно, имеющие подходящие размерности.

Математическим ожиданием случайного вектора r называется вектор

$$\mathbb{E}(r) = \begin{pmatrix} \mathbb{E}(r_1) \\ \mathbb{E}(r_2) \\ \dots \\ \mathbb{E}(r_n) \end{pmatrix}.$$

Ковариационная матрица вектора r определяется следующим образом:

$$\mathbb{V}\mathrm{ar}(r) = \begin{pmatrix} \mathbb{C}\mathrm{ov}(r_1, r_1) & \mathbb{C}\mathrm{ov}(r_1, r_2) & \dots & \mathbb{C}\mathrm{ov}(r_1, r_n) \\ \mathbb{C}\mathrm{ov}(r_2, r_1) & \mathbb{C}\mathrm{ov}(r_2, r_2) & \dots & \mathbb{C}\mathrm{ov}(r_2, r_n) \\ \dots & & \dots & \dots \\ \mathbb{C}\mathrm{ov}(r_n, r_1) & \mathbb{C}\mathrm{ov}(r_n, r_2) & \dots & \mathbb{C}\mathrm{ov}(r_n, r_n) \end{pmatrix}.$$

Ковариационная матрица векторов r и s определяется следующим образом:

$$\mathbb{C}\mathrm{ov}(r, s) = \begin{pmatrix} \mathbb{C}\mathrm{ov}(r_1, s_1) & \mathbb{C}\mathrm{ov}(r_1, s_2) & \dots & \mathbb{C}\mathrm{ov}(r_1, s_k) \\ \mathbb{C}\mathrm{ov}(r_2, s_1) & \mathbb{C}\mathrm{ov}(r_2, s_2) & \dots & \mathbb{C}\mathrm{ov}(r_2, s_k) \\ \dots & \dots & \dots & \dots \\ \mathbb{C}\mathrm{ov}(r_n, s_1) & \mathbb{C}\mathrm{ov}(r_n, s_2) & \dots & \mathbb{C}\mathrm{ov}(r_n, s_k) \end{pmatrix}.$$

Далее рассмотрим свойства для вектора математических ожиданий и ковариационной матрицы:

а) $\mathbb{E}(Ar + b) = A\,\mathbb{E}(r) + b$

б) $\mathbb{C}\mathrm{ov}(r, s) = \mathbb{E}(rs^T) - \mathbb{E}(r)\,\mathbb{E}(s^T)$

в) $\mathbb{C}\mathrm{ov}(Ar + b, s) = A\,\mathbb{C}\mathrm{ov}(r, s)$

г) $\mathbb{C}\mathrm{ov}(r, As + b) = \mathbb{C}\mathrm{ov}(r, s)A^T$

д) $\mathbb{V}\mathrm{ar}(r) = \mathbb{C}\mathrm{ov}(r, r) = \mathbb{E}(rr^T) - \mathbb{E}(r)\,\mathbb{E}(r^T)$

е) $\mathbb{V}\mathrm{ar}(Ar + b) = A\,\mathbb{V}\mathrm{ar}(r)A^T$

ж) $\mathbb{E}(r^T Ar) = \mathrm{trace}(A\,\mathbb{V}\mathrm{ar}(r)) + \mathbb{E}(r^T)A\,\mathbb{E}(r)$

6.2. Иерархия зависимостей случайных величин

Можно выделить три степени независимости случайных величин. Рассмотрим их на примере пары произвольных величин r и s.

$$\begin{tikzcd}
\boxed{r \text{ и } s \text{ независимы}} \arrow[r, Leftrightarrow] \arrow[d, Rightarrow] & \boxed{\mathbb{C}\mathrm{ov}(h_1(r), h_2(s)) = 0 \text{ для любых функций } h_1 \text{ и } h_2} \\
\boxed{\mathbb{E}(r \mid s) = \mathbb{E}(r)} \arrow[r, Leftrightarrow] \arrow[d, Rightarrow] & \boxed{\mathbb{C}\mathrm{ov}(r, h(s)) = 0 \text{ для любой функции } h} \\
\boxed{\begin{array}{c} r \text{ и } s \text{ линейно-независимы,} \\ \mathrm{BestLin}(r \mid s) = \mathbb{E}(r) \end{array}} \arrow[r, Leftrightarrow] & \boxed{\mathbb{C}\mathrm{ov}(r, s) = 0}
\end{tikzcd}$$

Напомним, что случайные величины r и s называются независимыми, если для любых числовых множеств A и B независимы события $\{r \in A\}$ и $\{s \in B\}$:

$$\mathbb{P}(r \in A, s \in B) = \mathbb{P}(r \in A) \cdot \mathbb{P}(s \in B)$$

Определить линейную независимость величин r и s можно по-разному. Некоторые авторы считают условие $\mathbb{C}\mathrm{ov}(r, s) = 0$ определением линейной независимости, в этом случае нижняя эквивалентность на графике тривиальна. Нам кажется более логичным другой подход. Величины r и s линейно независимы, если наилучшее линейное приближение r с помощью s не зависит от s.

Задача 6.1 *Покажем, что из равенства условного и безусловного математических ожиданий не следует независимость случайных величин:*

	1/3	1/3	1/3
r	-1	1	0
s	0	0	1

Определение 6.2 *Наилучшее линейное приближение величины r с помощью величины s — это линейная функция от s,*

$$\mathrm{BestLin}(r \mid s) = \alpha + \beta s,$$

где константы α и β находятся из решения задачи оптимизации $\mathbb{E}((r - \mathrm{BestLin}(r, s)^2) \to \min_{\alpha,\beta}$.

6.3. Теорема Гаусса — Маркова

Чтобы исследовать свойства полученной точечной оценки $\hat{\beta}$ нам потребуются предпосылки о математическом ожидании и ковариационной матрице вектора u.

Мы предположим, что случайные ошибки в среднем равны нулю, а именно,

$$\mathbb{E}(u \mid X) = 0.$$

Предпосылку о математическом ожидании можно записать и в скалярном виде,

$$\mathbb{E}(u_i \mid X) = 0, \quad \text{при } \forall i \in \{1, \ldots, n\}.$$

Теорема 6.3 *Если*

а) *Модель линейна по параметрам: $y = X\beta + u$;*

б) *Матрица X размера $[n \times k]$ имеет полный ранг k.*

в) *Условное ожидание ошибок равно нулю, $\mathbb{E}(u \mid X) = 0$;*

г) *Условная ковариационная матрица ошибок пропорциональна единичной, $\mathbb{V}\mathrm{ar}(u \mid X) = \sigma^2 I$;*

д) *Оценка $\hat{\beta}$ получена методом наименьших квадратов, $\hat{\beta} = (X^T X)^{-1} X^T y$;*

е) *Альтернативная оценка $\hat{\beta}^{alt}$ является условно несмещёнными $\mathbb{E}(\hat{\beta}^{alt} \mid X) = \beta$ и линейной по y;*

то

а) *Оценка $\hat{\beta}$ является линейной по y;*

б) Оценка $\hat{\beta}$ является условно несмещённой, $\mathbb{E}(\hat{\beta} \mid X) = \beta$ и несмещённой, $\mathbb{E}(\hat{\beta}) = \beta$;

в) Оценка любого коэффициента $\hat{\beta}_j$ является более эффективной, чем альтернативная оценка $\hat{\beta}_j^{alt}$:

$$\mathbb{V}\mathrm{ar}(\hat{\beta}_j \mid X) \leq \mathbb{V}\mathrm{ar}(\hat{\beta}_j^{alt} \mid X).$$

Вывод теоремы можно усилить, для любой линейной комбинации коэффициентов $w^T\beta$ МНК-оценка $w^T\hat{\beta}$ эффективнее альтернативной оценки $w^T\hat{\beta}^{\mathrm{alt}}$:

$$\mathbb{V}\mathrm{ar}(w^T\hat{\beta}_j \mid X) \leq \mathbb{V}\mathrm{ar}(w^T\hat{\beta}_j^{\mathrm{alt}} \mid X).$$

6.4. Задачи для доски:

Задача 6.4 *Исследовательница Мишель собрала данные по 20 студентам. Переменная y_i — количество решённых задач по эконометрике i-ым студентом, а x_i — количество просмотренных серий любимого сериала за прошедший год. Оказалось, что $\sum y_i = 10$, $\sum x_i = 0$, $\sum x_i^2 = 40$, $\sum y_i^2 = 50$, $\sum x_i y_i = 60$.*

а) Найдите МНК-оценки коэффициентов парной регрессии.

б) В рамках предположения $\mathbb{E}(u_i \mid X) = 0$ найдите $\mathbb{E}(y_i \mid X)$, $\mathbb{E}(\hat{\beta}_j \mid X)$, $\mathbb{E}(\hat{u}_i \mid X)$, $\mathbb{E}(\hat{y}_i \mid X)$.

в) Предположим дополнительно, что $\mathbb{V}\mathrm{ar}(u_i \mid X) = \sigma^2$ и u_i при фиксированных X независимы. Найдите $\mathbb{V}\mathrm{ar}(y_i \mid X)$, $\mathbb{V}\mathrm{ar}(y_i(x_i - \bar{x}) \mid X)$, $\mathbb{V}\mathrm{ar}(\sum y_i(x_i - \bar{x}) \mid X)$, $\mathbb{V}\mathrm{ar}(\hat{\beta}_2 \mid X)$.

Задача 6.5 *Рассмотрим классическую линейную модель $y = X\beta + u$ с предпосылками Гаусса — Маркова: $\mathbb{E}(u \mid X) = 0$ и $\mathbb{V}\mathrm{ar}(u \mid X) = \sigma^2 I$. Для всех случайных векторов (y, $\hat{y}$, $\hat{\beta}$, u, $\hat{u}$, $\bar{y}$) найдите все возможные ожидания и ковариационные матрицы $\mathbb{E}(\cdot)$, $\mathbb{V}\mathrm{ar}(\cdot)$, $\mathbb{C}\mathrm{ov}(\cdot, \cdot)$.*

Задача 6.6 *Рассмотрим модель $y_i = \beta x_i + u_i$ с двумя наблюдениями, $x_1 = 1$, $x_2 = 2$. Величины u_1 и u_2 независимы и равновероятно равны $+1$ или -1.*

а) Найдите оценку $\hat{\beta}_{ols}$ для β с помощью метода наименьших квадратов.

б) Чему равна дисперсия $\mathbb{V}\mathrm{ar}(\hat{\beta}_{ols} \mid x)$ и ожидание $\mathbb{E}(\hat{\beta}_{ols} \mid x)$?

в) Постройте несмещённую оценку $\hat{\beta}_{best}$ с наименьшей дисперсией.

г) Чему равна дисперсия $\mathbb{V}\mathrm{ar}(\hat{\beta}_{best} \mid x)$?

д) А как же теорема Гаусса — Маркова? Почему в данном примере удаётся построить оценку с дисперсией меньше, чем у оценки методом наименьших квадратов?

а) $\hat{\beta}_{ols} = (y_1 + 2y_2)/5$;

б) $\mathbb{V}\mathrm{ar}(\hat{\beta}_{ols} \mid x) = 1/5$;

в) Заметим, что по величине $2y_1 - y_2$ можно однозначно восстановить величины ошибок u_1 и u_2. Например, если $2y_1 - y_2 = 3$, то $u_1 = 1$, $u_2 = -1$.

$$\hat{\beta}_{best} = \begin{cases} y_1 + 1, & \text{если } 2y_1 - y_2 < 0, \\ y_1 - 1, & \text{если } 2y_1 - y_2 > 0. \end{cases}$$

г) Шок контент, $\mathbb{V}\mathrm{ar}(\hat{\beta}_{best} \mid x) = 0$.

д) Построенная оценка $\hat{\beta}_{best}$ *является нелинейной по* y*, а теорема Гаусса — Маркова гарантирует только, что метод наименьших квадратов порождает несмещённую оценку с наименьшей дисперсией среди линейных по* y *оценок.*

Задача 6.7 *(Hansen 4.14)*

Задана модель $y = X\beta + u$*, для которой выполняются предпосылки теоремы Гаусса — Маркова. Вас интересует величина* $\theta = \beta^2$*. Пусть получены МНК-оценки коэффициентов:* $\hat{\beta}$*,* $V_{\hat{\beta}} = \mathbb{V}\mathrm{ar}[\hat{\beta}|X]$*. Тогда кажется неплохой идеей оценить* θ *как* $\hat{\theta} = \hat{\beta}^2$*.*

а) Найдите $\mathbb{E}[\hat{\theta}|X]$*. Является ли* $\hat{\theta}$ *смещённой?*

б) Предложите способ коррекции смещения для получения несмещённой оценки $\hat{\theta}^*$*, используя результаты предыдущего пункта.*

Задача 6.8 *Рассмотрим модель регрессии* $y_i = \beta_1 + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + u_i$*. Пусть все предпосылки теоремы Гаусса — Маркова выполнены. Дополнительно предположим, что* $u_i \sim N(0, \sigma^2)$*,* $i = 1, \ldots, n$*.. Дополнительно известно, что на самом деле* $\beta_2 = \ldots = \beta_k = 0$*.*

а) Найдите $\mathbb{E}(R^2)$*.*

б) Найдите $\mathbb{E}(R^2_{adj})$*.*

6.5. Задачи для колаб:

Генерация R2 для вывода распределения
Генерация смещения
Генерация лишних регерессоров
Реальный пример с лишним регрессорами (тип знаки зодиака и ретроградный)
Какая-то длинная задача, которую из темы в тему и в ней находить потом нарушения предпосылок?
https://colab.research.google.com/drive/1wFrLyGcVVETx96jS93I4z8asgAQwqIdw?usp=sharing

6.6. Чёрный трэк:

Умножение блочных матриц. Если размеры блоков допускают операцию умножения, то:

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \cdot \left[\begin{array}{c|c} E & F \\ \hline G & H \end{array}\right] = \left[\begin{array}{c|c} AE + BG & AF + BH \\ \hline CE + DG & CF + DH \end{array}\right].$$

Формула Фробениуса (блочное обращение).

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]^{-1} = \left[\begin{array}{c|c} A^{-1} + A^{-1}BH^{-1}CA^{-1} & -A^{-1}BH^{-1} \\ \hline -H^{-1}CA^{-1} & H^{-1} \end{array}\right],$$

где A — невырожденная квадратная матрица размерности $n \times n$, D — квадратная матрица размерности $k \times k$, $H = D - CA^{-1}B$.

Задача 6.9 *Пусть истинной является модель $y = X_1\beta_1 + X_2\beta_2 + u$, где X_1, X_2 — матрицы признаков размерностей $n \times k_1$ и $n \times k_2$ соответственно. Вместо истинной модели вы оцениваете модель вида $y = X_1\beta_1 + v$, где v — вектор случайной ошибки, удовлетворяющий предпосылкам теоремы Гаусса — Маркова.*

а) *Будет ли МНК-оценка вектора параметров β_1 несмещённой?*

б) *Будет ли несмещённой МНК-оценка дисперсии случайной ошибки?*

в) *Рассчитайте $\mathbb{V}\mathrm{ar}(\hat{\beta}_1)$. Не противоречит ли полученной результат теореме Гаусса — Маркова?*

Задача 6.10 *Пусть истинной является модель $y = X_1\beta_1 + u$, где X_1 — матрица признаков размерности $n \times k_1$. Вместо истинной модели вы оцениваете модель вида $y = X_1\beta_1 + X_2\beta_2 + v$, где X_2 — матрица признаков размерности $n \times k_2$, v — вектор случайной ошибки, удовлетворяющий предпосылкам теоремы Гаусса — Маркова.*

а) *Будет ли МНК-оценка вектора параметров β_1 несмещённой?*

б) *Будет ли несмещённой МНК-оценка дисперсии случайной ошибки?*

в) *Рассчитайте $\mathbb{V}\mathrm{ar}(\hat{\beta}_1)$. Не противоречит ли полученной результат теореме Гаусса — Маркова?*

7. Доверительные интервалы для коэффициентов

Построение доверительных интервалов для МНК оценок. Проверка гипотез. Асимптотика без нормальности ошибок. Нормальность ошибок.

8. Бутстрэп

Бутстрэп. Классический бутстрэп до регрессии и бутстрэп в регрессии. Метод наименьших модулей.

Чёрный трэк: возможно, разные варианты бутстрэпа в регрессии? BCA-бутстрэп до регрессии?

9. Выбор функциональной формы

Дамми-переменные и их интерпретация. Функциональные формы: полиномы, логарифмы, интерпретация коэффициентов. Информационные критерии.

Чёрный трэк: Структурные сдвиги. Тест Чоу. Локально-линейная регрессия (LOESS).

10. Гетероскедастичность

Гетероскедастичность. Тестирование гетероскедастичности. Робастные оценки. Доступный обобщённый МНК.

Задачи для доски:

Хансен: во сколько раз может быть недооценена дисперсия из-за гетероскедастичности

Коммент: акцент на робастных ошибках, тестирование и обобщённый МНК — кратко.

11. Мультиколлинеарность и метод главных компонент

Мультиколлинеарность и метод главных компонент.
Чёрный трэк: несколько взглядов на метод главных компонент? LASSO?

12. Эндогенность

Эндогенность. Инструментальные переменные. Ошибка измерения регрессора. Двухшаговый МНК.

13. Эффекты воздействия

Оценка эффектов воздействия. ATE. LATE. Четкий (sharp) и нечеткий (fuzzy) разрывный регрессионный дизайн (RDD).
Чёрный трэк: Метод разность разностей (DiD). Динамический метода разность разностей (Event Study).

14. Логистическая регрессия: точечные оценки

Логистическая регрессия: Бинарный и упорядоченный логит. Точечные оценки, прогнозы. Интерпретация предельных эффектов.
Чёрный трэк: Множественные логиты. Неупорядоченные, условные, смешанные логиты.

15. Логистическая регрессия: доверительные интервал

Логистическая регрессия: доверительные интервалы ипроверка гипотез.