22. Теорема Гаусса-Маркова: выражение $Cov\left(\tilde{a},\tilde{a}\right)$ и его обоснование.

$Cov\left(\tilde{a},\tilde{a}\right) = \tilde{\sigma}_u^2 \cdot \left(X^T \cdot X\right)^{-1} = \tilde{\sigma}_u^2 \cdot Q$

Доказательство:

$\vec{\tilde{a}} = \left(X^T P X\right)^{-1} X^T P \vec{y} = M \cdot \vec{y}$ – оценивание вектора $\vec{a}$; $\vec{\tilde{a}}$ – линейное преобразование $\vec{y}$.

$\vec{\tilde{a}} = A * \vec{y} + \vec{b}$, где $A = M$, $\vec{b} = \vec{0}$

Тогда по теореме Фишера

$$Cov\left(\tilde{a},\tilde{a}\right) = M \cdot Cov(\vec{y},\vec{y}) \cdot M^T = \left(X^T P X^{-1}\right) \cdot X^T P \cdot \left(\sigma_0^2 P^{-1}\right) X \left(X^T P X\right)^{-1} =$$
$$= \sigma_0^2 \cdot \left(X^T \cdot X\right)^{-1} = \sigma_0^2 \cdot Q \ \blacksquare$$

Более подробное доказательство смотри пункт 13

23. Теорема Гаусса-Маркова: предпосылки и свойство наименьших квадратов $\tilde{u}^T \cdot \tilde{u} \to min$.

Пусть в уравнениях наблюдений $\vec{y} = X\vec{a} + \vec{u}$:

0. Столбцы X линейно независимы;
1. $E(u_1) = \ldots = E(u_n) = 0$;
2. $Var(u_1) = \ldots = Var(u_n) = \sigma_u^2$; - не зависят от объясняющих переменных
3. $Cov(u_i,\ u_j) \neq 0, i \neq j$; - Случайные остатки попарно некоррелированные
4. $Cov(u_i,\ x_{mj}) = 0$. - Значения объясняющих переменных не коррелированы со значениями случайных возмущений

Тогда выполняются необходимые утверждения (не все, только те, которые требуются в вопросе):

А) $\vec{\tilde{a}} = \left(X^T X\right)^{-1} X^T \vec{y}$ – оптимальная линейная процедура оценивания коэффициентов функции регрессии.

С) Оценки, вычисленные в А, обладают замечательным свойством наименьших квадратов, то есть $\sum_{i=1}^{n} \tilde{u}_i^2 \to min$. Именно это свойство является причиной общепринятого названия процедуры А – МНК.

24. Теорема Гаусса-Маркова: выражение $\tilde{\sigma}_0^2$

2-я предпосылка теоремы Гаусса-Маркова о гомоскедастичности случайного остатка не выполнена, то есть дисперсия зависит от объясняющих переменных, а остаток гетероскедастичен. В таком случае оценки параметров модели утрачивают свое свойство оптимальности (свойство минимальных дисперсий). Для построения оптимальной процедуры оценивания модели с гетероскедастичным остатком потребуется модель гетероскедастичности остатка, вот простейший вид такой модели:

$$Var(u) = \sigma_u^2 = \sigma_0^2 \left(\sum_{j=0}^{k} |x_j| \right)^{\lambda}$$

σ_0^2 имеет смысл дисперсии такой случайной величины, вес которой равен 1, поэтому называется дисперсией единицы веса.

λ-некоторое априорно заданное число (подбирается эксперементально)

$p = \dfrac{\sigma_0^2}{\sigma_u^2}$ – вес случайного возмущения

Величина $\tilde{\sigma}_0^2 = \dfrac{\tilde{\vec{u}}^T P \cdot \tilde{\vec{u}}}{n-(k+1)} = \dfrac{\sum_{i=1}^{n} p_i \cdot \tilde{u}_i^2}{n-(k+1)}$ - несмещенная оценка σ_0^2

25. Взвешенный метод наименьших квадратов (ВМНК). Практическая реализация ВМНК.

В данном случае предпосылка 2 теоремы Гаусса-Маркова нарушается (случайное возмущение гетероскедастично).

Алгоритм взвешенного метода наименьших квадратов (ВМНК) состоит в предварительной трансформации ЛММР с гетероскедастичным остатком к модели с гомоскедастичным остатком, далее проверке гомоскедастичности остатка в трансформированной модели и, наконец, в применении процедуры МНК.

$\vec{y} = X\vec{a} + \vec{u}$ - уравнения наблюдений

Ковариационная матрица имеет следующий вид:

$$\Omega_{\vec{u}} = \sigma_0^2 P^{-1} = \sigma_0^2 \cdot \begin{pmatrix} \frac{1}{p_1} & 0 & \dots & 0 \\ 0 & \frac{1}{p_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{p_n} \end{pmatrix}$$

Составим уравнения наблюдений:

$$\begin{cases} \sqrt{p} \cdot y = \sqrt{p} \cdot a_0 + \sqrt{p} \cdot a_1 x_1 + \sqrt{p} \cdot a_2 x_2 + \ldots + \sqrt{p} \cdot a_k x_k + \sqrt{p} \cdot u \\ E\left(\sqrt{p} \cdot u\right) = 0; \;\; E\left(\left(\sqrt{p} \cdot u\right)^2\right) = \sigma_0^2 \end{cases}$$

$$\underbrace{P^{\frac{1}{2}} \cdot \vec{y}}_{(\vec{y}')} = \underbrace{P^{\frac{1}{2}} \cdot X}_{(X')} \cdot \vec{a} + \vec{v} \qquad (1)$$

$$P^{\frac{1}{2}}=\begin{pmatrix} \frac{1}{\sqrt{p_1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\sqrt{p_2}} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \frac{1}{\sqrt{p_n}} \end{pmatrix}$$

(1) удовлетворяют всем предпосылкам теоремы Гаусса-Маркова, следовательно, можем оценить параметры с помощью МНК.

A-D теоремы Г-М превращаются в следующие утверждения:

$$A)\ \tilde{\vec{a}} = \left(X^T \cdot P \cdot X\right)^{-1} \cdot X^T \cdot P \cdot \vec{y} = Q \cdot X^T \cdot P \cdot \vec{y}$$

$$B)\ \tilde{\sigma}_0^2 = \frac{\sum_{i=1}^{n} \tilde{v}_i^2}{n-(k+1)} = \frac{\sum_{i=1}^{n} p_i \cdot \tilde{u}_i^2}{n-(k+1)};\ \ -\text{оценка дисперсии}$$

$$\tilde{u}_i = y_i - \left(\tilde{a}_0 + \tilde{a}_1 \cdot x_{1,i} + \ldots + \tilde{a}_k \cdot x_{k,i}\right);\ \tilde{v}_i = \sqrt{p_i} \cdot \tilde{u}_i$$

$$C)\ \sum_{i=1}^{n} \tilde{v}_i^2 = \sum_{i=1}^{n} p_i \cdot \tilde{u}_i^2 \to \min$$

$$D) \begin{cases} S\tilde{a}_j = \tilde{\sigma}_0 \cdot \sqrt{q_{j+1\ j+1}} \\ j\ = 0, 1,\ \dots, k \end{cases}$$

Свойство C) оценок коэффициентов из A) принято называть *взвешанными наименьшими квадратами,* откуда A) - взвешенный метод наименьших квадратов.

26. Обобщённый метод наименьших квадратов (ОМНК).

Приступаем к доказательству утверждений теоремы Гаусса-Маркова. Мы расширим предпосылки с номерами 2 и 3 этой теоремы отказавшись от них и предполагая, что вектор случайных возмущений $\vec{u}$ имеет нулевое математическое ожидание и коварионную матрицу $Cov\ (\vec{u}, \vec{u}) = \sigma_0^2 \cdot P^{-1}$.

Докажем утверждение А:

1) Мы будем разыскивать оценку вектора $\vec{a}$ в классе всех линейных функций определённых на векторе значений эндогенной переменной $\vec{y}$, так что определению подлежит матрица M этого линейного преобразования мы собираемся разыскать M.

$$\tilde{\vec{a}} = M \cdot \vec{y}$$

2) Поиск матрицы M удобно осуществить создавая оптимальную статистическую процедуру оцениевания значения y_0 произвольной линейной функцией вектора

коэффициентов модели

$$y_0 = \vec{x}_0^T \cdot \vec{a}$$

3) Процедуру оценивания числа y_0 мы будем отыскивать в классе линейных функций $\vec{y}$, где $\vec{m}$ – это строка линейных коэффициентов.

$$\tilde{y}_0 = \vec{m}_0 \cdot \vec{y}$$

И будем отыскивать опираясь на два требования оптимальности:

$$\begin{cases} E\left(\tilde{y}_0\right) = y_0 = \vec{x}_0^T \cdot \vec{a} \\ Var\left(\tilde{y}_0\right) \to \min \end{cases}$$

Вычислим мат ожидание символа $\tilde{y}_0$. Первое требование оптимальности приводит к следующим уравению отностительно искомых коэффициентов $\vec{m}$:

$$E\left(\tilde{y}_0\right) = \vec{m} \cdot X \cdot \vec{a}, \ \Rightarrow \vec{m}^T \cdot X = \vec{a} = \vec{x}_0^T \cdot \vec{a}$$

Теперь найдём дисперсию опираясь на следующее выражение:
$Cov\ (\vec{u}, \vec{u}) = \sigma_0^2 \cdot P^{-1}$.

Значит дисперсия: $Var\left(\tilde{y}_0\right) = \sigma_0^2 \cdot \vec{m}^T \cdot P^{-1} \cdot \vec{m}$

Следовательно искомые оценки коэффициентов нужно найти в процессе решения следующей оптимизационной задачи:

$$\begin{cases} \vec{m}^T \cdot P^{-1} \cdot \vec{m} \to \min \\ X^T \cdot \vec{m} = \vec{x}_0 \end{cases} \qquad (5.11')$$

Задачу 5.11' решаем методом Лагранжа:
Шаг 1. Составим функция Лагранжа:

$$L(\vec{m}, \vec{l}) = \vec{m}^T \cdot P^{-1} \cdot \vec{m} + \vec{l}^T \cdot \left(\vec{x}_0 - X^T \cdot \vec{m}\right)$$

символом l обозначен множитель Лагранжа.
Шаг 2. Составим необходимые условия экстремума функции Лагранжа:

$$\begin{cases} \dfrac{\partial L}{\partial \vec{m}} = 2 \cdot P^{-1} \cdot \vec{m} - X \cdot \vec{l} = 0, \\ \dfrac{\partial L}{\partial l} = \vec{x}_0 - X^T \cdot \vec{m} = 0. \end{cases} \qquad (5.14)$$

Шаг 3. Система 5.14 решается аналитически или методом подстановки:

$$\vec{m} = P \cdot X \cdot \left(X^T \cdot P \cdot X\right)^{-1} \cdot \vec{x}_0 \qquad (5.15)$$

Следовательно, наилучшая оценка расчитавается по правилу 5.16:

$$\tilde{y}_0 = \vec{m}^T \cdot \vec{y} = \vec{x}_0^T \cdot \left(X^T \cdot P \cdot X\right)^{-1} \cdot \vec{y} \qquad (5.16)$$

И последнее дейстие

$$y_0 = \vec{x}_0^T \cdot \vec{a} \Rightarrow \vec{\tilde{a}} = \left(\left(X^T \cdot P \cdot X\right)^{-1} \cdot X^T\right)(= M) \cdot P \cdot \vec{y} \blacksquare$$

Это и есть ОМНК.

27. Система нормальных уравнений и явный вид её решения при оценивании методом наименьших квадратов (МНК) линейной модели парной регрессии.

Из $\vec{\tilde{a}} = \left(X^T P X\right)^{-1} X^T P \vec{y}$ видно, что $\vec{\tilde{a}}$ вычисляется в процессе решения системы из $k+1$ линейных алгебраических уравнений с $k+1$ неизвестными: $\left(X^T P X\right)\vec{\tilde{a}} = X^T P \vec{y}$.

Эта система называется системой нормальных уравнений.

В ситуации процедуры МНК, т.е. $P = E$ её подробная запись принимает следующий вид:

$$\begin{cases} n\tilde{a}_0 + [x]\tilde{a}_1 = [y] \\ [x]\tilde{a}_0 + \left[x^2\right]\tilde{a}_1 = [xy] \end{cases}$$

где $[x] = \sum_{i=1}^{n} x_i;\ [y] = \sum_{i=1}^{n} y_i;\ \left[x^2\right] = \sum_{i=1}^{n} x_i^2\ ;\ [xy] = \sum_{i=1}^{n} x_i y_i$

Тогда явный вид решения этой системы:

$$\left\{ \tilde{a}_1 = \frac{Cov(x,y)}{Var(x)};\quad \tilde{a}_0 = \overline{y} - \overline{x}\tilde{a}_1 \right.$$

28. Ковариационная матрица оценок коэффициентов линейной модели парной регрессии: явные выражения $Var\left(\tilde{a}_0\right),\ Var\left(\tilde{a}_1\right),\ Cov\left(\tilde{a}_0,\ \tilde{a}_1\right)$.

Линейная модель парной регрессии имеет вид:

$$\begin{cases} y = a_0 + a_1 x + u \\ E(u) = 0,\ Var(u) = \sigma_u^2 \end{cases}$$

$$Cov\left(\vec{\tilde{a}},\ \vec{\tilde{a}}\right) = \begin{pmatrix} Var\left(\tilde{a}_0\right) & Cov\left(\tilde{a}_0,\ \tilde{a}_1\right) \\ Cov\left(\tilde{a}_0,\ \tilde{a}_1\right) & Var\left(\tilde{a}_1\right) \end{pmatrix}$$

$$Cov\left(\vec{\tilde{a}},\ \vec{\tilde{a}}\right) = \sigma_u^2 Q = \sigma_u^2 \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix}$$

Учитывая, что $R = X^T X = \begin{pmatrix} n & [x] \\ [x] & \left[x^2\right] \end{pmatrix}$ - матрица, образованная коэффициентами и свободными членами системы $\begin{cases} n\tilde{a}_0 + [x]\tilde{a}_1 = [y] \\ [x]\tilde{a}_0 + \left[x^2\right]\tilde{a}_1 = [xy] \end{cases}$

Получаем:

$$Var\left(\tilde{a}_0\right) = \sigma^2 Q_{11} = \sigma^2\left(\frac{1}{n} + \frac{\overline{x}^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}\right)$$

$$Var\left(\tilde{a}_1\right) = \sigma^2 Q_{22} = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

$$Cov\left(\tilde{a}_0,\ \tilde{a}_1\right) = \sigma^2 Q_{12} = \sigma^2 Q_{21} = \frac{-\sigma^2\overline{x}}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$