

## Семинар 11

---

24 ноября 2020 г.

### Задача 1. Теорема Гаусса-Маркова для стохастических регрессоров.

Если:

1. Модель задана как  $y = X\beta + u$ , и при помощи МНК оценивается линейная регрессия  $\hat{y} = X\hat{\beta}$ .
2.  $\beta$  – вектор констант.
3.  $\mathbb{P}(\{\text{у матрицы } X \text{ есть линейно зависимые столбцы}\}) = 0$ .
4.  $\mathbb{E}(u|X) = 0$ ,  $\text{Var}(u|X) = \sigma^2 I$ ,

то

1.  $\hat{\beta}$  существуют и единственны с вероятностью один.
  2.  $\hat{\beta}$  линейны по  $y$ .
  3.  $\mathbb{E}(\hat{\beta}|X) = \beta$ .
  4.  $\hat{\beta}$  эффективны в классе линейных по  $y$  и условно несмещённых оценок.
- а) Покажите, что если предпосылки ТГМ со стохастическими регрессорами выполнены, то оценки  $\hat{\beta}$  являются состоятельными.
- б) При нарушении какого условия ТГМ оценки МНК перестанут быть состоятельными?

### Задача 2. Причины эндогенности.

- а) Рассмотрим модель  $y_i = \beta x_i + u_i$  со стохастическими регрессорами, в которой выполнены все предпосылки ТГМ, кроме  $\mathbb{E}(u_i|X) \neq 0$  по какой-то причине. Покажите, что если оценивается регрессия  $\hat{y}_i = \hat{\beta}x_i$  при помощи МНК, то  $\hat{\beta}$  не являются состоятельными.
- б) (*Пропущенные переменные*) Рассмотрим модель  $y_i = \beta x_i + \gamma z_i + u_i$  со стохастическими переменными, в которой выполнены все предпосылки ТГМ. Исследователь Дорофей оценивает регрессию  $\hat{y}_i = \hat{\beta}x_i$  при помощи МНК. Покажите, что оценки в регрессии Дорофея не являются состоятельными.
- в) (*Ошибки измерения*) Отчаянный исследователь Афанасий пытается предсказать результаты контрольной по статистике при помощи модели со стохастическими регрессорами  $y_i = \beta x_i + u_i$ , где  $x_i$  – количество съеденных бургеров известной фирмы. Для сбора данных производится опрос студентов. Афанасий догадывается, что полученные данные отражают не настоящие  $x_i$ , а строятся следующим образом:

$$x_i^* = x_i + \alpha + v_i,$$

где  $\alpha$  – показатель русской народной скромности, а  $v_i$  – случайная величина, отражающая несовершенство человеческой памяти. Предположим, что  $\alpha$  – константа, а  $\mathbb{E}(x_i) = \mu_x$ ,  $\mathbb{E}(u_i) = 0$ ,  $\mathbb{E}(v_i) = 0$ ,  $\text{Var}(x_i) = \sigma_x^2$ ,  $\text{Var}(u_i) = \sigma_u^2$ ,  $\text{Var}(v_i) = \sigma_v^2$ . Предположим, что  $x_i$  не коррелирует с  $v_i$  и  $u_i$ , а все  $u_i$  и  $v_i$  не коррелируют друг с другом и между собой.

Афанасий оценивает регрессию  $\hat{y}_i = \hat{\beta} x_i^*$  при помощи МНК по собранным данным. Покажите, что  $\hat{\beta}$  не является состоятельной. Как смещение зависит от  $\alpha$ ?

### Задача 3. Двухшаговый МНК.

Переменные  $z_i$  называются инструментальными, если

1. (релевантность)  $\mathbb{E}(x_i z_i) \neq 0 \sim \text{Cov}(x_i, z_i) \neq 0 \sim \text{plim}_{n \rightarrow \infty} Z_j' X \neq 0$ .
2. (валидность)  $\mathbb{E}(u_i z_i) = 0 \sim \text{Cov}(u_i, z_i) = 0 \sim \text{plim}_{n \rightarrow \infty} Z_j' u = 0$ .

Если подобрали подходящие инструменты, то можно применить двухшаговый МНК:

**Шаг 1.** Оценить регрессию  $X$  на  $Z$ , получить прогнозы  $\hat{X}$ .

**Шаг 2.** Оценить регрессию  $y$  на  $\hat{X}$ , получить оценки  $\hat{\beta}_{2SLS}$ .

- a) Выведите оценки двухшагового МНК в общем случае ( $Z$  имеет размеры  $n \times m$ ,  $m \geq k$ ).
- b) Выведите оценки метода инструментальных переменных ( $Z$  имеет размеры  $n \times m$ ,  $m = k$ ).
- c) Покажите, что оценки метода инструментальных переменных являются состоятельными.