# Эконометрика: доказательства

*Борис Демешев*

*2019-03-12*

# Оглавление

# Глава 1

# Введение

1. Метод главных компонент и кластеризация
2. МНК без статистических предпосылок
3. МНК и дисперсия
4. МНК и нормальные ошибки
5. МНК и большие выборки
6. Гетероскедастичность
7. Эндогенность
8. Метод максимального правдоподобия
9. Логит и пробит
10. Деревья и леса
11. Временные ряды
12. Немного панельных данных

## Глава 2

# Метод главных компонент

# Глава 3

# МНК без статистических предпосылок

Определение.

Кросс-валидация с поочередным выкидыванием отдельных наблюдений.

Leave one out cross validation.

Рассмотрим модель $y = X\beta + u$.

Оценим модель без первого наблюдения. Получим МНК-оценки $\hat{\beta}^{(-1)}$. С помощью этих оценок спрогнозируем первое наблюдение, получим прогноз $\hat{y}_1^{CV}$ и ошибку прогноза $\hat{u}_1^{CV}$.

Вернём первое наблюдение в выборку и удалим второе наблюдение. Получим МНК-оценки $\hat{\beta}^{(-2)}$. С помощью этих оценок спрогнозируем второе наблюдение, получим прогноз $\hat{y}_2^{CV}$ и ошибку прогноза $\hat{u}_2^{CV}$.

Поступим так с каждым наблюдением. На выходе получим вектор кросс-валидационных прогнозов $\hat{y}^{CV}$ и вектор кросс-валидационных ошибок прогнозов $\hat{u}^{CV}$.

Теорема.

Если модель $y = X\beta + u$ оценивается с помощью МНК и проводится кросс-валидации с поочередным выкидыванием отдельных наблюдений, то:

$$\hat{u}_i = (1 - H_{ii}) \cdot \hat{u}_i^{CV},$$

где $H$ — матрица-шляпница $H = X(X'X)^{-1}X'$, $\hat{u}$ — остатки регрессии, а $\hat{u}^{CV}$ — кросс-валидационные ошибки прогнозов.

Доказательство.

Оценим модель без последнего наблюдения, $\hat{y}^- = X^-\hat{\beta}^-$.

Создадим вектор $y^*$, который будет отличаться от $y$ только последним, $n$-м элементом: вместо настоящего $y_n$ там будет стоять прогноз по модели без последнего наблюдения $\hat{y}_n^-$.

Раз уж мы добавили новую точку лежащую ровно на выборочной регрессии, то при оценки модели $\hat{y}^* = X\hat{\beta}^*$ мы получим в точности старые оценки $\hat{\beta}^* = \hat{\beta}^-$. Следовательно, и прогнозы эти две модели дают одинаковые, $\hat{y}_i^* = \hat{y}_i^-$.

А теперь посмотрим на последний элемент вектора $v = H(y^* - y)$.

С одной стороны, он равен последней строке матрицы $H$ умножить на вектор $(y^* - y)$. В векторе $(y^* - y)$ только последний элемент ненулевой, поэтому $v_n = H_{nn}(\hat{y}_n^- - y_n)$.

С другой стороны, мы можем раскрыть скобки, и заметить, что $v = Hy^* - Hy$. И окажется, что $v_n = \hat{y}_n^* - \hat{y}_n = \hat{y}_n^- - \hat{y}_n$.

Отсюда

$$\hat{y}_n^- - \hat{y}_n = H_{nn}(\hat{y}_n^- - y_n)$$

Приводим подобные слагаемые и добавляем слева и справа $y_n$, получаем как раз то, что нужно:

$$y_n - \hat{y}_n = (1 - H_{ii})(y_n - \hat{y}_n^-)$$

# Глава 4

# МНК и дисперсия

Теорема Гаусса-Маркова

Модель $y = X\beta + u$ оценивается с помощью МНК.

TODO: дописать оставшуюся часть теоремы и доказательства

Доказательство эффективности МНК-оценок.

Эффективность МНК-оценок — это реинкарнация теоремы Пифагора. Мы докажем, что дисперсия МНК-оценки — это квадрат длины катета, дисперсия альтернативной несмещённой оценки — квадрат длины гипотенузы.

Для примера рассмотрим первый коэффициент бета. Доказательство не меняется ни капли, если рассмотреть произвольную линейную комбинацию коэффициентов бета. У нас есть две оценки, $\hat{\beta}_1$ и $\hat{\beta}_1^{alt}$. Обе они линейны по $y$, следовательно, $\hat{\beta}_1 = a'y$ и $\hat{\beta}_1^{alt} = a'_{alt}y$.

Замечаем, что $Var(\hat{\beta}_1) = \sigma^2 a'a$, и $Var(\hat{\beta}_1^{alt}) = \sigma^2 a'_{alt}a_{alt}$. То есть дисперсии пропорциональны квадратм длин векторов $a$ и $a^{alt}$. Осталось доказать, что вектор $a$ не длиннее вектора $a^{alt}$ :)

Для этого мы докажем, что $a^{alt}$ — это гипотенуза, а $a$ — катет. То есть нужно доказать, что вектор $a - a^{alt}$ перпендикулярен вектору $a$.

Разобъём доказательство перпендикулярности $a$ и $a - a^{alt}$ на два шага:

Шаг 1. Вектор $a - a^{alt}$ перпендикулярен любому столбцу матрицы $X$.

Шаг 2. Вектор $a$ является линейной комбинацией столбцов матрицы $X$.

TODO: здесь картинка!

Приступаем к шагу 1. Обе оценки несмещённые, поэтому для любых $\beta$ должно выполняться:

$$E(\hat{\beta}_1) = E(\hat{\beta}_1^{alt})$$

Переносим всё в левую сторону:

$$E((a' - a'_{alt})(X\beta + u)) = 0$$

Получаем, что для любых $\beta$ должно быть выполнено условие:

$$(a - a_{alt})' X\beta = 0$$

Но это возможно только если вектор $(a - a_{alt})'X$ равен нулю. Следовательно, вектор $(a - a_{alt})$ перпендикулярен любому столбцу $X$.

Приступаем к шагу 2.

Вспоминаем, что $\hat{\beta} = (X'X)^{-1}X'y$. Следовательно, нужная строка весов $a'$ — это первая строка в матрице $(X'X)^{-1}X'$. Замечаем, что выражение имеет вид $A \cdot X'$.

Вспоминаем из линейной алгебры, что при умножении матриц $AB$ получается матрица $C$, на которую можно взглянуть двумя способами. Можно считать, что $C$ — это разные линейные комбинации столбцов левой матрицы $A$, а можно считать, что $C$ — это разные линейные комбинаций строк правой матрицы $B$.

Применим второй взгляд :) Получаем, что строка $a'$ — линейная комбинация строк матрицы $X'$. Или, другими словами, столбец $a$ — линейная комбинация столбцов матрицы $X$.

# Глава 5

# Обозначения

$n$ — количество наблюдений

$k$ — количество коэффициентов бета

$y$ — вектор зависимой переменной размера $(n \times 1)$

$\beta$ — вектор истинных значений коэффициентов размера $(k \times 1)$

$\hat{u}$ — остатки модели;

$\hat{u}^{CV}$ — ошибки прогнозов, полученных с помощью кросс-валидации с поочередным выкидыванием отдельных наблюдений;

$H$ — матрица-шляпница, $H = X(X'X)^{-1}X'$