

# 1 МНК — это...

Минитеория:

1. Истинная модель. Например,  $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + u_i$ .
2. Формула для прогнозов. Например,  $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i + \hat{\beta}_3 z_i$ .
3. Метод наименьших квадратов,  $\sum (y_i - \hat{y}_i)^2 \rightarrow \min$ .

Задачи:

**1.1** Каждый день Маша ест конфеты и решает задачи по эконометрике. Пусть  $x_i$  — количество решённых задач, а  $y_i$  — количество съеденных конфет.

$x_i$	$y_i$
1	1
2	2
2	4

1. Рассмотрим модель  $y_i = \beta x_i + u_i$ :
  - (a) Найдите МНК-оценку  $\hat{\beta}$  для имеющихся трёх наблюдений.
  - (b) Нарисуйте исходные точки и полученную прямую регрессии.
  - (c) Выведите формулу для  $\hat{\beta}$  в общем виде для  $n$  наблюдений.
2. Рассмотрим модель  $y_i = \beta_1 + \beta_2 x_i + u_i$ :
  - (a) Найдите МНК-оценки  $\hat{\beta}_1$  и  $\hat{\beta}_2$  для имеющихся трёх наблюдений.
  - (b) Нарисуйте исходные точки и полученную прямую регрессии.
  - (c) Выведите формулу для  $\hat{\beta}_2$  в общем виде для  $n$  наблюдений.

**1.2** Упростите выражения:

1.  $n\bar{x} - \sum x_i$
2.  $\sum (x_i - \bar{x})\bar{x}$
3.  $\sum (x_i - \bar{x})\bar{z}$
4.  $\sum (x_i - \bar{x})^2 + n\bar{x}^2$

**1.3** При помощи метода наименьших квадратов найдите оценку неизвестного параметра  $\theta$  в следующих моделях:

1.  $y_i = \theta + \theta x_i + \varepsilon_i$ ;
2.  $y_i = 1 + \theta x_i + \varepsilon_i$ ;
3.  $y_i = \theta/x_i + \varepsilon_i$ ;
4.  $y_i = \theta x_i + (1 - \theta)z_i + \varepsilon_i$ .

**1.4** Найдите МНК-оценки параметров  $\alpha$  и  $\beta$  в модели  $y_i = \alpha + \beta y_i + \varepsilon_i$ .

**1.5** Рассмотрите модели  $y_i = \alpha + \beta(y_i + z_i) + \varepsilon_i$ ,  $z_i = \gamma + \delta(y_i + z_i) + \varepsilon_i$ .

1. Как связаны между собой  $\hat{\alpha}$  и  $\hat{\gamma}$ ?
2. Как связаны между собой  $\hat{\beta}$  и  $\hat{\delta}$ ?

**1.6** Как связаны МНК-оценки параметров  $\alpha, \beta$  и  $\gamma, \delta$  в моделях  $y_i = \alpha + \beta x_i + \varepsilon_i$  и  $z_i = \gamma + \delta x_i + v_i$ , если  $z_i = 2y_i$ ?

**1.7** Для модели  $y_i = \beta_1 x_i + \beta_2 z_i + \varepsilon_i$  решите условную задачу о наименьших квадратах:

$$Q(\beta_1, \beta_2) := \sum_{i=1}^n (y_i - \hat{\beta}_1 x_i - \hat{\beta}_2 z_i)^2 \rightarrow \min_{\beta_1 + \beta_2 = 1}.$$

**1.8** 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.

**1.9** 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.

**1.10** Есть двести наблюдений. Вовочка оценил модель  $\hat{y} = \hat{\beta}_1 + \hat{\beta}_2 x$  по первой сотне наблюдений. Петечка оценил модель  $\hat{y} = \hat{\gamma}_1 + \hat{\gamma}_2 x$  по второй сотне наблюдений. Машенька оценила модель  $\hat{y} = \hat{m}_1 + \hat{m}_2 x$  по всем наблюдениям.

1. Возможно ли, что  $\hat{\beta}_2 > 0$ ,  $\hat{\gamma}_2 > 0$ , но  $\hat{m}_2 < 0$ ?
2. Возможно ли, что  $\hat{\beta}_1 > 0$ ,  $\hat{\gamma}_1 > 0$ , но  $\hat{m}_1 < 0$ ?
3. Возможно ли одновременное выполнение всех упомянутых условий?

**1.11** У эконометриста Вовочки есть переменная  $1_f$ , которая равна 1, если  $i$ -ый человек в выборке — женщина, и 0, если мужчина. Есть переменная  $1_m$ , которая равна 1, если  $i$ -ый человек в выборке — мужчина, и 0, если женщина. Вовочка попробовал оценить 4 регрессии.

1.  $y$  на константу и  $1_f$ ;
  2.  $y$  на константу и  $1_m$ ;
  3.  $y$  на  $1_f$  и  $1_m$  без константы;
  4.  $y$  на константу,  $1_f$  и  $1_m$ .
1. Какой смысл будут иметь оцениваемые коэффициенты?
  2. Как связаны между собой оценки коэффициентов этих регрессий?

**1.12** Эконометрист Вовочка оценил методом наименьших квадратов модель 1,  $y = \beta_1 + \beta_2 x + \beta_3 z + \varepsilon$ , а затем модель 2,  $y = \beta_1 + \beta_2 x + \beta_3 z + \beta_4 w + \varepsilon$ . Сравните полученные  $ESS$ ,  $RSS$ ,  $TSS$  и  $R^2$ .

**1.13** Что происходит с  $TSS$ ,  $RSS$ ,  $ESS$ ,  $R^2$  при добавлении нового наблюдения? Если величина может изменяться только в одну сторону, то докажите это. Если возможны и рост, и падение, то приведите пример.

**1.14** Эконометресса Аглая подглядела, что у эконометрессы Жозефины получился  $R^2$  равный 0.99 по 300 наблюдениям. От чёрной зависти Аглая не может ни есть, ни спать.

1. Аглая добавила в набор данных Жозефины ещё 300 наблюдений с такими же регрессорами, но противоположными по знаку играками, чем были у Жозефины. Как изменится  $R^2$ ?
2. Жозефина заметила, что Аглая добавила 300 наблюдений и вычеркнула их, вернув в набор данных в исходное состояние. Хитрая Аглая решила тогда добавить всего одно наблюдение так, чтобы  $R^2$  упал до нуля. Удается ли ей это сделать?

**1.15** На работе Феофан построил парную регрессию по трём наблюдениям и посчитал прогнозы  $\hat{y}_i$ . Придя домой он отчасти вспомнил результаты:

$y_i$	$\hat{y}_i$
0	1
6	?
6	?

Поднапрягшись, Феофан вспомнил, что третий прогноз был больше второго. Помогите Феофану восстановить пропущенные значения.

## 2 МНК в матрицах!

Минитеория.

Дифференциал для матриц подчиняется правилам:

1.  $da = 0, dA = 0$ ;
2.  $d(RS) = dR\dot{S} + R \cdot dS$

**2.16** Пусть  $t$  — скалярная переменная,  $r, s$  — векторные переменные,  $R, S$  — матричные переменные. Кроме того,  $a, b$  — векторы констант,  $A, B$  — матрицы констант.

Применив базовые правила дифференцирования найдите:

1.  $d(ARB)$ ;
2.  $d(r'r)$ ;
3.  $d(r'Ar)$ ;
4.  $d(R^{-1})$ , воспользовавшись тем, что  $R^{-1} \cdot R = I$ ;
5.  $d \cos(r'r)$ ;
6.  $d(r'Ar/r'r)$ .

**2.17** В методе наименьших квадратов минимизируется функция

$$Q(\hat{\beta}) = (y - X\hat{\beta})'(y - X\hat{\beta}).$$

1. Найдите  $dQ(\hat{\beta})$  и  $d^2Q(\hat{\beta})$ ;
2. Выпишите условия первого порядка для задачи МНК;
3. Выразите  $\hat{\beta}$  предполагая, что  $X'X$  обратима.

**2.18** В методе LASSO минимизируется функция

$$Q(\hat{\beta}) = (y - X\hat{\beta})'(y - X\hat{\beta}) + \lambda hb'hb,$$

где  $\lambda$  — положительный параметр, штрафующий функцию за слишком большие значения  $\hat{\beta}$ .

1. Найдите  $dQ$ ;
2. Выпишите условия первого порядка для задачи LASSO;
3. Выразите  $\hat{\beta}$ .

**2.19**

## 3 Распределение

**3.20** Компоненты вектора  $x = (x_1, x_2)'$  независимы и имеют стандартное нормальное распределение. Вектор  $y$  задан формулой  $y = (2x_1 + x_2 + 2, x_1 - x_2 - 1)$ .

1. Выпишите совместную функцию плотности вектора  $x$ ;
2. Нарисуйте на плоскости линии уровня функции плотности вектора  $x$ ;
3. Выпишите совместную функцию плотности вектора  $y$ ;
4. Найдите собственные векторы и собственные числа ковариационной матрицы вектора  $y$ ;
5. Нарисуйте на плоскости линии уровня функции плотности вектора  $y$ .

**3.21** Компоненты вектора  $x = (x_1, x_2, x_3)'$  независимы и имеют стандартное нормальное распределение.

1. Как выглядят в пространстве поверхности уровня совместной функции плотности?
2. Рассмотрим три апельсина с кожурой одинаковой очень маленькой толщины: бэби-апельсин радиуса 0.1, стандартный апельсин радиуса 1 и гранд-апельсин радиуса 10. В кожуру какого апельсина вектор  $x$  попадает с наибольшей вероятностью?
3. Мы спроецируем случайный вектор на  $x$  на плоскость  $2x_1 + 3x_2 - 7x_3 = 0$ . Какое распределение имеет квадрат длины проекции?
4. Введём вектор  $y$  независимый от  $x$  и имеющий такое же распределение. Спроецируем вектор  $x$  на плоскость проходящую через начало координат и перпендикулярную вектору  $y$ . Какое распределение имеет квадрат длины проекции?

## 4 Решения

1.1.

1.2. Ответы: 0, 0, 0,  $\sum x_i^2$ .

1.3.

$$1. \hat{\theta} = \sum ((y_i - z_i)(x_i - z_i)) / \sum (x_i - z_i)^2$$

1.4.  $\hat{\alpha} = 0, \hat{\beta} = 1$

1.5. Рассмотрим регрессию суммы  $(y_i + z_i)$  на саму себя. Естественно, в ней

$$\widehat{y_i + z_i} = 0 + 1 \cdot (y_i + z_i).$$

Отсюда получаем, что  $\hat{\alpha} + \hat{\gamma} = 0$  и  $\hat{\beta} + \hat{\delta} = 1$ .

1.6.

Исходя из условия, нужно оценить методом МНК коэффициенты двух следующих моделей:

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

$$y_i = \frac{\gamma}{2} + \frac{\delta}{2} x_i + \frac{1}{2} v_i$$

Заметим, что на минимизацию суммы квадратов остатков коэффициент  $1/2$  не влияет, следовательно:

$$\hat{\gamma} = 2\hat{\alpha}, \hat{\delta} = 2\hat{\beta}$$

1.7. Выпишем задачу:

$$\begin{cases} RSS = \sum_{i=1}^n (y_i - \hat{\beta}_1 x_i - \hat{\beta}_2 z_i)^2 \rightarrow \min_{\hat{\beta}_1, \hat{\beta}_2} \\ \hat{\beta}_1 + \hat{\beta}_2 = 1 \end{cases}$$

Можем превратить ее в задачу минимизации функции одного аргумента:

$$RSS = \sum_{i=1}^n (y_i - x_i - \hat{\beta}_2 (z_i - x_i))^2 \rightarrow \min_{\hat{\beta}_2}$$

Выпишем условия первого порядка:

$$\frac{\partial RSS}{\partial \hat{\beta}_2} = \sum_{i=1}^n 2(y_i - x_i - \hat{\beta}_2(z_i - x_i))(x_i - z_i) = 0$$

Отсюда:

$$\sum_{i=1}^n (y_i - x_i)(x_i - z_i) + \hat{\beta}_2 \sum_{i=1}^n (z_i - x_i)^2 = 0 \Rightarrow \hat{\beta}_2 = \frac{\sum_{i=1}^n (y_i - x_i)(z_i - x_i)}{\sum_{i=1}^n (z_i - x_i)^2}$$

А  $\hat{\beta}_1$  найдется из соотношения  $\hat{\beta}_1 + \hat{\beta}_2 = 1$ .

**1.8.** Обозначив вес первого слитка за  $\beta_1$ , вес второго слитка за  $\beta_2$ , а показания весов за  $y_i$ , получим, что

$$y_1 = \beta_1 + \varepsilon_1, \quad y_2 = \beta_2 + \varepsilon_2, \quad y_3 = \beta_1 + \beta_2 + \varepsilon_3$$

Тогда

$$(300 - \beta_1)^2 + (200 - \beta_2)^2 + (400 - \beta_1 - \beta_2)^2 \rightarrow \min_{\beta_1, \beta_2}$$

$$\hat{\beta}_1 = \frac{800}{3}, \quad \hat{\beta}_2 = \frac{500}{3}$$

**1.9.** Можем воспользоваться готовой формулой для регрессии на константу:

$$\hat{\beta} = \bar{y} = \frac{10 + 10 + 3}{3} = \frac{23}{3}$$

(можно решить задачу  $2(10 - \beta)^2 + (3 - \beta)^2 \rightarrow \min$ )

**1.10.**

**1.11.**

**1.12.**

**1.13.** Пусть  $\bar{y}$  — средний  $y$  до добавления нового наблюдения,  $\bar{y}'$  — после добавления нового наблюдения. Будем считать, что изначально было  $n$  наблюдений. Заметим, что

$$\bar{y}' = \frac{(y_1 + \dots + y_n) + y_{n+1}}{n+1} = \frac{n\bar{y} + y_{n+1}}{n+1} = \frac{n}{n+1}\bar{y} + \frac{1}{n+1}y_{n+1}$$

Покажем, что  $TSS$  может только увеличиться при добавлении нового наблюдения (остается неизменным при  $y_{n+1} = \bar{y}$ ):

$$\begin{aligned} TSS' &= \sum_{i=1}^{n+1} (y_i - \bar{y}')^2 = \sum_{i=1}^n (y_i - \bar{y} + \bar{y} - \bar{y}')^2 + (y_{n+1} - \bar{y}')^2 = \\ &= \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - \bar{y}')^2 + (y_{n+1} - \bar{y}')^2 = TSS + \frac{n}{n+1}(y_{n+1} - \bar{y})^2 \end{aligned}$$

Следовательно,  $TSS' \geq TSS$ .

Также сумма  $RSS$  может только вырасти или остаться постоянной при добавлении нового наблюдения. Действительно, новое  $(n+1)$ -ое слагаемое в сумме неотрицательно. А сумма  $n$  слагаемых минимальна при старых коэффициентах, а не при новых.

$ESS$  и  $R^2$  могут меняться в обе стороны. Например, рассмотрим ситуацию, где точки лежат симметрично относительно некоторой горизонтальной прямой. При этом  $ESS = 0$ . Добавим наблюдение —  $ESS$  вырастет, удалим наблюдение —  $ESS$  вырастет.

#### 1.14.

1.  $R^2$  упал до нуля.
2. Да, можно. Если добавить точку далеко слева внизу от исходного набора данных, то наклон линии регрессии будет положительный. Если далеко справа внизу, то отрицательный. Будем двигать точку так, чтобы поймать нулевой наклон прямой. Получим  $ESS = 0$ .

**1.15.** На две неизвестных  $a$  и  $b$  нужно два уравнения. Эти два уравнения — ортогональность вектора остатков плоскости регрессоров. А именно:

$$\begin{cases} \sum_i (y_i - \hat{y}_i) = 0 \\ \sum_i (y_i - \hat{y}_i) \hat{y}_i = 0 \end{cases}$$

В нашем случае

$$\begin{cases} -1 + (6 - a) + (6 - b) = 0 \\ -1 + (6 - a)a + (6 - b)b = 0 \end{cases}$$

Решаем квадратное уравнение и получаем два решения:  $a = 4$  и  $a = 7$ . Итого:  $a = 4$ ,  $b = 7$ .

#### 2.16.

#### 2.17.

#### 2.18.

#### 2.19.

#### 3.20.

**3.21.** Сферы с центром в начале координат. Проекция имеет хи-квадрат распределение с тремя степенями свободы. Для нахождения максимальной вероятности максимизируем функцию

$$\exp(-R^2/2) \cdot ((R+t)^3 - R^3) \rightarrow \max_R$$

, где  $R$  — радиус мякоти, а  $t$  — толщина кожуры апельсина. Оставляем только линейную часть по  $t$  и затем максимизируем.

Наибольшая вероятность попасть в апельсин радиуса  $R = 1$ .

## 5 Источники мудрости