handout_1 2 3

Содержание

Лекция 1. Метод наименьших квадратов без статистики

Лекция 2.

Лекция 3.

Геометрическая иллюстрация МНК

Тут картинка.

Вывод формулы для оценок коэффициентов

Если вектор y перпендикулярен вектору x, то их скалярное произведение должно быть равно нулю, т.к. $cos(90°) = 0$, а скалярное произведение равно:

$$(x, y) = |x| \cdot |y| \cdot cos(x, y)$$

С другой стороны, скалярное произведение равно

$$(x, y) = x_1y_1 + x_2y_2 + \ldots + x_ny_n = \sum_{i=1}^{n} x_iy_i = x'y$$

Значит условие перпендикулярности векторов x и y можно кратко записать как $x'y = 0$. Столбца матрицы регрессоров X ортогональны остаткам регрессии, вектору $\hat{\varepsilon}$:

$$X'\hat{\varepsilon} = 0$$

Заметим, что здесь 0 — это вектор размера $k \times 1$. Подставляем формулу для остатков, $\hat{\varepsilon} = y - X'\hat{\beta}$:

$$X'(y - X\hat{\beta}) = 0$$

Раскрываем скобки и переносим в разные стороны уравнения:

$$X'y - X'X\hat{\beta} = 0$$

$$X'X\hat{\beta} = X'y$$

Матрица X' имеет размер $k \times n$, поэтому на неё сокращать нельзя. Хотя иногда хочется :) А вот обратная матрица к матрице $X'X$ существует, если среди столбцов X нет линейно зависимых и $n \geq k$. Домножаем обе части уравнения слева на $(X'X)^{-1}$:

$$\hat{\beta} = (X'X)^{-1}X'y$$

Ура! Мы получили формулу для МНК-оценок множественной регрессии! Заметьте, что она подозрительно похожа на формулу МНК-оценки для случая одного оцениваемого параметра. В модели $y_i = \beta x_i + \varepsilon_i$ МНК-оценка коэффициента β имела вид $\hat{\beta} = \sum x_iy_i / \sum x_i^2$.

Матрица-шляпница

Если $\hat{\beta} = (X'X)^{-1}X'y$, то вектор прогнозов, $\hat{y}$, будет равен $\hat{y} = X\hat{\beta} = X(X'X)^{-1}X'y$. Матрицу $H = X(X'X)^{-1}X'$ по-английски называют "hat-matrix", матрицей-шляпницей, потому, что она надевает на y шляпку: $\hat{y} = H{\cdot}y$. Умножение любого вектора на матрицу H проецирует этот вектор на пространство, порождаемое регрессорами. Поскольку сами регрессоры уже лежат в этом пространстве, то $H \cdot X = X$. Матрица H идемпотентная, то есть возведенная в произвольную натуральную степень даст саму себя, $H^n = H$. В этом легко можно убедиться либо перемножив руками H на H,

$$H \cdot H = X(X'X)^{-1}X'X(X'X)^{-1}X' = X(X'X)^{-1}X' = H$$

либо из геометрических соображений. Умножение на H несколько раз подряд, это проецирование результата проецирования. А проекция от проекции совпадает с проекцией.

Собственными числами матрицы H могут быть только нули или единицы. Действительно, при проецировании часть векторов сохраняются (те, что лежали в пространстве регрессоров), часть превращается в ноль (те, что были ортогональны пространству регрессоров), а все другие при проецировании меняют направление.

Ранг матрицы-шляпницы можно посчитать, воспользовавшись тем, что $rk(AB) = rk(BA)$:

$$rk(X(X'X)^{-1}X') = rk(X'X(X'X)^{-1}) = rk(I_{k\times k}) = k$$

Математическое ожидание и ковариационная матрица

Если $y = (y_1, \ldots, y_n)'$ — случайный вектор, то для него определены математическое ожидание, $E(y)$, и ковариационная матрица, $Var(y)$.

Определение. Если y — вектор-столбец случайных величин,

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \text{ то } E(y) = \begin{pmatrix} E(y_1) \\ E(y_2) \\ \vdots \\ E(y_n) \end{pmatrix}$$

Определение. $Var(y) = E(yy') - E(y)E(y')$.

Согласно определению ковариационной матрицы:

$$Var(y) = \begin{pmatrix} Var(y_1) & Cov(y_1, y_2) & \cdots & Cov(y_1, y_n) \\ Cov(y_2, y_1) & Var(y_2) & \cdots & Cov(y_2, y_n) \\ Cov(y_3, y_1) & Cov(y_3, y_2) & \cdots & Cov(y_3, y_n) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(y_n, y_1) & Cov(y_n, y_2) & \cdots & Var(y_n) \end{pmatrix}$$

Определение. $Cov(y, z) = E(yz') - E(y)E(z')$

Свойства:

Если A и B — неслучайные матрицы, a и b — неслучайные вектора, y и w — случайные вектора подходящих размеров, все математические ожидания и ковариационные матрицы существуют, то:

$E(a) = a$

$E(Ay + b) = AE(y) + b$ и $E(yA + b) = E(y)A + b$

$E(y + w) = E(y) + E(w)$

$Var(a) = 0$

$Var(Ay + b) = AVar(y)A'$

$Var(y + z) = Var(y) + Var(z) + Cov(y, z) + Cov(z, y)$

$Cov(Ay + a, Bz + b) = ACov(y, z)B'$

$Cov(y, z) = Cov(z, y)'$

$Cov(y + z, w) = Cov(y, w) + Cov(z, w)$ и $Cov(y, z + w) = Cov(y, z) + Cov(y, w)$

$Cov(y, y) = Var(y)$

Статистические свойства МНК оценок

Если:

1. Истинная зависимость имеет вид $y_i = \beta_1 + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i$
   В матричном виде: $y = X\beta + \varepsilon$
2. С помощью МНК оценивается регрессия y на константу, $x_{.2}$, $x_{.3}$, $\ldots$, $x_{.k}$
   В матричном виде: $\hat{\beta} = (X'X)^{-1}X'y$
3. Наблюдений больше, чем оцениваемых коэффициентов β: $n > k$
4. Строгая экзогенность: $E(\varepsilon_i |$ все $x_{ij}) = 0$
   В матричном виде: $E(\varepsilon_i|X) = 0$
5. Условная гомоскедастичность: $E(\varepsilon_i^2 |$ все $x_{ij}) = \sigma^2$
   В матричном виде: $E(\varepsilon_i^2|X) = \sigma^2$
6. $Cov(\varepsilon_i, \varepsilon_j|X) = 0$ при $i \neq j$
7. вектора $(x_{i.}, y_i)$ — независимы и одинаково распределены
8. с вероятностью 1 среди регрессоров нет линейно зависимых $rank(X) = k$ $det(X'X) \neq 0$ $(X'X)^{-1}$ существует

То (свойства для конечных выборок, не требующие нормальности ε):

тГМ МНК оценки $\hat{\beta}$ линены по y: $\hat{\beta}_j = c_1 y_1 + ... + c_n y_n$

тГМ $E(\hat{\beta}|X) = \beta$, и в частности $E(\hat{\beta}) = \beta$

тГМ Для любой альтернативной оценки $\hat{\beta}^{alt}$ удовлетворяющей свойствам 1 и 2: $Var(\hat{\beta}_j^{alt}|X) \geq Var(\hat{\beta}_j|X)$ $Var(\hat{\beta}_j^{alt}) \geq Var(\hat{\beta}_j)$

1. $Var(\hat{\beta}|X) = \sigma^2 (X'X)^{-1}$
2. $Cov(\hat{\beta}, \hat{\varepsilon}|X) = 0$
3. $E(\hat{\sigma}^2|X) = \sigma^2$, и $E(\hat{\sigma}^2) = \sigma^2$?остается ли при условной ГК?

свойства для конечных выборок, требующие нормальности ε Если дополнительно известно, что $\varepsilon|X \sim N$, (в частности ε и X независимы) то:

1. $t|X \sim t_{n-k}$, $t \sim t_{n-k}$
2. $RSS/\sigma^2|X \sim \chi^2_{n-k}$, $RSS/\sigma^2 \sim \chi^2_{n-k}$
3. F тест $F|X \sim F$

Асимптотические свойства:

1. $\hat{\beta} \to \beta$ по вероятности
2. $t \to N(0,1)$
3. $rF \to \chi^2_r$, r — число ограничений
4. $nR^2 \to \chi^2_{k-1}$ $\frac{RSS}{n-k} \to \sigma^2$

Лекция 4. Мультиколлинеарность. Метод главных компонент.

Лекция 5. Гетероскедастичность.

Лекция 6. Автокорреляция.

Лекция 7. Метод максимального правдоподобия. Логит и пробит-модели.

Лекция 8. Процессы авторегрессии и скользящего среднего

Лекция 9. Инструментальные переменные

Лекция 10. Квантильная регрессия. Случайный лес. Байесовский подход.