УДК 51.77

МАТРИЧНЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Пашкевич П.А.
Научный руководитель – Бричикова А.П., ассистент

Эконометрика – это наука, которая изучает различные экономические взаимосвязи с помощью математических методов и построения моделей. Важным инструментом при проведении эконометрических исследований является регрессионный анализ.

Линейной регрессией называется зависимость вида: $y_i = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik} + \varepsilon_i$, $i = \overline{1, n}$,где n – число наблюдений, i – номер наблюдения, $x_{i1}, \ldots x_{ik}$ – независимые переменные, y_i – зависимая переменная, ε_i – случайная ошибка. Неизвестные коэффициенты $\beta_0, \beta_1, \ldots \beta_k$необходимо приближенно оценить.

Для этого можно использовать метод наименьших квадратов (МНК). Суть метода состоит в минимизации суммы квадратов отклонений:

$F(\beta_0, \beta_1, \ldots, \beta_k) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_{i1} - \cdots - \beta_k x_{ik})^2$.

Линейная регрессия может быть записана в матричной форме: $Y = X \cdot \beta + \varepsilon$, где

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ 1 & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_\kappa \end{pmatrix}, \ \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Если $\det(X^T \cdot X) \neq 0$, то формула МНК-оценок коэффициентов примет вид:

$$\hat{\beta} = \begin{pmatrix} \widehat{\beta_0} \\ \widehat{\beta_1} \\ \vdots \\ \widehat{\beta_k} \end{pmatrix} = (X^T \cdot X)^{-1} X^T Y.$$

Пример.

Пусть модель регрессии задана в матричном виде: $Y = X \cdot \beta + \varepsilon$, где

$$y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, \ x = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \ \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}.$$

Дополнительно, предположим, что $M(\varepsilon) = 0$, $D(\varepsilon) = \delta^2 \cdot E$.Определим основные составляющие модели.

Число наблюдений: n=5.Число регрессов без учета свободного члена: k=2.

Рассчитаем $X^T \cdot X = \begin{pmatrix} 5 & 3 & 1 \\ 3 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ и $(X^T \cdot X)^{-1} = \begin{pmatrix} 0,5 & -0,5 & 0 \\ -0,5 & 1 & -0,5 \\ 0 & -0,5 & 1,5 \end{pmatrix}$.

При помощи МНК оценим неизвестные коэффициенты:

$\hat{\beta} = (X^T \cdot X)^{-1} \cdot X^T \cdot Y = \begin{pmatrix} 1,5 \\ 2 \\ 1,5 \end{pmatrix}$.

Прогнозное значениеY: $\hat{Y} = X \cdot \hat{\beta} = \begin{pmatrix} 1,5 \\ 1,5 \\ 3,5 \\ 3,5 \\ 5 \end{pmatrix}$.

Для того, чтобы определить качество построенной модели вычислимсумму квадратов остатков $RSS = \sum_{i=0}^{n} (y_i - \hat{y})^2 = 1$ и общую сумму квадратов $TSS = \sum_{i=0}^{n} (y_i - \bar{y})^2 = 10$ (среднее значение $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$).

Коэффициент детерминации:$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{1}{10} = 0,9$.Так как$R^2$принимает значение близкое к 1, то качество регрессии – хорошее.

Оценим дисперсию ошибок регрессии: $\hat{\sigma}^2 = \frac{RSS}{n-k-1} = \frac{1}{2}$.

Тогда оценка для ковариационной матрицы коэффициентов $\hat{\beta}$имеет вид: $\hat{V}(\hat{\beta}) = \hat{\sigma}^2 \cdot (X^T \cdot X)^{-1} = \begin{pmatrix} 0,25 & -0,25 & 0 \\ -0,25 & 0,5 & -0,25 \\ 0 & -0,25 & 0,75 \end{pmatrix}$.

Проверим значимость переменной x_1в построенном уравнении.

Нулеваягипотеза H_0: $\beta_1=0$, альтернативная гипотеза H_a: $\beta_1 \neq 0$. Уровень значимости $\alpha=0,01$. Несмещенную оценку дисперсии МНК – коэффициента $\hat{\beta}_1$находим из ковариационной матрицы: $\hat{D}(\hat{\beta}_1)=0,5$ (находится на пересечении 2-ой строки и 2-го столбца).

Тестовая статистика: $t_{набл} = \frac{\hat{\beta}1 - 0}{\sqrt{D(\hat{\beta}1)}} = \frac{2}{\sqrt{0,5}} = 2,83$. $t_{кр}=t(n-k-1)=t(2)=2,92$ (t-критическое находим в таблице Стьюдента).Так как $2,83 \in (-2,92; 2,92)$, то на уровне значимости 10% нельзя отвергнуть нулевую гипотезу. Следовательно, можем считать x_1незначимой переменной в построенной регрессии.

Протестируем значимость регрессии «в целом».Уровень значимости $\alpha=0,05$.

Нулевая гипотеза H_0: $\begin{cases} \beta_1 = 0 \\ \beta_2 = 0 \end{cases}$, альтернативная гипотеза H_a: $\left[\begin{array}{l} \beta_1 \neq 0 \\ \beta_2 \neq 0 \end{array}\right.$.

Тестовая статистика: $F = \frac{R^2}{1-R^2} \cdot \frac{n-k-1}{k} = 9$. $F_{кр}$= F(k; n-k-1)= F(2; 2)=19 (F-критические находим в таблице Фишера-Снедекора).

Так как $9 \in [0; 19]$, то на уровне значимости 5% нельзя отвергнуть нулевую гипотезу, что говорит о незначимости регрессии.

Данный пример носит иллюстративный характер. Его цель заключалась в демонстрации использования матричной формы записи регрессии. Для работы же с реальными (большими) данными удобнее использовать статистические пакеты (Stata, Statistica, SPSS, MSExcel).

Литература

1. Борзых, Д. А. Эконометрика в задачах: Базовый курс. С примерами в среде MATLAB. – М.: ЛЕНАНД, 2021.
2. Сток, Джеймс; Уотсон, Марк. Введение в эконометрику/Джеймс Сток, Марк Уотсон; пер. с англ.; под науч. ред. М.Ю.Турунцевой. – М.: Издательский дом «Дело» РАНХиГС, 2015.

УДК 336.76

АНАЛИЗ СОСТОЯНИЯ ФОНДОВОГО РЫНКА США НА ОСНОВЕ СТАТИСТИЧЕСКИХ ДАННЫХ НА 2024 ГОД

Киянко М.В., Толкач И.В.
Научный руководитель – Щукин М.В., к. ф.-м. н., доцент

Финансовые кризисы являются одним из наиболее серьезных вызовов, с которыми сталкиваются экономические системы во всем мире. Они могут приводить к значительным потерям для инвесторов и общества в целом, а также иметь негативный эффект на весь мировой рынок. В данной работе мы исследуем возможность использования статистических данных для анализа фондового рынка США. Особое внимание уделяется анализу кризисов разных годов, анализу влияния кризисов на инвестиционные стратегии и использованию индекса Баффета в качестве одного из инструментов для оценки состояния финансового рынка. Работа направлена на разработку предиктивной модели, которая поможет инвесторам и экономическим аналитикам прогнозировать возможные финансовые кризисы в США и принимать эффективные меры для минимизации рисков.

Для анализа состояния экономики, мы рассмотрим два показателя: валовой внутренний продукт и общую рыночную капитализацию 500 крупнейших компаний США с 1970 года. Используя отношение ОРК к ВВП, мы можем получить Индекс Баффета. На рисунке 3, проанализировав