Анализ панельных данных

Грабовой Андрей

Московский физико-технический институт

Москва,
2019г

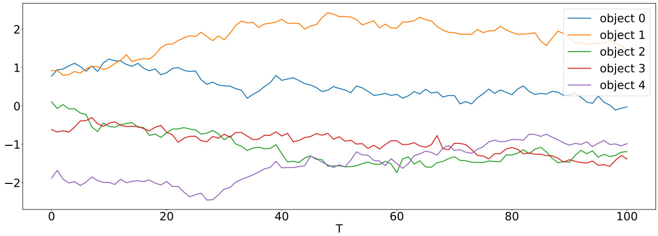
Литература

- *Я. Р. Магнус, П. К. Катышев, А. А. Пересецкий* Эконометрика. 2004. 591 с.

Что анализируется?:

- Есть набор объектов (например разные люди/фирмы/акции), которые описываются набором признаков, и по этим данным что либо прогнозируем.
- Также известно изменения прогноза во времени.

Пример:



Обозначения:

$$\mathbf{y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \dots \\ y_{iT} \end{bmatrix}, \quad \mathbf{X}_i = \begin{bmatrix} \mathbf{x}_{i1} \\ \mathbf{x}_{i2} \\ \dots \\ \mathbf{x}_{iT} \end{bmatrix}, \quad \boldsymbol{\varepsilon}_i = \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \dots \\ \varepsilon_{iT} \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \dots \\ \mathbf{y}_N \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \dots \\ \mathbf{X}_N \end{bmatrix}, \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \dots \\ \boldsymbol{\varepsilon}_N \end{bmatrix},$$

где $\mathbf{y}, \boldsymbol{\varepsilon}$ — вектора размерности $(N \cdot T) \times 1$, $\mathbf{X}$ — матрица размерности $(N \cdot T) \times n$.

Простейшая модель

Модель регрессии:

$$y_{it} = \mathbf{x}_{it}^{\mathsf{T}}\mathbf{w} + \varepsilon_{it}.$$

Модель регрессии в матричном виде:

$$\mathbf{y} = \mathbf{X}\mathbf{w} + \boldsymbol{\varepsilon},$$

данная модель не учитывает панельность данных.

Модели

Модификация:

$$y_{it} = \alpha_i + \mathbf{x}_{it}^{\mathsf{T}}\mathbf{w} + \varepsilon_{it},$$

где α_i выражает индивидуальный эффект объекта i, не зависящий от времени t.

Класификация моделей:

- *Модель с фиксированным эффектом*: предполагает, что α_i это неизвестный параметр.
- *Модель со случайным эффектом*: предполагает, что $\alpha_i = \mu + u_i$, где μ некоторый постоянный параметр, а u_i — ошибки.

Модель с фиксированным эффектом

Модель:

$$y_{it} = \alpha_i + \mathbf{x}_{it}^{\mathsf{T}}\mathbf{w} + \varepsilon_{it} = \sum_{j=1}^{N} \alpha_j d_{ij} + \mathbf{x}_{it}^{\mathsf{T}}\mathbf{w} + \varepsilon_{it},$$

где $d_{ij} = \delta_{ij}$.

Модель в матричном виде:

$$\mathbf{y} = \mathbf{D}\boldsymbol{\alpha} + \mathbf{X}\mathbf{w} + \boldsymbol{\varepsilon}.$$

Данная модель уже рассматривается как стандартная модель регрессии и при помощи МНК находятся $\boldsymbol{\alpha}$ и $\mathbf{w}$. Оценка параметров будет несмещеной и эффективной.

В случае, когда $N \gg T$ оценки $\boldsymbol{\alpha}$ не являются состоятельными. В случае $T \gg N$ оценки являются состоятельными.

Модель с фиксированным эффектом

Переход:

$$\overline{y}_i = \alpha_i + \overline{\mathbf{x}}_i^{\mathsf{T}}\mathbf{w} + \overline{\varepsilon}_i,$$

где $\overline{y}_i, \overline{\mathbf{x}}_i, \overline{\varepsilon}_i$ — усредненные по времени величины.

$$y_{it} - \overline{y}_i = (\mathbf{x}_i - \overline{\mathbf{x}}_i)^{\mathsf{T}}\mathbf{w} + \varepsilon_{it} - \overline{\varepsilon}_i.$$

Матричная форма:

$$\mathbf{My} = \mathbf{MXw} + \mathbf{M}\boldsymbol{\varepsilon}, \quad \mathbf{M} = \mathbf{I} - \mathbf{D}\left(\mathbf{D}^{\mathsf{T}}\mathbf{D}\right)^{-1}\mathbf{D}^{\mathsf{T}}.$$

Решение:

$$\hat{\mathbf{w}} = \left(\mathbf{X}^{\mathsf{T}}\mathbf{MX}\right)^{-1}\mathbf{X}^{\mathsf{T}}\mathbf{My}, \quad \hat{\alpha}_i = \overline{y}_i - \overline{\mathbf{x}}_i^{\mathsf{T}}\hat{\mathbf{w}},$$

данные оценки состоятельны для $N = const$ и при $T \gg N$.

Модель:

$$y_{it} = \alpha_i + \mathbf{x}_{it}^{\mathsf{T}}\mathbf{w} + \varepsilon_{it} = \mu + \mathbf{x}_{it}^{\mathsf{T}}\mathbf{w} + u_i + \varepsilon_{it},$$

где μ — константа, а u_i — случайная ошибка не зависящая от времени для всех объектов.

Модель в матричном виде:

$$\mathbf{y} = \boldsymbol{\mu} + \mathbf{X}\mathbf{w} + \mathbf{v},$$

где $\boldsymbol{\mu} = \mu\mathbf{i}$, а ошибки $v_{it} = u_i + \varepsilon_{it}$. Данная модель не удовлетворяет условиям Гауса-Маркова, поэтому не являются эффективными, но являются несмещеными и состоятельными.

Модель с фиксированным эффектом:

- Если $T \gg N$

Модель со случайным эффектом:

- Объекты могут рассматриваться как сэмплы и $\mathsf{E}\left(\varepsilon_i \mathbf{x}_i\right) = 0$
- Если $T \gg N$
- Если $T \gg N$