

## Семинар 11.

## 1. Пространственные статистики (аналоги коэффициента корреляции).

## • Глобальный индекс Морана (Global Moran's I)

Глобальный индекс Морана оценивает общую пространственную автокорреляцию для всего набора данных:

$$I = \frac{n}{S_0} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2},$$

где

- $n$  — количество объектов
- $y_i$  и  $y_j$  — значения признака в точках  $i$  и  $j$
- $\bar{y}$  — среднее значение признака
- $w_{ij}$  — вес пространственной связи между объектами  $i$  и  $j$  (например, бинарный вес или обратное расстояние)
- $S_0 = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$  — нормировочный коэффициент

Необходимо тестировать гипотезу о незначимости индекса Морана:

$$H_0 : I = 0,$$

$$H_1 : I \neq 0.$$

Для проверки данной гипотезы используется  $z$ -статистика вида

$$z = \frac{I - \mathbb{E}(I)}{sd(I)} \sim N(0, 1).$$

Если гипотеза  $H_0$  отвергается, то интерпретация будет следующей:

Интерпретация:  $I \approx 0$  — отсутствие автокорреляции (случайное распределение)  $I > 0$  — положительная автокорреляция (похожие значения группируются)  $I < 0$  — отрицательная автокорреляция (соседние значения различаются)

## \* Локальный индекс Морана (Local Moran's I, LISA)

Локальный индекс Морана оценивает пространственную автокорреляцию для каждого конкретного объекта.

$$I_i = \frac{(y_i - \bar{y})}{s^2} \sum_{j=1}^n w_{ij} (y_j - \bar{y}),$$

где  $S^2 = \frac{\sum_{j=1}^n (y_j - \bar{y})^2}{n}$  — дисперсия.

- положительный  $I_i$  — объект  $i$  окружен схожими значениями (кластер)
- отрицательный  $I_i$  — объект  $i$  окружен противоположными значениями (выброс)
- Около нуля — отсутствие значимой локальной автокорреляции
- Глобальный индекс показывает общий паттерн автокорреляции
- Локальный индекс помогает найти конкретные кластеры и аномалии

Оба индекса широко применяются в экономике, биологии, социологии для анализа пространственных данных.

## 2. Рассмотрим модели пространственной эконометрики:

- SAR-модель

$$y = \rho W y + X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n)$$

- SEM-модель

$$y = X\beta + u,$$

$$u = \lambda W u + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I_n).$$

- (a) Будут ли МНК-оценки SAR- и SEM-моделей несмещёнными, эффективными, состоятельными?
- (b) Для моделей SAR и SEM рассчитайте предельные эффекты. Сравните их с предельным эффектом в обычной модели множественной регрессии.

## 3. Семинарское занятие составлено на основе материалов лекций А. Niebuhr “Spatial Econometrics” (2012) в Университете г. Киля.

В работе требуется оценить неоклассическую модель роста и проверить гипотезу о конвергенции.

Для оценки того, имеют ли страны с изначально более низкими доходами более высокие темпы экономического роста, используют концепцию абсолютной и условной  $\beta$ -конвергенции. Гипотеза проверки абсолютной  $\beta$ -конвергенции подразумевает отрицательную статистическую зависимость между темпом роста показателя и его начальным уровнем.

Идея условной конвергенции является доработанной моделью абсолютной конвергенции. Эта концепция предполагает, что сближение зависит от структурных характеристик каждой экономики (в том числе и степени ее открытости),

и эти структурные различия означают, что разные страны имеют различные стационарные состояния относительно доходов на душу населения. Другими словами, страны имеют «различные траектории пропорционального роста, которые определяются специфическими региональными факторами развития». Если при условной конвергенции страна приближается к собственному устойчивому состоянию, то абсолютная конвергенция означает, что все страны сходятся к общему устойчивому состоянию.

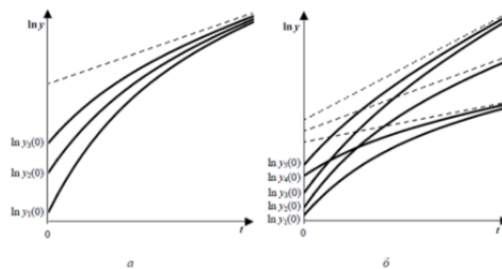


Рисунок 1. Безусловная (а) и условная (б) конвергенция. Пунктиром обозначены траектории равновесного роста.  
Источник: Gluschenko, 2009.

Файл EU27.dta. Данные по 728 территориальным единицам стран Европейского союза. Отсутствуют данных по Латвии, Малте, Кипру, Болгарии и Румынии. Имеются следующие показатели:

- $y_{95}$  и  $y_{04}$  — ВВП на душу в сопоставимых ценах 1995 и 2004 год;
- $nms\_dummy$  — дамми переменная на новые страны-члены ЕС;
- страновые дамми.

Файл inverse\_travel\_time\_EU27.dta. В файле содержатся данные об обратных расстояниях между центрами регионов в минутах.

В работе требуется оценить неоклассическую модель роста и проверить гипотезу о конвергенции:

$$\frac{1}{T} \ln \left( \frac{y_{it}}{y_{it-T}} \right) = \alpha - \frac{1 - e^{-bT}}{T} \ln(y_{it-T}) + \varepsilon_{it}.$$

Скорость сходимости:

$$\beta = -\frac{\ln(1 - T\beta_1)}{T},$$

где  $\beta_1$  — это коэффициент из регрессии

$$\ln \left( \frac{y_{it}}{y_{it-T}} \right) = \beta_0 - \beta_1 \ln(y_{it-T}) + X\gamma + \varepsilon_{it}.$$

Время, необходимое для сокращения различий между регионами в два раза:

$$h = -\frac{\ln(0.5)}{\beta}.$$

В задании необходимо оценить модель конвергенции с пространственными эффектами.

- (a) Создайте переменные: средний темп роста ВВП на душу за период 1995-2004 и натуральный логарифм ВВП на душу в 1995 году.
- (b) Сравните средний уровень среднего темпа роста ВВП на душу и ВВП на душу в 1995 году для новых и старых стран-членов Европейского союза. Опишите различия. Что можно сказать относительно гипотезы абсолютной конвергенции для стран из разных групп на основании дескриптивных статистик?
- (c) Оцените модель абсолютной бета-конвергенции для ВВП на душу. Проинтерпретируйте свой результат. Рассчитайте скорость сходимости и время сокращения разрыва в два раза ( $h$ ). Учтите гетероскедастичность: рассчитайте робастные стандартные отклонения.
- (d) Включите в модель дамми переменную на старых и новых членов ЕС: ЕС15 (старые) и ЕС10 (новые). Можно ли сказать, что включение новых стран в ЕС способствовало конвергенции?
- (e) Протестируйте наличие пространственной автокорреляции для зависимой и независимой переменных. Рассчитайте статистику I Морана и постройте диаграмму Морана. Используйте в качестве матрицы весов стандартизованную матрицу обратных расстояний, измеренных во времени. Ожидаете ли Вы наличие пространственной автокорреляции в остатках регрессии из пункта 3? Проверьте вашу гипотезу.
- (f) Оцените модели условной конвергенции с пространственными лагами (SAR и SEM) методом максимального правдоподобия. Какую модель вы предпочтете? Проинтерпретируйте результаты моделей. Что показывают коэффициенты?
- (g) Оцените модели условной бета-конвергенции, добавив в модель страновые дамми-переменные. Тем самым вы проверяете гипотезу об индивидуальной траектории сбалансированного роста для каждой страны (country-specific steady state). Изменились ли ваши выводы относительно гипотезы о конвергенции? Повторите шаги 5 и 6 для расширенной модели. Как вы считаете, может ли наличие пространственной автокорреляции быть вызвано пропуском важных переменных?