Эконометрика

ГЕТЕРОСКЕДАСТИЧНОСТЬ СЛУЧАЙНОЙ СОСТАВЛЯЮЩЕЙ

Гетероскедастичность

Одной из предпосылок применения метода наименьших квадратов являлось требование **гомоскедастичности,** предполагающей независимость случайной составляющей модели от факторных переменных и равенство дисперсий случайных ошибок в каждом наблюдении между собой ($\sigma_{\varepsilon i}^2=\sigma_{\varepsilon j}^2=$const).

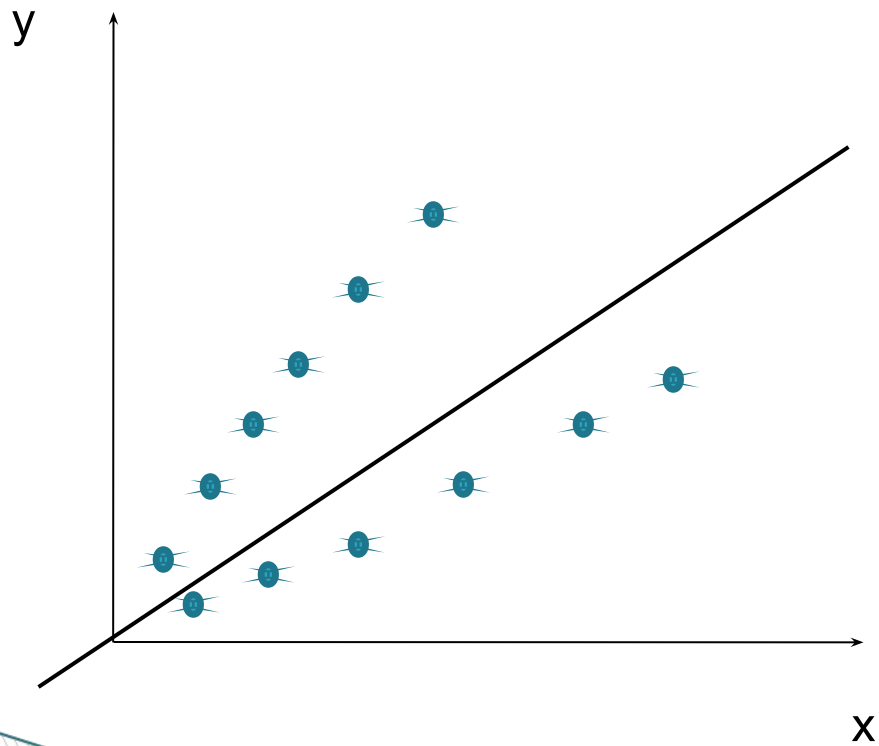
Это требование означало, что нет оснований ожидать больших случайных отклонений в любом наблюдении.

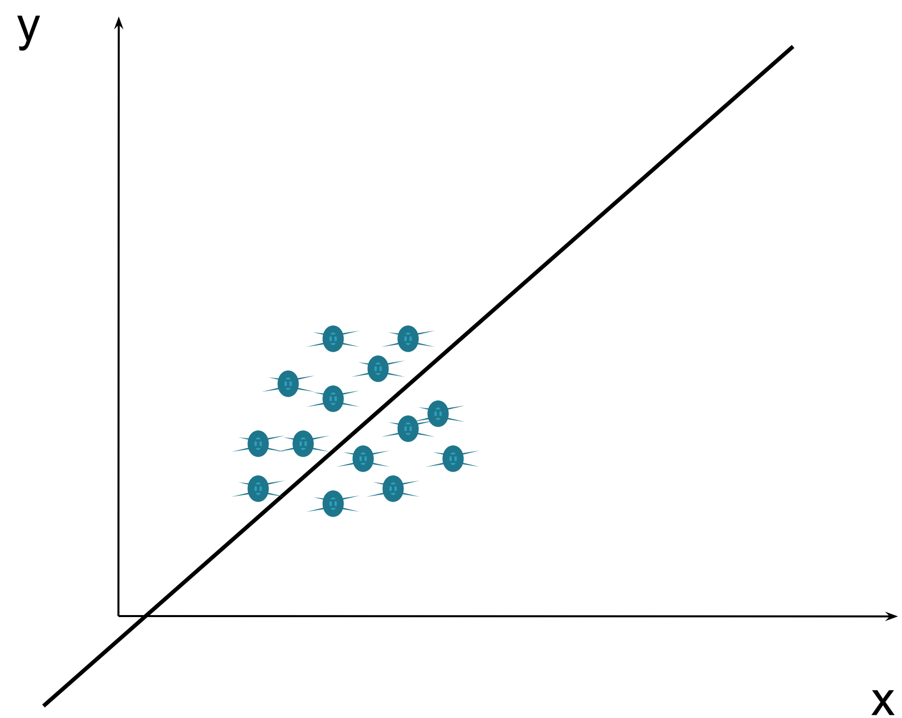
Нарушение этого требования приводит к развитию **гетероскедастичности случайной составляющей модели**.

Гетероскедастичность

- это ситуация не равенства дисперсий случайных составляющих друг другу ($\sigma_{\varepsilon i}^2 \neq \sigma_{\varepsilon j}^2 \neq const$).

Гетероскедастичность имеет место в следующих случаях:





Последствия гетероскедастичности

- Неэффективность оценок параметров регрессии;
- Неточность стандартных ошибок параметров регрессии (следовательно, неверная интерпретация значимости параметров регрессии и неверность вывода о надежности уравнения регрессии).

Обнаружение гетероскедастичности

Осуществляется по тесту Голдфелда–Квандта, который применяется в случае, когда среднее квадратическое отклонение случайной составляющей $\sigma_{\varepsilon i}$ пропорционально значению фактора в i-м наблюдении, ε_i распределено нормально.

Процедура Голдфелда–Квандта предполагает:

1. Оценку регрессии $\hat{y}_{1i}$ по первым *n* переменным (*n*<*N*/2).
2. Оценку регрессии $\hat{y}_{2i}$ по оставшимся *N*–*n* наблюдениям.
3. Расчет сумм квадратов отклонений фактических значений результата от его расчетных значений для обеих регрессий:

Процедура Голдфелда–Квандта

3. $$Q_1 = \sum_{i=1}^{n} (y_i - \hat{y}_{1i})^2 \quad u \quad Q_2 = \sum_{i=n+1}^{N} (y_i - \hat{y}_{2i})^2$$

4. Расчет отношения сумм квадратов отклонений Q_1 / Q_2 или Q_2 / Q_1, при этом в числителе должна быть наибольшая из сумм. Данное отношение имеет *F*–распределение со степенями свободы: $k_1=n-h$ и $k_2=N-h$, где *h* – число оцениваемых параметров модели. Если наблюдаемое отношение больше табличного значения *F*–распределения, то гетероскедастичность имеет место.

Устранение гетероскедастичности

Возможно при помощи деления всего уравнения регрессии на величину $\sigma_{\varepsilon i}$ и замены переменных на новые (например, в случае парной линейной регрессии):

$$Y_i = \frac{y_i}{\sigma_{\varepsilon i}}, \quad X_i = \frac{x_i}{\sigma_{\varepsilon i}}, \quad u_i = \frac{y_i}{\sigma_{\varepsilon i}}, \quad \mathrm{E}_i = \frac{\varepsilon_i}{\sigma_{\varepsilon i}}$$

Этот способ применим, в случае, когда известны фактические значения $\sigma_{\varepsilon i}$.

Кроме того, можно предположить, что $\sigma_{\varepsilon i}$ приблизительно пропорциональная x_i, следовательно, можно разделить все уравнение регрессии на x_i и ввести новые переменные, тогда гетероскедастичность тоже будет устранена.

Эконометрика

АВТОКОРРЕЛЯЦИЯ СЛУЧАЙНЫХ СОСТАВЛЯЮЩИХ

Автокорреляция

- это корреляционная зависимость между текущими уровнями каждой переменной и уровнями этой же переменной, сдвинутыми на несколько периодов времени назад.

Автокорреляция случайной составляющей – корреляционная зависимость текущих ε_i и предыдущих ε_{i-L} значений случайной составляющей. Величина *L* называется **запаздыванием** или **лагом** (сдвигом во времени). Лаг определяет порядок автокорреляции.

Автокорреляция нарушает условие независимости случайных составляющих в различных наблюдениях нормальной линейной модели регрессии. Обычно автокорреляция встречается при использовании временных рядов.

Автокорреляция

Автокорреляция может быть как положительной, так и отрицательной. **Положительная автокорреляция** означает постоянное однонаправленное действие неучтенных факторов на результат. Например, спрос на прохладительные напитки всегда выше тренда летом ($\varepsilon>0$) и ниже – зимой ($\varepsilon<0$).

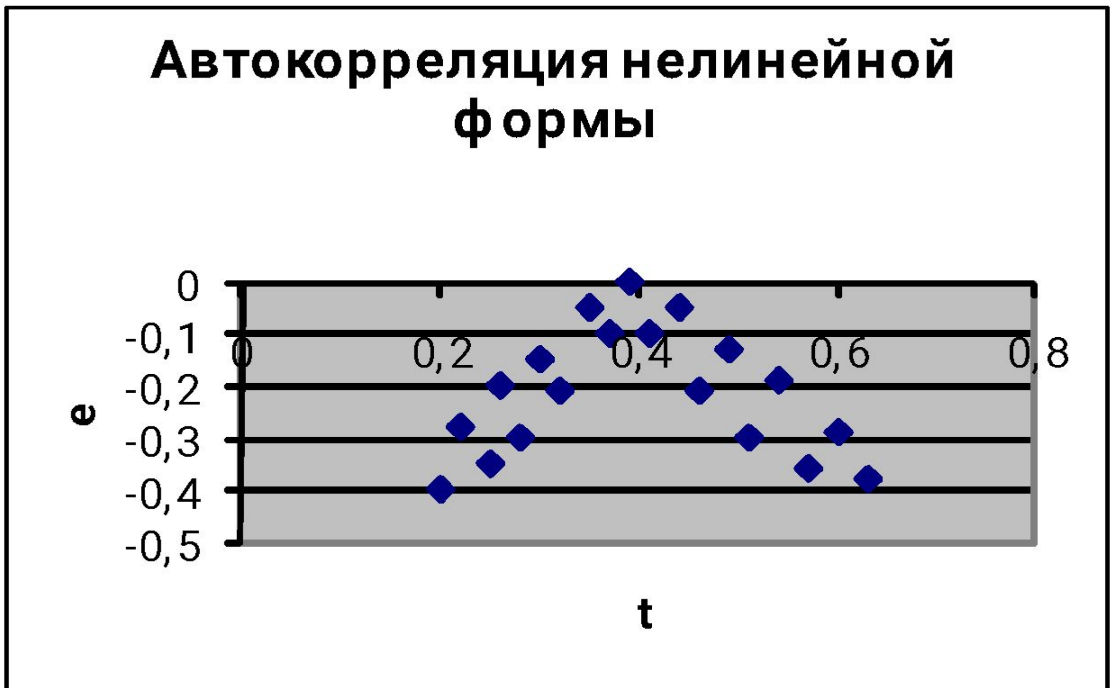
Отрицательная автокорреляция означает разнонаправленное действие неучтенных факторов, что приводит к отрицательной корреляции между последовательными значениями случайной составляющей (то есть за положительными значениями случайной составляющей в одном наблюдении идут отрицательные – в следующем). Отрицательная автокорреляция в экономике встречает крайне редко.

Последствия автокорреляции

- Неэффективность коэффициентов регрессии (при наличии несмещенности и состоятельности);
- Занижение стандартных ошибок коэффициентов регрессии.

Обнаружение автокорреляции

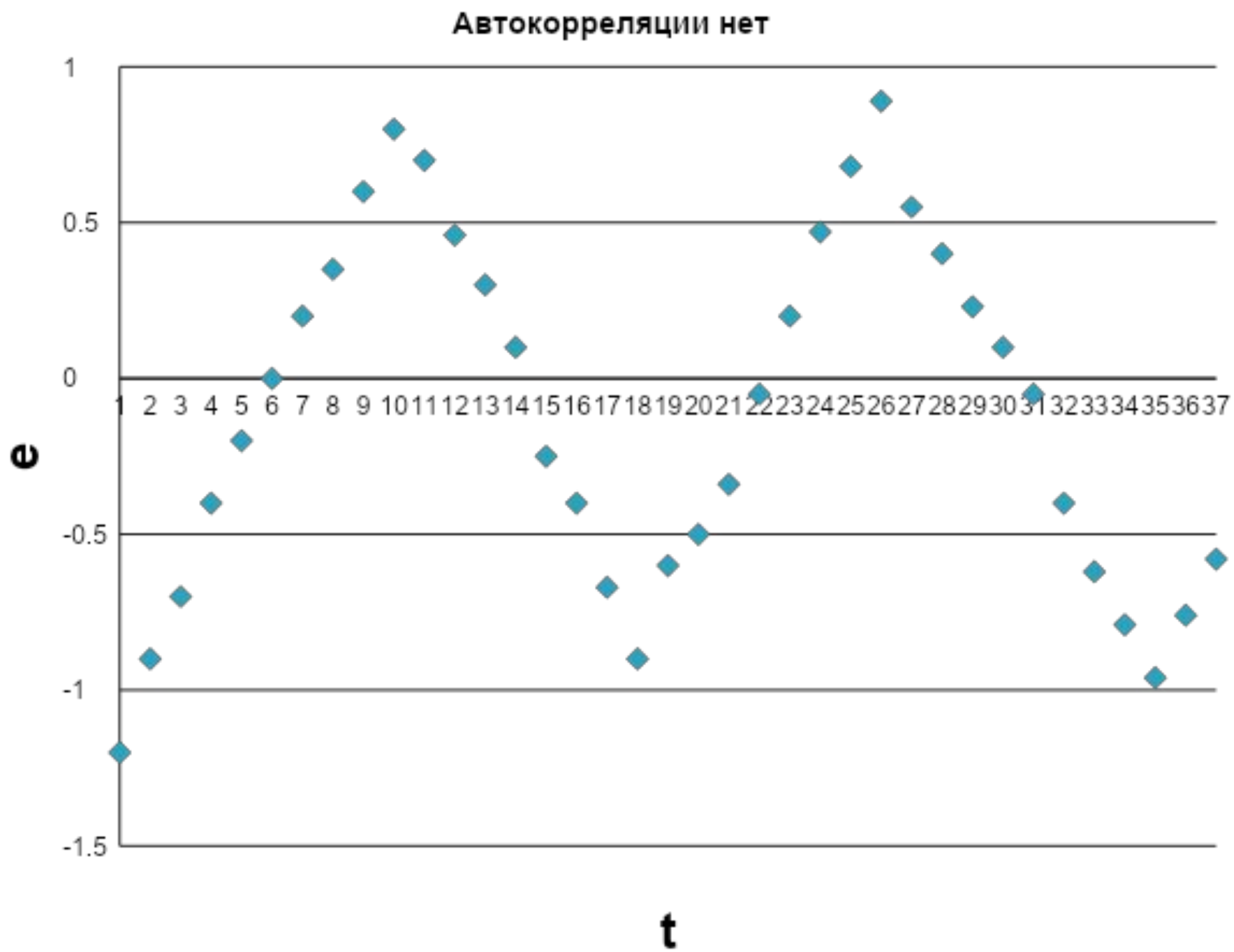
Обнаружить наличие автокорреляции можно (наиболее простым способом) с помощью анализа остатков между фактическим и рассчитанным по уравнению регрессии значением результата. Далее можно воспользоваться графическим методом. Для этого с помощью МНК–процедуры рассчитываются остатки e_t и строится график зависимости остатков от номера наблюдения



Обнаружение автокорреляции

Обнаружение автокорреляции

Обнаружение автокорреляции



Устранение автокорреляции

Необходимо:

- Выделить фактор, ответственный за автокорреляцию и включить его в уравнение регрессии (но это сложно);
- Рассчитаем коэффициент автокорреляции ρ: оценить регрессию с помощью МНК, вычислить остатки e_t для всех наблюдений, оценить регрессионную

зависимость e_t от e_{t-1}. Тогда

$$\hat{e}_t = \rho \cdot e_{t-1},$$

где

$$\rho = r_{e_t, e_{t-1}} = \frac{\sum_{t=2}^{N} e_t \cdot e_{t-1}}{\sum_{t=2}^{N} e_t^2},$$

здесь e_t - остаток регрессии по N наблюдениям, e_{t-1} – остаток регрессии по $t-1$ наблюдению.

.

Устранение автокорреляции

Произвести преобразование координат уравнения регрессии. В случае парной линейной регрессии уравнение в новых переменных будет:

$$y_t^* = a_0 \cdot q_t^* + a_1 \cdot x_t^* + \varepsilon_t^*,$$

где $\boldsymbol{y}_t^* = \boldsymbol{y}_t - \rho \cdot \boldsymbol{y}_{t-1}$; $\boldsymbol{q}_t^* = 1 - \rho$; $\boldsymbol{x}_t^* = \boldsymbol{x}_t - \rho \cdot \boldsymbol{x}_{t-1}$; $\varepsilon_t^* = \varepsilon_t - \rho \cdot \varepsilon_{t-1}$

- Далее с помощью МНК–процедуры вычисляются параметры этого уравнения и снова оцениваются остатки, а затем процесс повторяется до успешного устранения автокорреляции.

Эконометрика

АВТОКОРРЕЛЯЦИЯ УРОВНЕЙ ВРЕМЕННОГО РЯДА

Автокорреляция уровней временного ряда

Модели, построенные по временным данным, называются моделями **временных рядов** – это ряды значений какого-либо показателя, характеризующие один и тот же объект за несколько последовательных моментов или периодов времени. Уровень временного ряда X_t складывается из следующих компонент:

- Трендовой компоненты, характеризующей основную тенденцию уровней ряда (*T*);
- Циклической (периодической) компоненты, характеризующей циклические колебания изучаемого явления. Выделяют конъектурную компоненту (*K*) и сезонную – (*S*);
- Случайной компоненты, которая является результатом воздействия множества случайных факторов (*ε*).

Автокорреляция уровней временного ряда

- Уровень ряда можно представить в виде функции *X=f*(*T*,*K*,*S*,*ε*). В зависимости от вида связи между компонентами может быть построена либо **аддитивная модель** *X= T+K+S+ε*, либо **мультипликативная модель**: *X=T·K·S·ε* **ряда динамики.**
- Для выявления **структуры ряда** (т. е. состава компонент) строят автокорреляционную функцию. **Автокорреляция уровней ряда** – корреляционная связь между последовательными уровнями ряда динамики (сдвинутыми на определенный промежуток времени *L* – лаг). Она может быть измерена **коэффициентом автокорреляции**:

$$r_{t,t-L} = \frac{\overline{X_t \cdot X_{t-L}} - \overline{X_t} \cdot \overline{X_{t-L}}}{\sigma_t \cdot \sigma_{t-L}}$$

- Лаг определяет порядок коэффициента автокорреляции. Рассчитав несколько коэффициентов автокорреляции, можно определить лаг, при котором автокорреляция наиболее высокая, выявив тем самым структуру временного ряда.

Коэффициент автокорреляции

Первого порядка равен:

$$r_{x_t, x_{t-1}} = \frac{\sum_{t=2}^{N} x_t \cdot x_{t-1}}{\sum_{t=2}^{N} x_t^2}$$

Второго порядка коэффициент рассчитывается по *N* и *N*–2 наблюдениям: $r_{x_t, x_{t-2}}$

Далее рассчитываются коэффициенты третьего, четвертого и далее порядка:

$r_{x_t, x_{t-3}}$ $r_{x_t, x_{t-4}}$ $r_{x_t, x_{t-5}}, \ldots..$

Эконометрика

ПОСТРОЕНИЕ ТРЕНДА ВРЕМЕННОГО РЯДА

Тренд временного ряда

Для выявления **основной тенденции** (**тренда**) в уровнях ряда, т. е. выравнивания ряда динамики, используются различные методики:

- Методы механического выравнивания (без количественной модели);
- Метод аналитического выравнивания (с использованием количественной модели).
- Методы механического выравнивания (скользящих средних, экспоненциального сглаживания и др.) подробно изучаются в статистике.
- В эконометрике основное внимание уделяется методу аналитического выравнивания.

Метод аналитического выравнивания

Данный метод заключается в построении уравнения регрессии, характеризующего зависимость уровней ряда от временной переменной $X=f(t)$. При выборе функции тренда можно воспользоваться методом конечных разностей (при равенстве интервалов между уровнями ряда).

Метод аналитического выравнивания

Конечными разностями первого порядка являются

$$\Delta_t^1 = X_t - X_{t-1}.$$

Второго порядка

$$\Delta_t^2 = X_t - 2X_{t-1} + X_{t-2},$$

j-го порядка:

$$\Delta_t^j = \Delta_t^{j-1} - \Delta_{t-1}^{j-1}$$

Метод аналитического выравнивания

Если **тенденция выражается линейным уравнением**, то конечные разности первого порядка постоянны, а разности второго порядка равны нулю. Если **тенденция выражается параболой второго порядка**, то постоянны конечные разности второго порядка, а третьего – равны нулю. Порядок конечных разностей *j*, становящихся приблизительно равными друг другу отвечает за **степень выравнивающего многочлена**:

$$X = \sum_{i=1}^{j} a_i \cdot t^i$$

Если примерно равными оказываются темпы роста, то для выравнивания применяют показательную функцию $X = a \cdot b^t$

Метод аналитического выравнивания

При выборе вида функции следует исходить из объема имеющейся информации. Чем больше параметров содержит уравнение, тем больше должно быть наблюдений при одной и той же степени надежности оценивания. Выбор функции осуществляется и на основе принятого критерия качества уравнения регрессии, из совокупности зависимостей выбирается та, которая дает минимальное значение критерия.

Например, в случае парной линейной регрессии параметры регрессии будут:

$$\boldsymbol{b} = \frac{\sum_{t=1}^{N} \boldsymbol{x}_t \cdot \boldsymbol{t}_t^y}{\sum_{t=1}^{N} (\boldsymbol{t}_t^y)^2} \qquad \boldsymbol{a} = \frac{\sum_{t=1}^{N} \boldsymbol{x}_t}{N} \qquad \hat{x}_t = a + bt_t^y$$

Эконометрика

МОДЕЛИРОВАНИЕ СЕЗОННЫХ И ЦИКЛИЧЕСКИХ КОЛЕБАНИЙ

Моделирование сезонных и циклических колебаний

При моделировании сезонных или циклических колебаний существует несколько классических подходов:

- Расчет значений сезонной компоненты и построение аддитивной или мультипликативной модели временного ряда;
- Применение сезонных фиктивных переменных;
- Использование рядов Фурье и др.

Самый простой подход

- Рассмотрим первый наиболее простой из этих подходов для моделирования сезонных колебаний. Выбор типа модели зависит от динамики амплитуды колебаний. Если амплитуда не меняется во времени, то применяют аддитивную модель, в противном случае – мультипликативную.
- Количество исходных уровней временного ряда X_{ij} (где i=1,…, L – число сезонов (квартала, месяца и т. п.), а j=1,…, k – число года) равно $L \cdot k = N$.
- **При построении модели** вначале строят сезонную компоненту, а только после этого рассчитывают трендовую. **Для аддитивной модели** в качестве сезонной компоненты применяют абсолютное отклонение, **для мультипликативной** – индекс сезонности. В случае **аддитивной модели** сумма всех сезонных компонент должна быть равна нулю, а в случае **мультипликативной** – их произведение должно равняться единице.

Индекс сезонности и абсолютное отклонение

- Перед расчетом сезонных компонент ряд динамики выравнивают (например, с помощью метода скользящей средней) и получают выровненный ряд $X_{ij}^{в}$. Абсолютное отклонение по выровненному ряду будет:

$$\Delta_i = \frac{1}{k}\sum_{i=1}^{k}(X_{ij} - X_{ij}^{в})$$

- Индекс сезонности:

$$Is_i = \frac{1}{k}\sum_{j=1}^{k}\frac{X_{ij}}{X_{ij}^{в}}$$

- Далее при построении трендовой компоненты используется аналитическое выравнивание.

Эконометрика

СИСТЕМЫ ЭКОНОМЕТРИЧЕСКИХ УРАВНЕНИЙ

Системы эконометрических уравнений

- Сложные социально–экономические явления обычно описываются с помощью целой системы взаимосвязанных эконометрических уравнений. В некоторых случаях трудно бывает определить, какие из переменных являются зависимыми, а какие свободными. Выделяют следующие **типы эконометрических систем**:
- **Системы независимых уравнений,** в которых каждая результирующая переменная рассматривается как функция ряда выделенных факторов;
- **Системы рекурсивных уравнений,** в которых результат каждого последующего уравнения является функцией от всех переменных предыдущих уравнений;
- **Системы взаимозависимых** (совместных, одновременных) **уравнений**, в которых факторные переменные в одних уравнения входят в левую часть, а в других – в правую (одновременно одни и те же переменные рассматриваются и как результаты и как факторы).

Системы взаимозависимых уравнений

Являются наиболее сложными.

Для них традиционный МНК не применим, так как нарушаются его предпосылки.

Здесь применяется понятие структурной и приведенной формы системы одновременных уравнений.

Структурная и приведенная форма системы уравнений

Структурная форма описывает реальный экономический процесс или явление, параметры таких моделей называются **структурными.** Некоторые ее уравнения могут быть представлены тождествами. От структурной формы можно перейти к **приведенной форме** – системе независимых уравнений, в которой все текущие эндогенные переменные представлены в модели. Параметры приведенной формы определяются независимо традиционным МНК.

Зная параметры приведенной формы, и, если оценки структурных параметров можно однозначно найти по приведенным коэффициентам, можно оценить коэффициенты структурной формы модели.