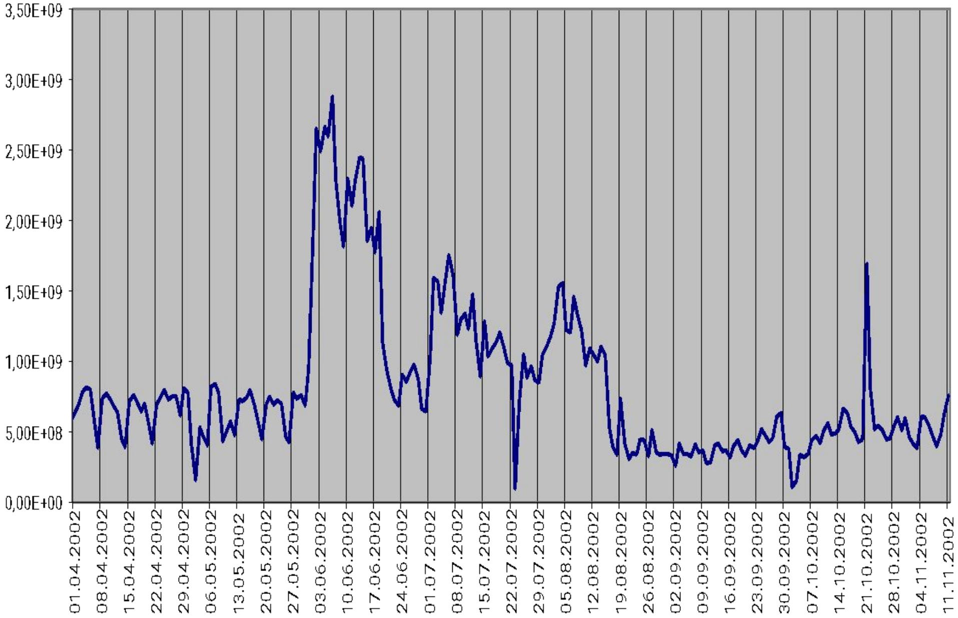
Лекция 8



- Анализ временных рядов
- Спектральный анализ (разложение в ряд Фурье, периодограмма)

Случайная функция

Это функция неслучайного аргумента, которая при каждом фиксированном значении аргумента является случайной величиной

Случайный процесс

Неслучайный аргумент t-время

Статистические характеристики случайной функции (изучает корреляционная теория)

1. Математическое ожидание - неслучайная функция - при каждом значении t = мат. ожиданию сечения.

2. Дисперсия - неслучайная функция,
   состоит из дисперсий сечений.

3. Корреляционная функция (автокорреляция) равна коэффициенту корреляции между двумя сечениями..

Стационарный случайный процесс

1.Математическое ожидание постоянно (стационарность в широком смысле).
2. Автокорреляционная функция зависит только от разности аргумента.

В.Е. Гмурман. Теория вер. и мат. статистика. Стр. 386-449

Анализ временных рядов

Временной ряд - реализация (траектория, выборочная функция) случайной функции.

временным рядом называют последовательность наблюдений, упорядоченных по времени

Аргумент (***t***) дискретно меняется через равные промежутки.

Временной ряд

Реализация случайного процесса -
неслучайная функция аргумента ***t* (времени)** -
результат экспериментов (опытов).

Временной ряд (случайная последовательность)
Аргумент (***t***) дискретно меняется через равные
промежутки.

Визуализация временного ряда

месячные международные авиаперевозки (в тысячах)

в течение 12 лет.

(Бокс и Дженкинс, 1976, стр. 531)

Модель временного ряда

тренд

случайная составляющая

$$Y(t) = f(t)+g(t)+\varepsilon(t)$$

периодическая (сезонная) составляющая

Анализ тренда

Не существует "автоматического" способа обнаружения тренда в временном ряде.

Два распространенных способа:

1) если тренд монотонный (возрастает или убывает), то используется регрессионный анализ;

2) если большая ошибка (разброс в значениях), то сначала делают сглаживание (окнами), потом регрессионный анализ.

Анализ периодической (сезонной) составляющей

1. Анализ автокорреляций (процесс авторегрессии и скользящего среднего *АРПСС* - модель не известна. Прогноз по предыдущим значениям с осреднением)
2. Анализ Фурье. Периодограмма. Отличие от АРПСС и экспоненциального сглаживания - периоды заранее неизвестны.

Анализ Фурье

модель

$$x(t_i) = \bar{x} + \sum_{k=1}^{n} \alpha_k \cos(2\pi f_k t_i) + \sum_{k=1}^{n} \beta_k \sin(2\pi f_k t_i) + e_t,$$

k – гармоника ряда $k = \overline{1, n}$;

$f_k = k/(N\Delta)$ - частота k-й гармоники ряда;

e_t – последовательность независимых, нормально распределенных случайных величин с нулевым средним значением и дисперсией σ_e^2.

Оценки коэффициентов по МНК

$$\overline{x} = \frac{1}{N}\sum_{i=1}^{N} x(t_i)$$

$$\alpha_k = \frac{2}{N}\sum_{i=1}^{N} x(t_i)\cos(2\pi f_k t_i) \ ,$$

$$\beta_k = \frac{2}{N}\sum_{i=1}^{N} x(t_i)\sin(2\pi f_k t_i),$$

$$k = \overline{1, n-1},$$

Оценки коэффициентов по МНК (продолжение)

$$\alpha_n = \frac{1}{N} \sum_{i=1}^{N} x(t_i)(-1)^i,$$

$$\beta_n = 0$$

Модель с амплитудой и фазой

$$x(t) = \bar{x} + \sum_{k=1}^{n} R_k \cos(2\pi f_k t - \varphi_k) + e_t$$

$$R_k = \sqrt{\alpha_k^2 + \beta_k^2},$$

$$\varphi_k = arctg(\frac{\beta_k}{\alpha_k}),$$

$$T_k = \frac{N\Delta}{k},$$

$$f_k = \frac{1}{T_k}$$

Периодограмма или *линейчатый спектр Фурье*

$$I(f_k) = \alpha_k^2 + \beta_k^2$$

Интенсивность k-той гармоники

Оценка дисперсии величины $x(t_i)$

$$S^2 = \frac{1}{N}\sum_{i=1}^{N}\left(x(t_i)-\bar{x}\right)^2 = \frac{1}{2}\sum_{k=1}^{n-1}\left(\alpha_k^2+\beta_k^2\right)+\alpha_n^2$$

где $N=2\cdot n$,

т.е. общее число наблюдений – четная величина. При нечетном $N=2\cdot n-1$ слагаемое α_n исчезает.

1. Выделение значимых гармоник по критерию Фишера

$$\frac{I(f_k)/\nu_1}{S^2/\nu_2} > F(\alpha,\nu_1,\nu_2)$$

ν_1 – число степеней свободы, приходящихся на гармонику ряда Фурье =2,

$\nu_2=N-1$ – число степеней свободы всего ряда,

α –уровень значимости.

Выделение значимых гармоник по вкладу доминирующих гармоник в дисперсию

$$\rho_k^2 = \frac{R_k^2}{2S^2} * 100\% = \frac{\alpha_k^2 + \beta_k^2}{\sum_k^{n-1}(\alpha_k^2 + \beta_k^2) + \alpha_n^2} * 100\%$$

Основными гармониками принято считать те,

у которых $\sum_{k=1}^{n} \rho_k^2 \geq 80\%$

Модель после выделения гармоник

$$x(t_i) = \bar{x} + \sum_{k=1}^{m} \alpha_k \cos(2\pi f_k t_i) + \sum_{k=1}^{m} \beta_k \sin(2\pi f_k t_i)$$

где m –

Алгоритм расчета коэффициентов разложения в ряд Фурье

1. Рассчитать оценку математического ожидания $\overline{x}$, оценку дисперсии S^2, стандартное отклонение S.
2. Определить количество наблюдений N, моменты времени, в которые производилось наблюдения $t_i=i\cdot\Delta$, где Δ - интервал между отдельными наблюдениями.

Алгоритм расчета коэффициентов разложения в ряд Фурье

3 Определить количество гармоник $k = \overline{1, n-1}$,

где $n = \begin{cases} \dfrac{N}{2}, & \text{если } N \text{ четное} \\ \dfrac{N-1}{2}, & \text{если } N \text{ нечетное} \end{cases}$

4 Для каждой k-ой гармоники рассчитать:

- частоту f_k и период T_k: $f_k = k/(N\Delta)$, $T_k = 1/f_k$
- коэффициенты при синусе и косинусе (см. оценки коэффициентов по МНК)

$$\alpha_k = \frac{2}{N}\sum_{i=1}^{N} x(t_i)\cos(2\pi f_k t_i), \quad \beta_k = \frac{2}{N}\sum_{i=1}^{N} x(t_i)\sin(2\pi f_k t_i)$$

Алгоритм расчета коэффициентов разложения в ряд Фурье

5 Вычислить:

- амплитуду и фазу $R_k = \sqrt{\alpha_k^2 + \beta_k^2}, \quad \varphi_k = arctg(\frac{\beta_k}{\alpha_k});$
- вклад в дисперсию ряда ρ_k^2 :

$$\rho_k^2 = \frac{R_k^2}{2S^2} * 100\% = \frac{\alpha_k^2 + \beta_k^2}{\sum_k^{n-1}(\alpha_k^2 + \beta_k^2) + \alpha_n^2} * 100\%$$

Алгоритм расчета коэффициентов разложения в ряд Фурье

6 Если N – четное, то вычислить коэффициент α_n

$$\alpha_n = \frac{1}{N}\sum_{i=1}^{N} x(t_i)(-1)^i$$

При нечетном N коэффициент $\alpha_n=0$.

7 Проанализировать полученные результаты.

Литература

1. Андерсон Т. Статистический анализ временных рядов. М.: Мир, 1976, 755 с.
2. Дженкинс Г., Ваттс Д. Спектральный анализ и его приложения. М.: Мир, 1972, вып 2., 287 с
3. Кендалл М. Дж., Стьюарт А. Многомерный статистический анализ и временные ряды. М.: Наука, 1976, 736 с.
4. Дрейпер Н., Симт Г. Прикладной регрессионный анализ. М.: Финансы и статистика, 1986, 366 с.
5. Чекотовский Э.В. Графический анализ статистических данных в Microsoft Excel 2000, Изд-во: Вильямс 2002, 464 с.
6. http://education.iet.ru/files/text/econometrics/lect/