Статистика – 4. Анализ временных рядов – Часть 1

Д-р Вольфганг Бёме

Январь 2013

Содержание

- 1. Введение
- 2. Дескриптивный одномерный анализ распределений
- 3. Дескриптивный корреляционный и регрессионный анализ
- **4. Дескриптивный анализ временных рядов**
- 5. Относительные показатели / Индексы
- 6. Введение в избранные многомерные методы анализа

- Статистическое программное обеспечение: SPSS

Содержание

- **4.0 Определения**
- 4.1 Скользящие средние
- 4.2 Экспоненциальное сглаживание
- 4.3 Трендовые функции
- 4.4 Фильтры
- 4.5 Автокорреляция
- 4.6 Перспективы: Множественный анализ временных рядов

Временные ряды

Временной ряд – это упорядоченная во времени последовательность значений статистического признака. Как правило, предполагается, что наблюдаемые значения представлены через равные промежутки времени.
Обозначение: $\{y_t,\ t = 1, 2, ..., n\}$

В экономических исследованиях различают:

- Годовые данные (напр., валовой внутренний продукт)
- Квартальные данные (напр., валовой внутренний продукт)
- Месячные данные (напр., индексы цен)
- Дневные данные (напр., биржевые цены, процентные ставки)

Временные ряды

Внутригодовой временной ряд $\{y_t,\ t = 1, 2, ..., n\}$ является результатом взаимодействия следующих компонент:

- Гладкая компонента (g_t) (трендовая компонента (T_t), конъюнктурная или циклическая компонента (Z_t))
- Сезонная компонента (s_t)
- Нерегулярная компонента (i_t) (не сводимая к (g_t) и (s_t), остаточная величина)

Две модели временного ряда (связь компонент):

1. Аддитивная связь:
$$y_t = g_t + s_t + i_t$$
2. Мультипликативная связь:
$$y_t = g_t \cdot s_t \cdot i_t$$

- **4.1 Скользящие средние**
- 4.2 Экспоненциальное сглаживание
- 4.3 Трендовые функции
- 4.4 Фильтры
- 4.5 Автокорреляция
- 4.6 Перспективы: Множественный анализ временных рядов

4.1 Скользящие средние – Содержание

- **4.1.1 Центрированные скользящие средние**
- 4.1.2 Ретроспективное скользящее среднее

Центрированные скользящие средние

Определение: Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\}$ и k – целая константа $(k \geq 0)$. Тогда временной ряд $\{\overline{y_j},\ j = k+1, k+2, ..., n-k\}$ из $n - 2k$ средних арифметических $\overline{y_j}$:

$$\overline{y_j} = \frac{1}{r} \sum_{i=-k}^{k} y_{i+j}$$

называется **центрированным скользящим средним** с **нечётным** опорным диапазоном $r = 2 \cdot k + 1$.

Определение: Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\}$ и k – целая константа $(k \geq 0)$. Тогда временной ряд $\{\overline{y_j},\ j = k+1, k+2, ..., n-k+1\}$ из $n - 2k$ средних арифметических $\overline{y_j}$:

$$\overline{y_j} = \frac{1}{r} \sum_{i=-k}^{k-1} y_{i+j}$$

называется **центрированным скользящим средним** с **чётным** опорным диапазоном $r = 2 \cdot k$.

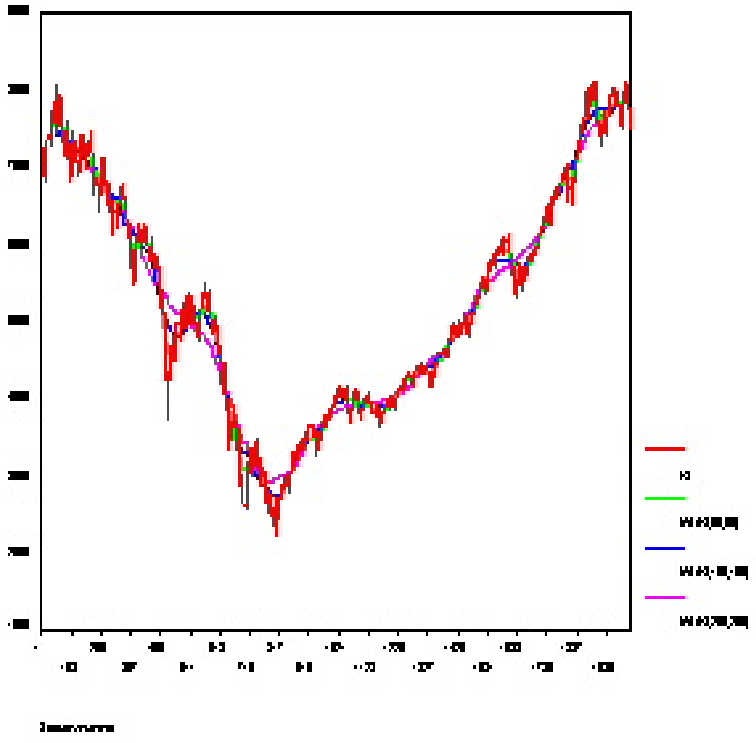
Простой пример

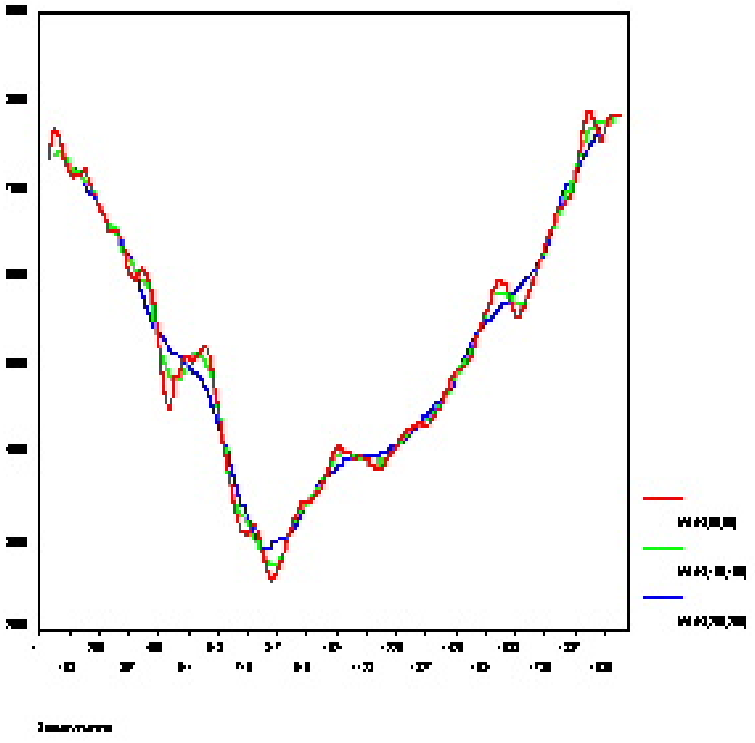
Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\} = \{1, 2, 4, 7, 9, 10, 11\}$

1. Нечётный опорный диапазон: $r = 3,\ k = 1$
2. Чётный опорный диапазон: $r = 2,\ k = 1$
3. Частный случай – нечётный опорный диапазон: $r = 1,\ k = 0$

Пример из финансовой математики

Динамика DAX и центрированные скользящие средние ($r = 50, 100, 200$)





4.1 Скользящие средние – Содержание

- 4.1.1 Центрированные скользящие средние
- **4.1.2 Ретроспективное скользящее среднее**

Ретроспективные скользящие средние

Определение: Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\}$ и k – целая константа ($k \geq 0$). Тогда временной ряд $\{\overline{y_j},\ j = 2k + 1, 2k + 2, ..., n\}$ из $n - 2k$ средних арифметических $\overline{y_j}$:

$$\overline{y_j} = \frac{1}{r} \sum_{i=-2k}^{0} y_{i+j}$$

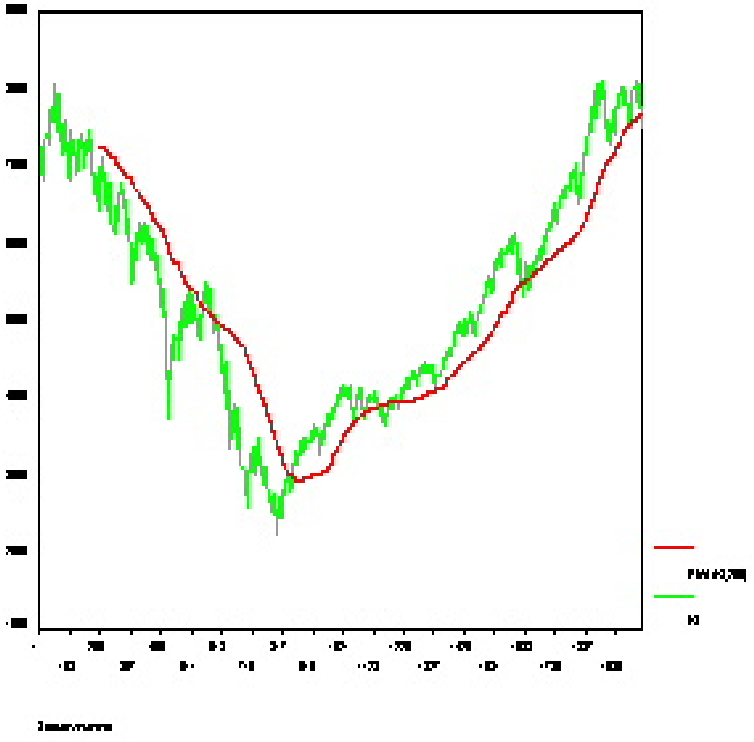
называется **ретроспективным скользящим средним** с опорным диапазоном $r = 2 \cdot k + 1$.

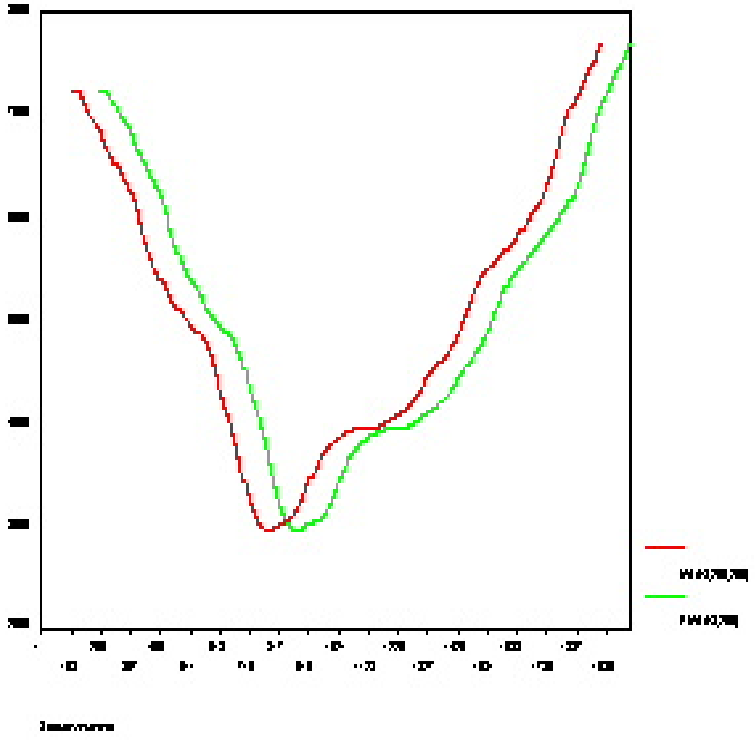
Пример: Пусть задан временной ряд
$\{y_t,\ t = 1, 2, ..., n\} = \{1, 2, 4, 7, 9, 10, 11\}$

1. Опорный диапазон: $r = 3,\ k = 1$
2. Частный случай: $r = 1,\ k = 0$

Пример из финансовой математики

Динамика DAX и ретроспективные скользящие средние ($r = 50, 100, 200$)





- 4.1 Скользящие средние
- **4.2 Экспоненциальное сглаживание**
- 4.3 Трендовые функции
- 4.4 Фильтры
- 4.5 Автокорреляция
- 4.6 Перспективы: Множественный анализ временных рядов

4.2 Экспоненциальное сглаживание – Содержание

- **4.2.1 Базовая модель**
- 4.2.2 Учёт линейного тренда
- 4.2.3 Учёт сезонной компоненты

Базовая модель

Цель: простая прогнозная модель для одношагового прогноза (для y_{n+1})

Определение: Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\}$. Простая прогнозная модель

$$y_t^* = \alpha y_t + (1 - \alpha) y_{t-1}^*$$

называется **базовой моделью экспоненциального сглаживания**.

Замечания:

1. y_t^* = прогнозное значение для y_{t+1}
2. α – коэффициент сглаживания (α близко к 1 – почти без сглаживания) $\Rightarrow$ Пример
3. Выбор параметра α по критерию $\sum_{t=2}^{n} (y_t - y_{t-1}^*)^2 = MIN$
4. $y_t^* = F(y_t, y_{t-1}^*)$ зависит от наблюдения y_t и оценки предыдущего периода y_{t-1}^*
5. $y_t^* = \alpha y_t + \alpha(1 - \alpha) y_{t-1} + (1 - \alpha)^2 y_{t-2} + ... \Rightarrow$ экспоненциально убывающие веса (текущим наблюдениям придаётся больший вес)
6. $y_t = F(y_t, y_{t-1}, ..., y_2, y_1^*) \Rightarrow$ начальное значение должно быть задано

Базовая модель

Выбор начального значения – разные подходы в статистическом ПО:

1. SPSS: среднее всех значений:

$$y_1^* = \frac{1}{n}\sum_{t=1}^{n} y_t$$

2. EVIEWS: среднее первых 4 значений:

$$y_1^* = \frac{1}{4}\sum_{t=1}^{4} y_t$$

Простой пример

Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\} = \{1, 2, 4, 7, 9, 10, 11\}$

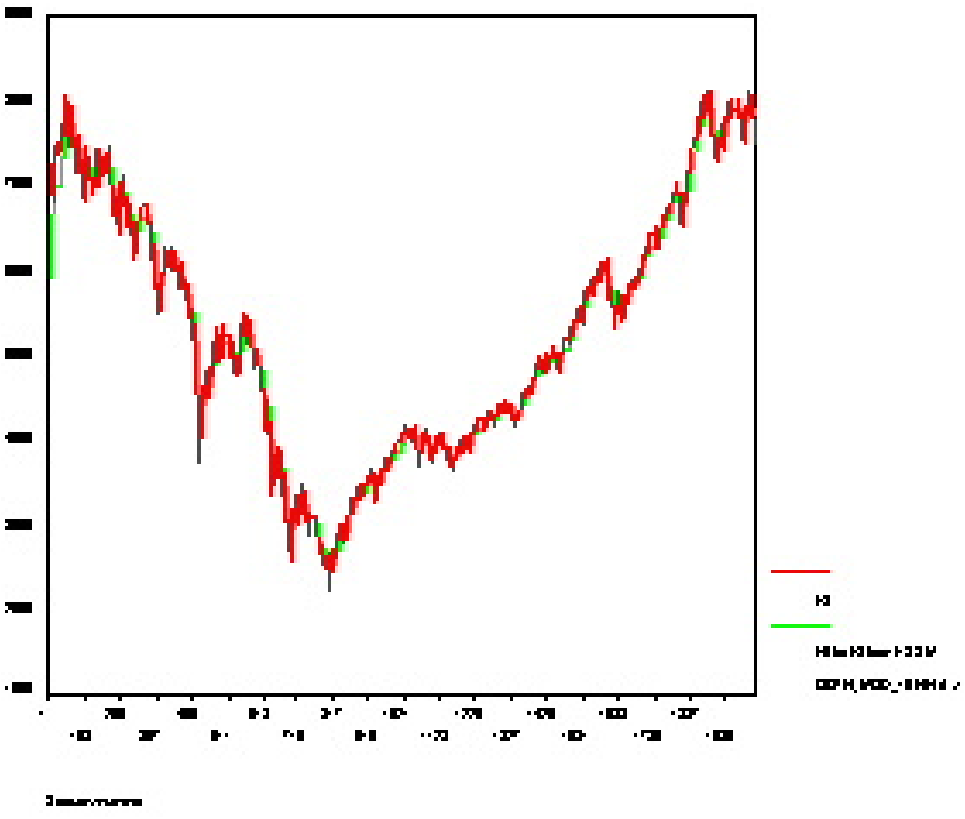
Начальное значение: $y_1^* = \frac{1}{n} \sum_{t=1}^{n} y_t$

Параметр α:

1. $\alpha = 0{,}01$
2. $\alpha = 0{,}3$
3. $\alpha = 0{,}5$
4. $\alpha = 0{,}9$

Пример из финансовой математики

Динамика DAX и базовая модель экспоненциального сглаживания



Экспоненциальное сглаживание – Содержание

- 4.2.1 Базовая модель
- **4.2.2 Учёт линейного тренда**
- 4.2.3 Учёт сезонной компоненты

Двойное сглаживание – метод ХОЛТА

Цель: учёт линейного тренда

Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\}$. Базовая модель:

$$y_t^* = \alpha y_t + (1 - \alpha) y_{t-1}^*$$

Введение трендовой компоненты $T_t \Longrightarrow$ модифицированный подход:

$$\begin{aligned} y_t^* &= \alpha y_t + (1 - \alpha)(y_{t-1}^* + T_{t-1}) \\ T_t &= \beta(y_t^* - y_{t-1}^*) + (1 - \beta) T_{t-1} \end{aligned}$$

Замечания:

1. Начальные значения y_1^* и T_1 должны быть заданы
2. y_t^* = прогнозное значение для y_{t+1}

Простой пример

Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\} = \{1, 2, 4, 7, 9, 10, 11\}$

Начальное значение: y_1^*, T_1 из SPSS

Параметр α:

1. $\alpha = 0{,}01$
2. $\alpha = 0{,}3$
3. $\alpha = 0{,}5$
4. $\alpha = 0{,}9$

Пример из финансовой математики

Динамика DAX и экспоненциальное сглаживание по методу ХОЛТА

Results of EXSMOOTH procedure for Variable DAX
MODEL= HOLT (Linear trend, no seasonality)

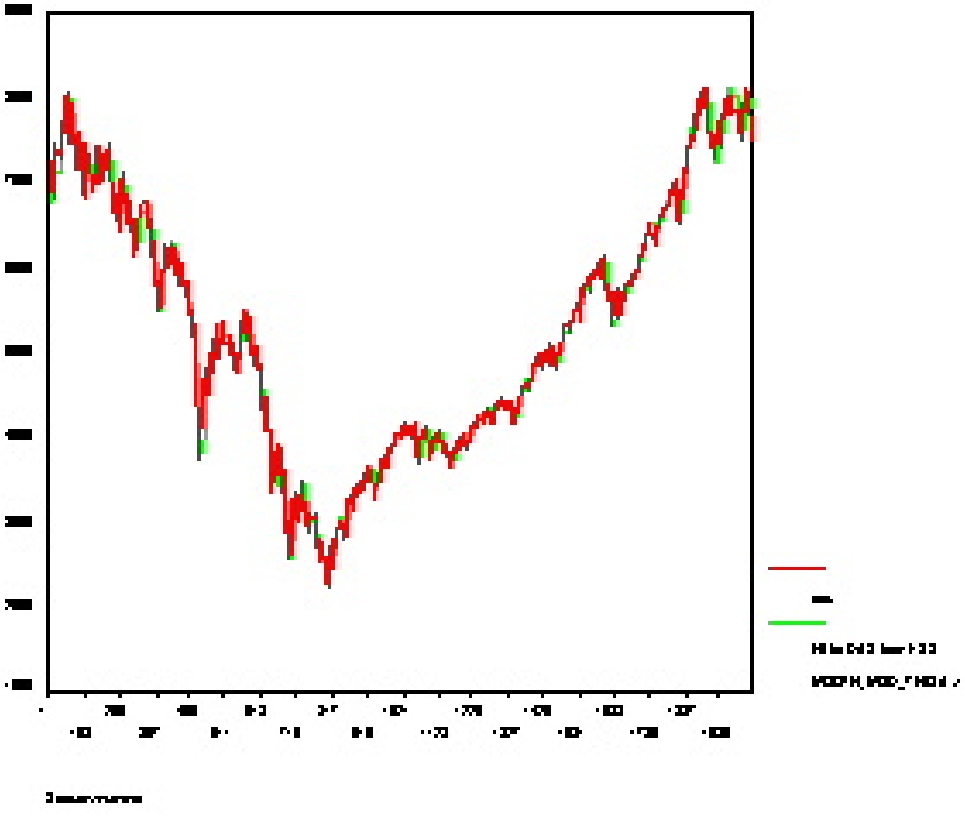
Initial values:	Series	Trend
	6958.01928	.24144

DFE = 2025.

The SSE is:	Alpha	Gamma	SSE
	.1000000	.1000000	51472963.140

Пример из финансовой математики

Динамика DAX и экспоненциальное сглаживание по методу ХОЛТА



Экспоненциальное сглаживание – Содержание

- 4.2.1 Базовая модель
- 4.2.2 Учёт линейного тренда
- **4.2.3 Учёт сезонной компоненты**

Двойное сглаживание – метод ХОЛТА

Цель: учёт линейного тренда и сезонной компоненты

Пусть задан временной ряд $\{y_t,\ t = 1, 2, \ldots, n\}$ и трендовая компонента T_t. Сглаживание по ХОЛТУ:

$$\begin{aligned} y_t^* &= \alpha y_t + (1-\alpha)(y_{t-\mathbf{1}}^* + T_{t-\mathbf{1}}) \\ T_t &= \beta(y_t^* - y_{t-\mathbf{1}}^*) + (1-\beta)T_{t-\mathbf{1}} \end{aligned}$$

Введение трендовой компоненты $c_t \Longrightarrow$ модифицированный подход:
Аддитивная сезонная компонента:

$$\begin{aligned} y_t^* &= \alpha(y_t - c_{t-s}) + (1-\alpha)(y_{t-\mathbf{1}}^* + T_{t-\mathbf{1}}) \\ T_t &= \beta(y_t^* - y_{t-\mathbf{1}}^*) + (1-\beta)T_{t-\mathbf{1}} \\ c_t &= \gamma(y_t - y_t^*) + (1-\gamma)c_{t-s} \end{aligned}$$

Замечания:

1. $s = 12$ для месячных данных
2. $s = 4$ для квартальных данных ...
3. Мультипликативная связь также возможна $\Longrightarrow$ модифицированный вид системы уравнений

Двойное сглаживание – метод ХОЛТА

Мультипликативная сезонная компонента:

$$\begin{aligned} y_t^* &= \alpha \frac{y_t}{c_{t-s}} + (1-\alpha)(y_{t-1}^* + T_{t-1}) \\ T_t &= \beta(y_t^* - y_{t-1}^*) + (1-\beta)T_{t-1} \\ c_t &= \gamma \frac{y_t}{y_t^*} + (1-\gamma)c_{t-s} \end{aligned}$$

Замечания:

1. $s = 12$ для месячных данных
2. $s = 4$ для квартальных данных ...
3. Мультипликативная связь = стандартная настройка в SPSS

- 4.1 Скользящие средние
- 4.2 Экспоненциальное сглаживание
- **4.3 Трендовые функции**
- 4.4 Фильтры
- 4.5 Автокорреляция
- 4.6 Перспективы: Множественный анализ временных рядов

4.3 Трендовые функции – Содержание

- **4.3.1 Определение**
- 4.3.2 Пример

Определение

Определение: Пусть X и Y – два метрически масштабированных признака. Значения получены у n статистических единиц:

$$(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$$

Функция $y^* = f(x)$ для описания статистической зависимости Y от X называется **простой регрессией** Y на X.

Частный случай:

Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\}$. Функция $y^* = f(t)$ для описания статистической зависимости признака Y от времени t называется **трендовой функцией**.

$$\begin{aligned} y &= f(x) + u = y^* + u \\ y &= f(t) + u = y^* + u \end{aligned}$$

4.3 Трендовые функции – Содержание

- 4.3.1 Определение
- **4.3.2 Пример**

Пример – обменный курс USD/EUR

Excel: трендовая функция обменного курса USD/EUR

2 периода:

1. 15.8 по 30.11 – линейный тренд
2. 2.1 по 30.4 – полином 2-й степени

Сравнение моделей – пример с обменным курсом USD/EUR

Следующие типы моделей:

1. Линейная
2. Квадратичная

Критерии выбора:

1. Коэффициент детерминации
2. Скорректированный коэффициент детерминации

$\Rightarrow$ SPSS

- 4.1 Скользящие средние
- 4.2 Экспоненциальное сглаживание
- 4.3 Трендовые функции
- **4.4 Фильтры**
- 4.5 Автокорреляция
- 4.6 Перспективы: Множественный анализ временных рядов

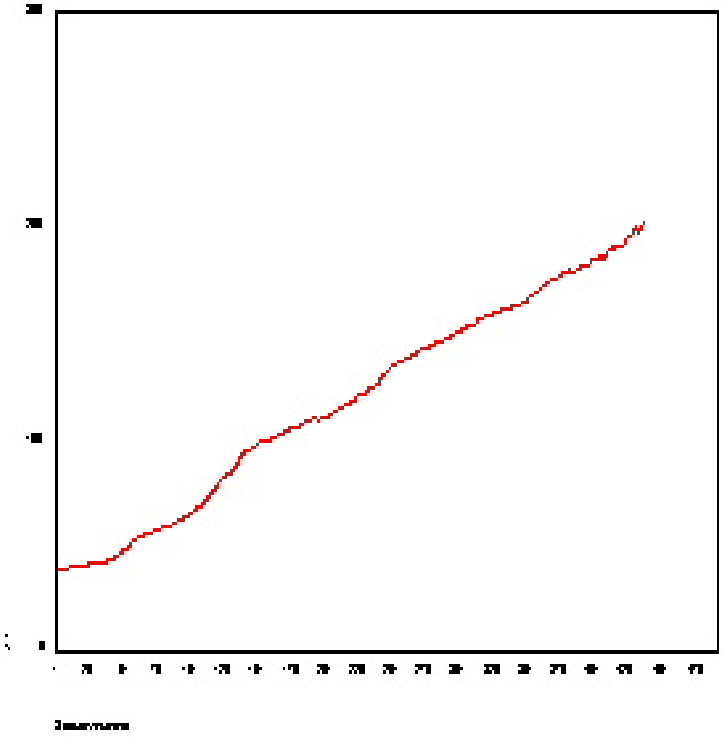
4.4 Фильтры – Содержание

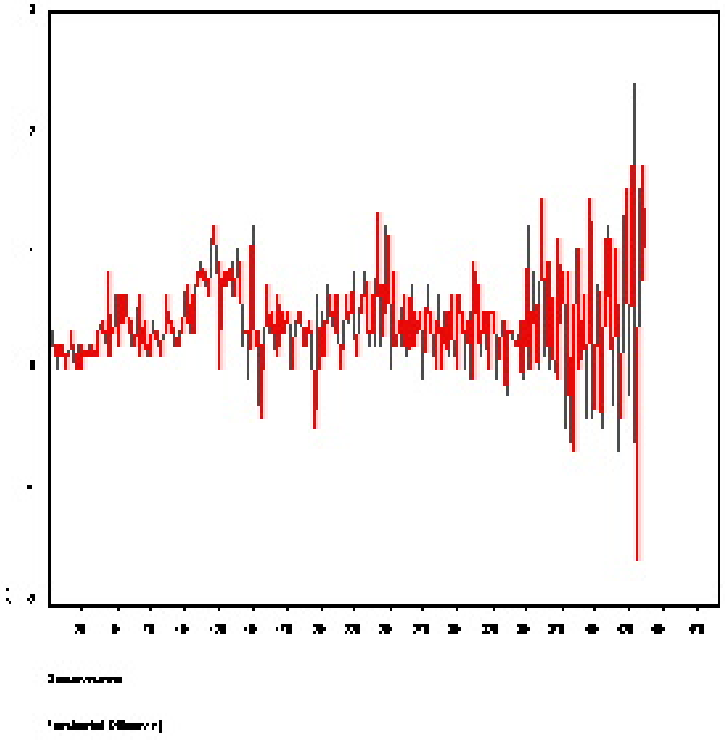
- **4.4.1 Разностный фильтр**
- 4.4.2 Сезонный разностный фильтр

Разностный фильтр

Цель: устранение тренда

Вводный пример: ИПЦ США





Разностный фильтр

Определение: Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\}$.

Разностный фильтр: $\Delta y_t = y_t - y_{t-1}$

Другие распространённые разностные фильтры:

Разности логарифмов (логарифмические доходности):

$$\Delta \widehat{y_t} = ln(y_t) - ln(y_{t-1}) = ln\left(\frac{y_t}{y_{t-\mathbf{1}}}\right)$$

Разности 2-го порядка:

$$\Delta^2 y_t = \Delta y_t - \Delta y_{t-1} = y_t - 2y_{t-1} + y_{t-2}$$

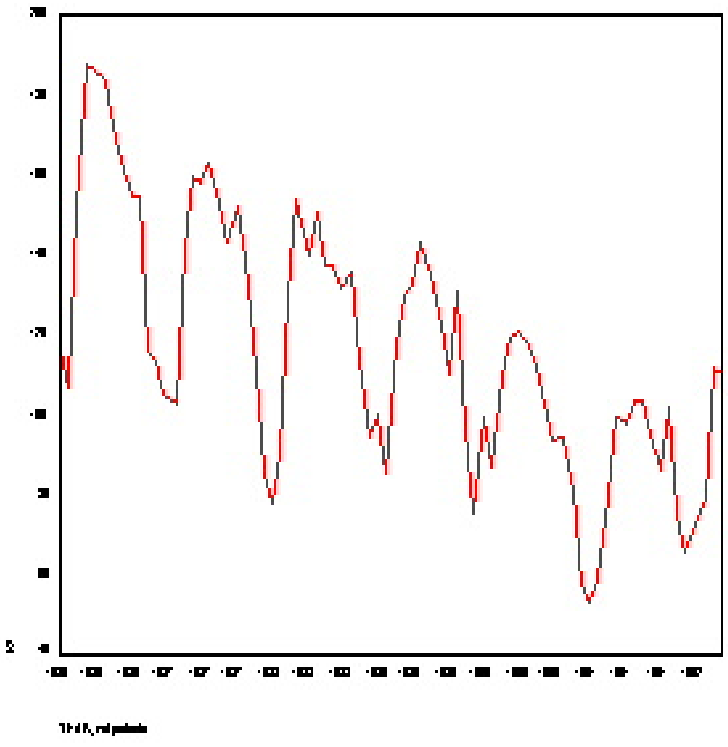
4.4 Фильтры – Содержание

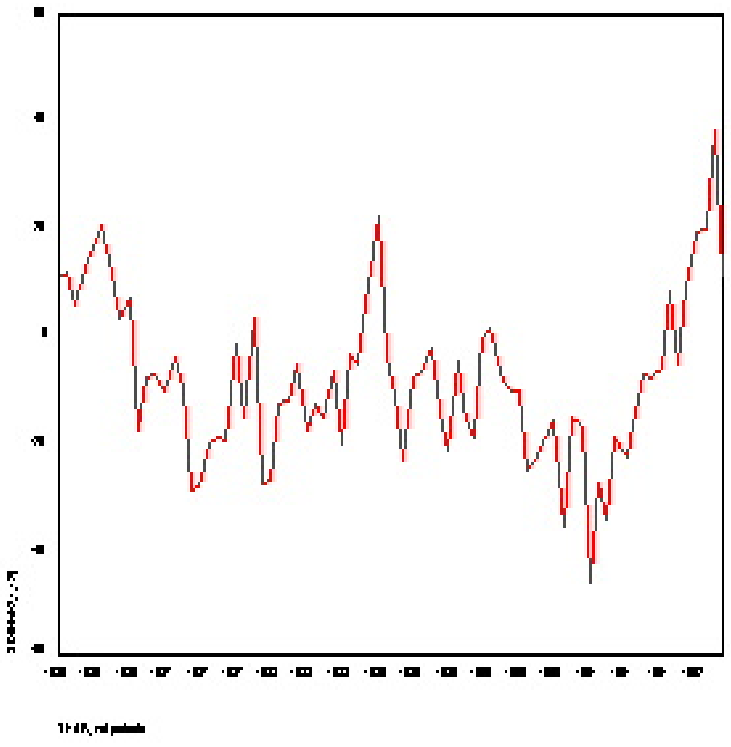
- 4.4.1 Разностный фильтр
- **4.4.2 Сезонный разностный фильтр**

Сезонная разность

Цель: устранение сезонного эффекта

Вводный пример: начало строительства жилья в США, 1959–1990





Сезонная разность

Определение: Пусть задан временной ряд $\{y_t,\ t = 1, 2, ..., n\}$.

Идентифицированная периодичность s

Разностный фильтр: $\Delta y_t = y_t - y_{t-s}$

1. $s = 12$ для месячных данных
2. $s = 4$ для квартальных данных ...