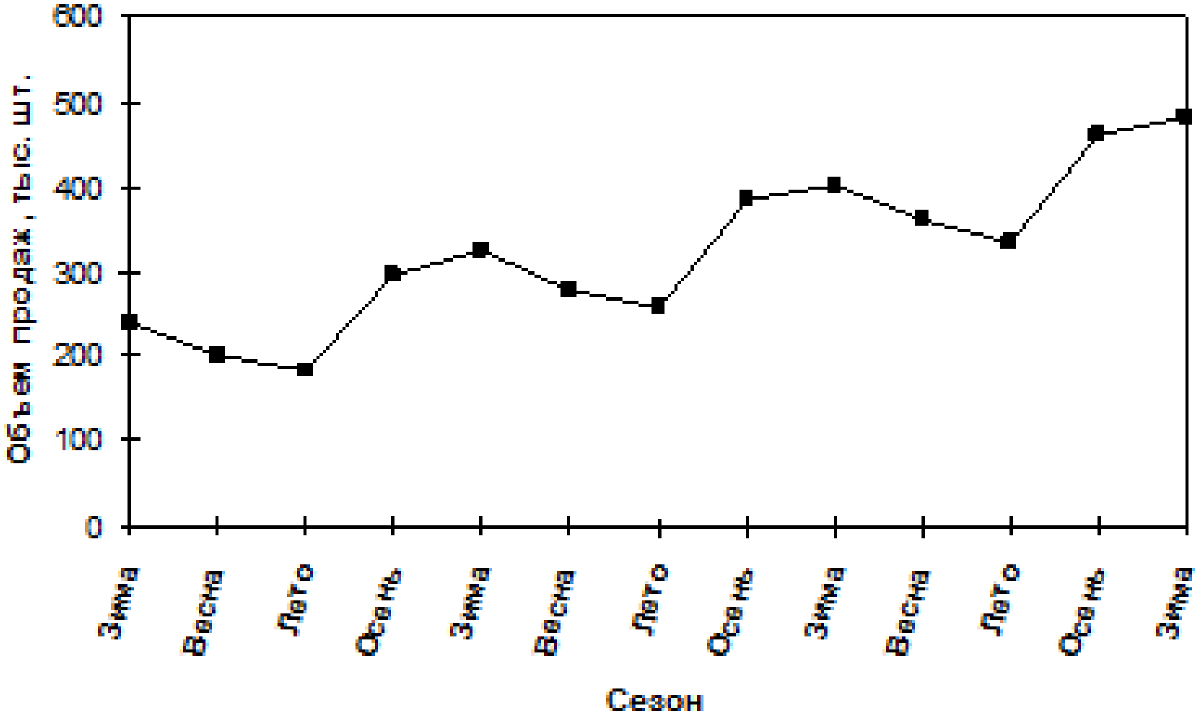
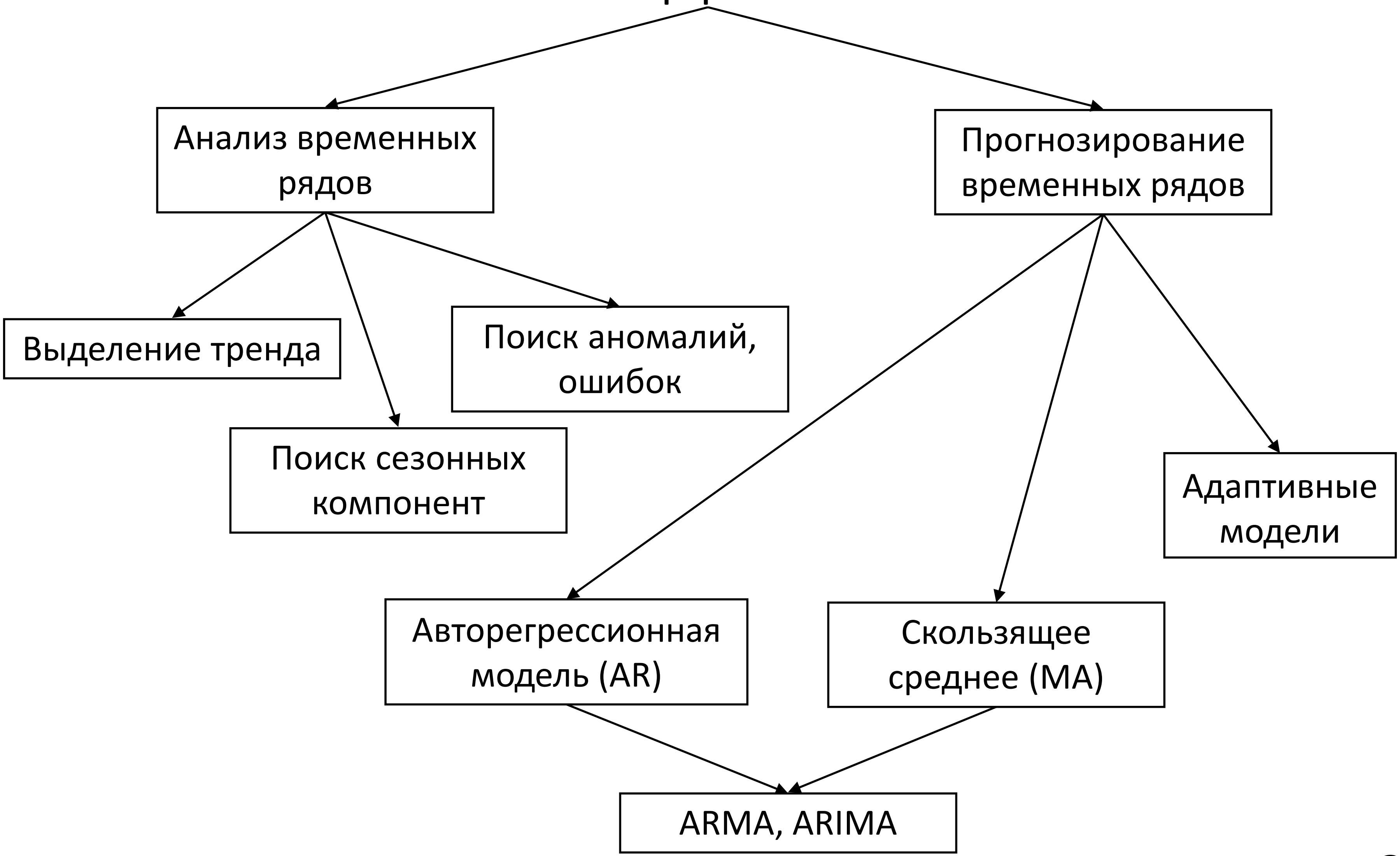
Временные ряды в R

Магистр 2 курс ПИ - Жибарь М.А.

Временной ряд - это совокупность значений какого-либо показателя за несколько последовательных моментов или периодов времени

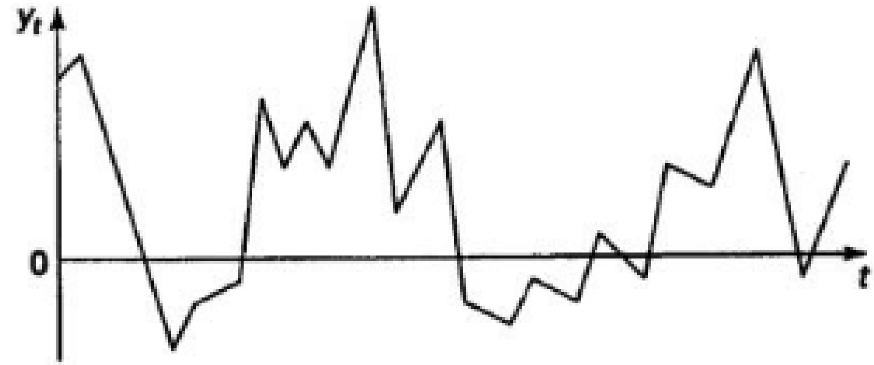


Задачи

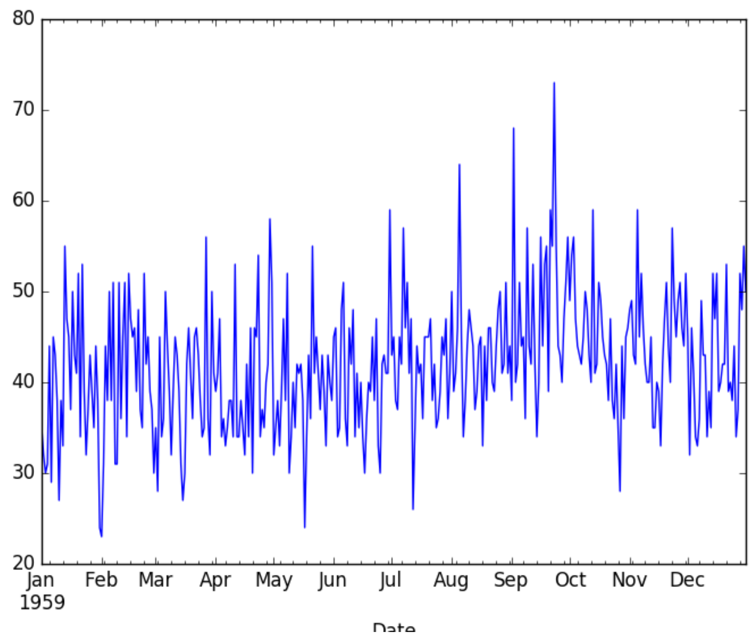

Анализ временных рядов
Прогнозирование временных рядов
Выделение тренда
Поиск аномалий, ошибок
Поиск сезонных компонент
Адаптивные модели
Авторегрессионная модель (AR)
Скользящее среднее (MA)
ARMA, ARIMA

Какими бывают временные ряды

- Стационарные

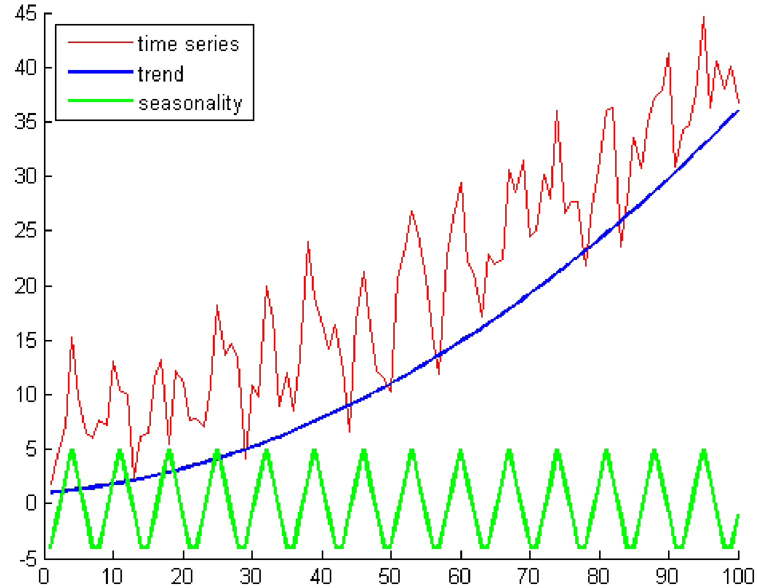


- Нестационарные

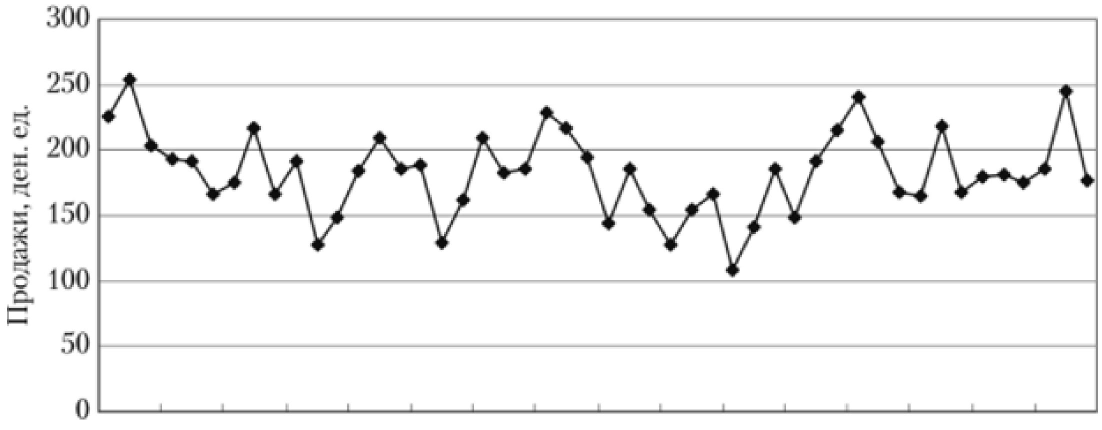


Какими бывают временные ряды

- Сезонность и тренд

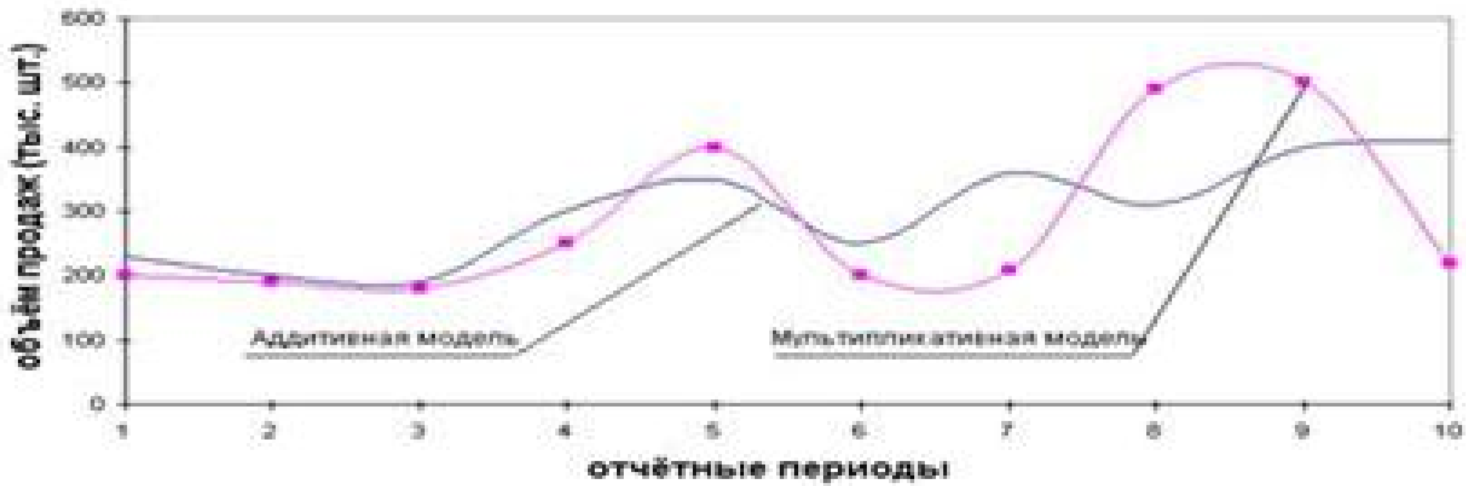


- Без сезонности



Модели для анализа структуры временных рядов

- Аддитивная модель $Y = T + S + E$
- Мультипликативная модель $Y = T \cdot S \cdot E$

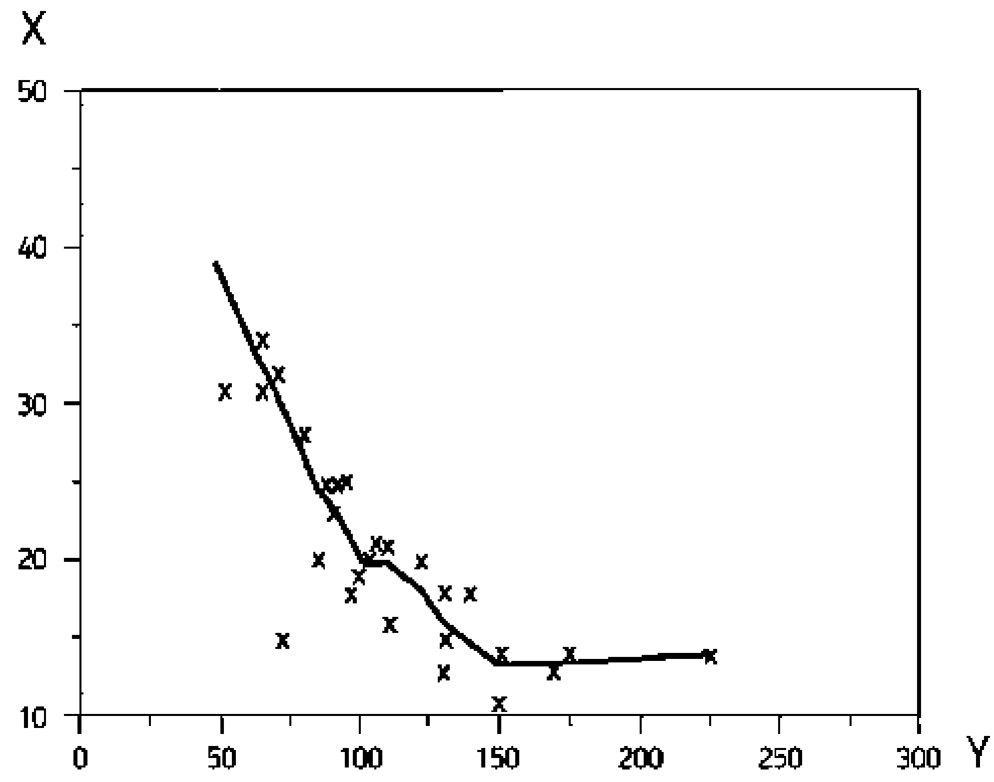


Метод STL

- STL (A Seasonal-Trend Decomposition Procedure Based on Loess) —это процедура декомпозиции временного ряда на сезонную, трендовую составляющие и остатки, которая использует метод локальных регрессий (LOESS).
- Применим только для аддитивных моделей.

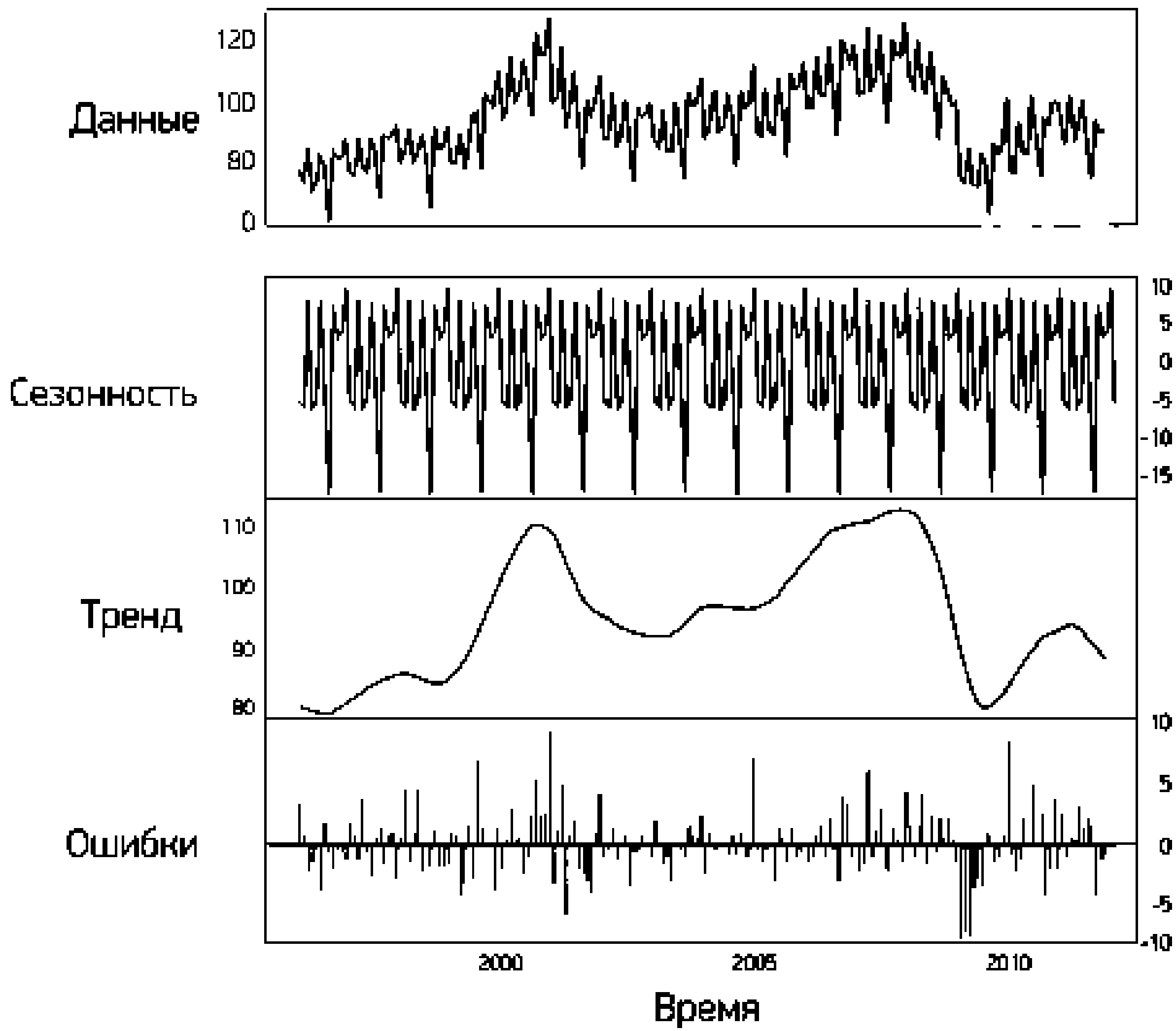
Метод Loess

- LOESS (Locally weighted scatterplot smoothing) — это методика для моделирования и сглаживания двумерных данных. Автор — Уильям Кливленд (1979).



- $X^m = (x_i; y_i)_{i=1}^m$

Что позволяет получить STL



Внутренний цикл STL

1. Вычесть тренд. $T_0^v = 0, p = 0$

$$Y_* = Y_v - T_k^v$$

Y_v — исходный временной ряд
T — тренд
p —веса для loess

2. Сортируем каждый сезонный компонент по периодам, сглаживаем LOESS.

$$C_{k+1}^v \leftarrow loess(Y_*)$$

C—ряд со сглаженными компонентами сезонности

Внутренний цикл STL

3. Глубокое сглаживание сезонных компонент методами MA и LOESS.

$$\overline{L_{k+1}^{v}}$$

L —ряд, полученный после MA, LOESS

4. Детрендирование сглаженных сезонных компонент.

$$S_{k+1}^{v} = C_{k+1}^{v} - L_{k+1}^{v}$$

S —сезонный компонент без тренда.

5. Десезонализация.

$$Y_{**} = Y_{v} - S_{k+1}^{v}$$

Y_{**} —исходный ряд без сезонности.

6. Сглаживание тренда.

$$T_{k+1}^{v} \leftarrow loess(Y_{**})$$

Внешний цикл STL

1. Считаем остатки.

$$R_v = Y_v - T_v - S_v$$

R —остатки
Y —исходный временной ряд
T —сглаженный тренд без сезонности
S —сезонный компонент без тренда
p —веса для LOESS

2. Пересчитываем веса.

$$p_v = B\left(\frac{|R_v|}{6 \cdot median(|R_v|)}\right)$$

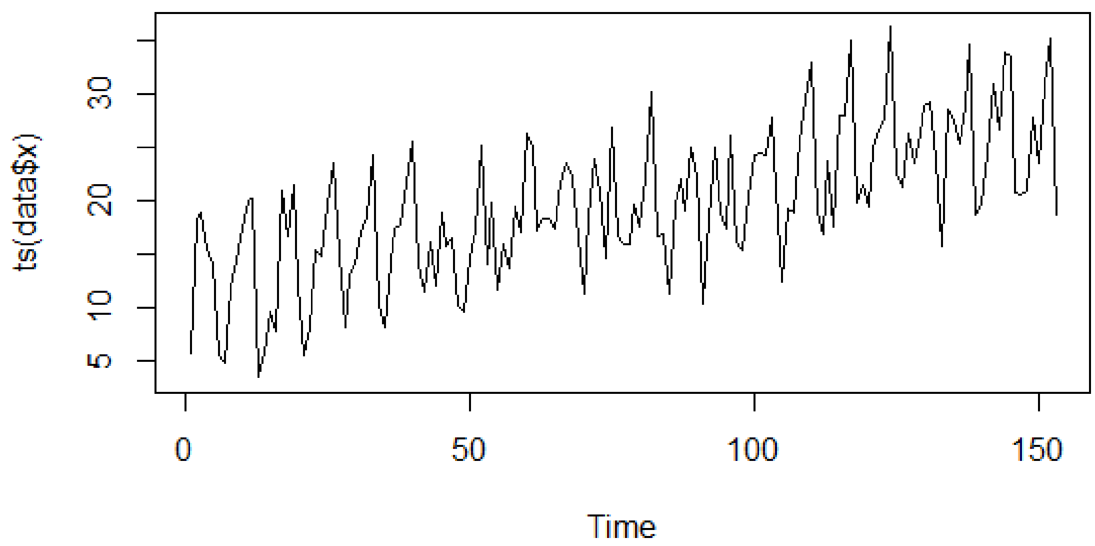
STL разложение с помощью R

Общая информация о данных:

```
> data<-read.csv("C:/dataTimeSeries.csv",sep=";",dec=".")
> summary(data)
      Date                  x
 Min.   :20150101   Min.   : 3.414
 1st Qu.:20150208   1st Qu.:15.617
 Median :20150318   Median :19.084
 Mean   :20150321   Mean   :19.593
 3rd Qu.:20150425   3rd Qu.:24.235
 Max.   :20150602   Max.   :36.167
```

График ряда:

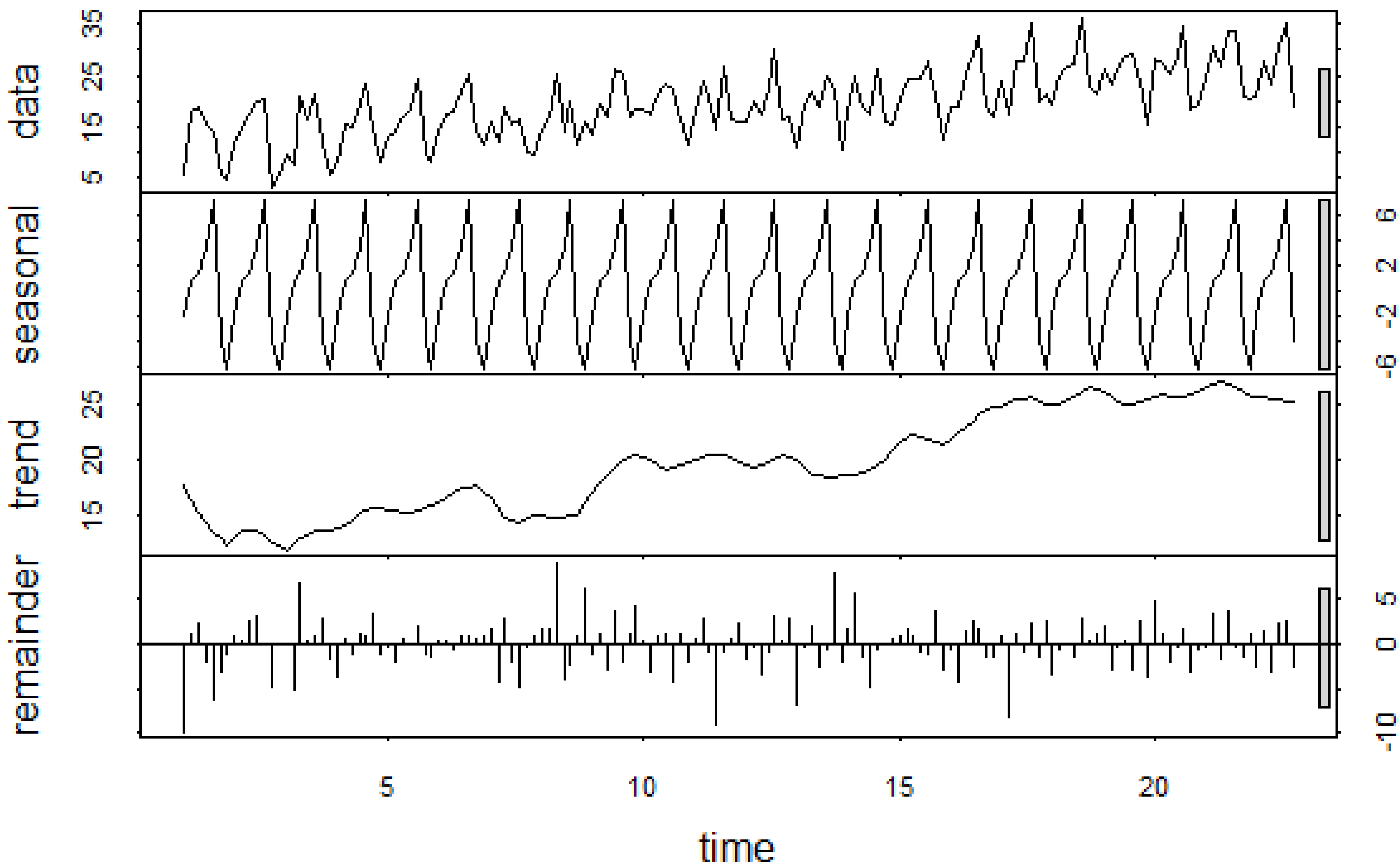
```
> plot(ts(data$x))
```



STL разложение с помощью R

```
> res<-stl((ts(data[,2],frequency=7)),s.window="periodic",robust=TRUE,inner=2)
> plot(res)
```

STL разложение с помощью R



AR-модель

Авторегрессионная (AR-) модель (англ. autoregressive model) — модель временных рядов, в которой значения временного ряда в данный момент линейно зависят от предыдущих значений этого же ряда.

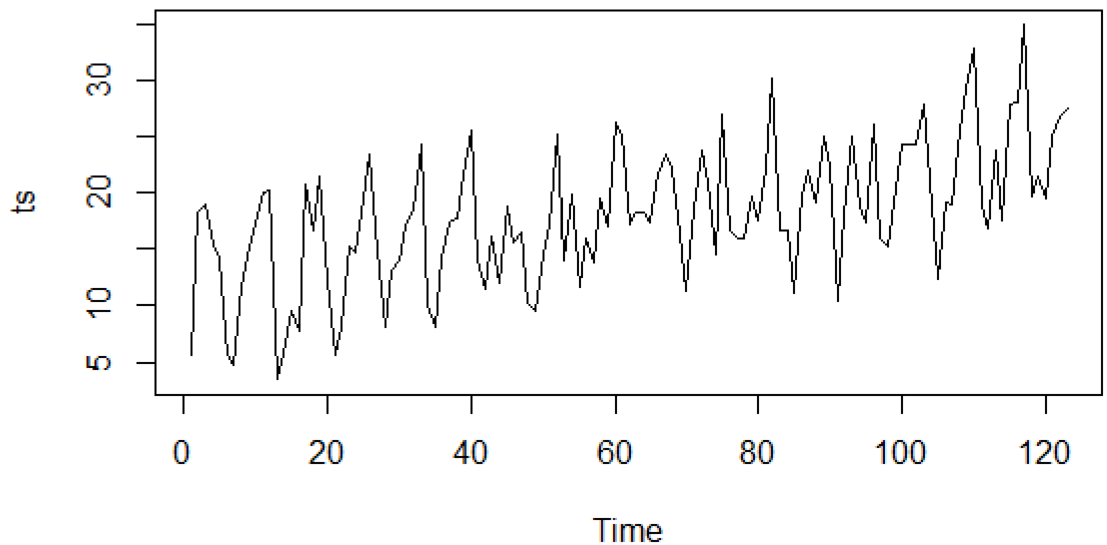
$$X_t = c + \sum_{i=1}^{p} a_i X_{t-i} + \varepsilon_t$$

- *AR(p)* — авторегрессионная модель
- *p* — порядок регрессии
- *a1, …, ap* — параметры модели
- *c* — постоянная
- *εt* — белый шум

Построение авторегрессионной модели прогнозирования временных рядов с помощью R.

Данные 30 последних наблюдений:

```
data<-read.csv("C:/dataTimeSeries.csv",sep=";",dec=".")
ts <- ts(head(data$x, (nrow(data)-30)))
newts <- ts(tail(data$x),30)
plot.ts(ts)
```



Построение авторегрессионной модели прогнозирования временных рядов с помощью R.

Авторегрессия первого порядка:

```
> model_ar1 <- ar(ts, aic=FALSE, order.max=1)
> model_ar1

Call:
ar(x = ts, aic = FALSE, order.max = 1)

Coefficients:
     1
0.4565

Order selected 1  sigma^2 estimated as  29.76
```

Прогноз модели на 30 последних точках:

```
> forecast_ar1 <- predict(model_ar1,n.ahead=30)
> forecast_ar1
$pred
Time Series:
Start = 124
End = 153
Frequency = 1
 [1] 22.32835 19.97937 18.90715 18.41772 18.19432 18.09234 18.04579 18.02455 18.01485
[10] 18.01042 18.00840 18.00748 18.00706 18.00686 18.00678 18.00674 18.00672 18.00671
[19] 18.00671 18.00670 18.00670 18.00670 18.00670 18.00670 18.00670 18.00670 18.00670
[28] 18.00670 18.00670 18.00670

$se
Time Series:
Start = 124
End = 153
Frequency = 1
 [1] 5.455600 5.997085 6.103864 6.125877 6.130454 6.131407 6.131606 6.131647 6.131656
[10] 6.131658 6.131658 6.131658 6.131658 6.131658 6.131658 6.131658 6.131658 6.131658
[19] 6.131658 6.131658 6.131658 6.131658 6.131658 6.131658 6.131658 6.131658 6.131658
[28] 6.131658 6.131658 6.131658
```

Построение авторегрессионной модели прогнозирования временных рядов с помощью R.

Авторегрессия порядка 10:

```
> model_ar <- ar(ts, aic=TRUE, order.max=10)
> model_ar

Call:
ar(x = ts, aic = TRUE, order.max = 10)

Coefficients:
      1        2        3        4        5        6        7  
 0.1948   0.0506  -0.0196  -0.0077  -0.0855   0.2743   0.4508  

Order selected 7  sigma^2 estimated as  19.4
```

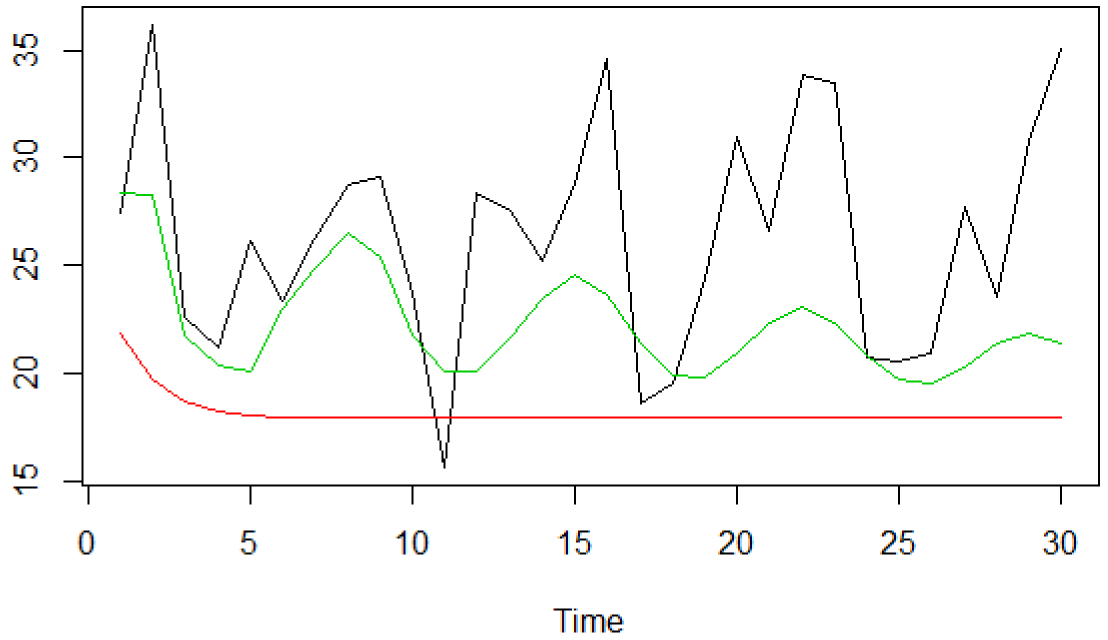
Прогноз модели на 30 последних точках:

```
> forecast_ar <- predict(model_ar, n.ahead=30)
> forecast_ar
$pred
Time Series:
Start = 124 
End = 153 
Frequency = 1 
 [1] 27.93185 21.78033 20.30328 20.18464 23.12244 24.70569 26.17839 25.13363 21.78264
[10] 20.08628 20.21245 21.85792 23.45111 24.51907 23.52922 21.35964 19.98284 19.98811
[19] 21.05729 22.40543 23.12804 22.39229 20.89458 19.83035 19.73655 20.48958 21.53287
[28] 22.03622 21.52144 20.46121

$se
Time Series:
Start = 124 
End = 153 
Frequency = 1 
 [1] 4.404754 4.487573 4.504495 4.504615 4.504683 4.521869 4.642924 5.214321 5.312126
[10] 5.334414 5.336076 5.339707 5.341555 5.433516 5.635768 5.711306 5.734083 5.735495
[19] 5.738690 5.738790 5.790287 5.882848 5.935504 5.954269 5.955240 5.956481 5.9573[illegible]1
[28] 5.984929 6.033392 6.068584
```

Построение авторегрессионной модели прогнозирования временных рядов с помощью R.

```
> plot(ts(data.frame(data=newts, ar1=forecast_ar1$pred, ar=forecast_ar$pred)),
+       plot.type="single", col=1:3)
```



СПАСИБО ЗА ВНИМАНИЕ