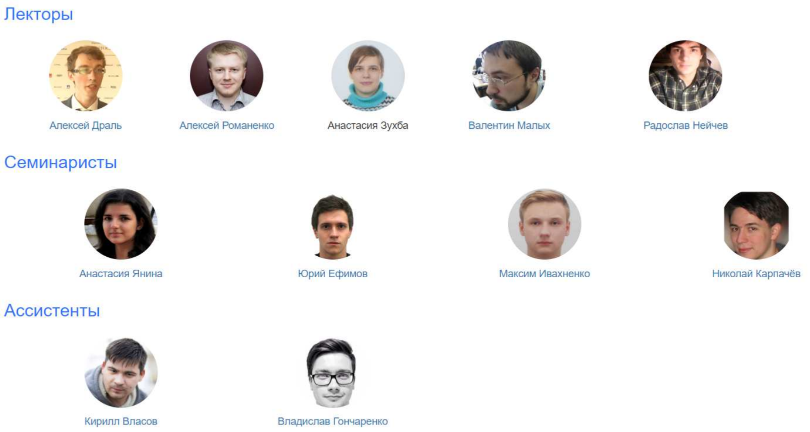
Time Series Analysis.
0. Intro into TS

Alexey Romanenko
alexromsput@gmail.com

FIVT MIPT, September 2018

Содержание

The Course team



site of the course with full info about: `https://ml-mipt.github.io/`

Let me introduce: Alexey Romanenko

Education:

- Masters Degree at MIPT, 2011
- PhD thesis: Composition of Time Series Forecasting Algorithm at MIPT, tbd

Work experience:

- SAS Institute (software implementation), 2016 – now
- MIPT (teaching assistant), 2011 – now
- Svyaznoy (one of the largest cellphone retailer in Russia), 2010–2016
- Forecsys (Machine Learning software for business), 2008–2010

View full profile: `https://www.linkedin.com/in/alexromsput/`

Date	Topic
05/09/2018	Intro in TS forecasting
12/09/2018	Exponential Smoothing models, Comparing Models
19/09/2018	ARIMAX and other autoregression models (ARCH, GARCH)
26/09/2018	Compositions, Hierarchial Forecasting, TS Forecasting with NN

Date	Topic
05/09/2018	Intro in TS forecasting
12/09/2018	Exponential Smoothing models, Comparing Models
19/09/2018	ARIMAX and other autoregression models (ARCH, GARCH)
26/09/2018	Compositions, Hierarchial Forecasting, TS Forecasting with NN

You will have to come through:

1. 4 seminars
2. 2 HW (for 1-2 hours)

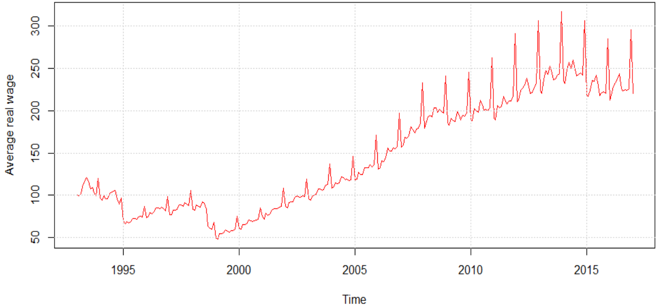
Date	Topic
05/09/2018	Intro in TS forecasting
12/09/2018	Exponential Smoothing models, Comparing Models
19/09/2018	ARIMAX and other autoregression models (ARCH, GARCH)
26/09/2018	Compositions, Hierarchial Forecasting, TS Forecasting with NN

1. 4 seminars
2. 2 HW (for 1-2 hours)

- It is interesting!
- It is difficult!
- We will do it!

Time Series definition

Time series: $y_1, \dots, y_T, \dots, \quad y_t \in \mathbb{R}$, — a sequence of values of some variable, detected in a constant time interval.

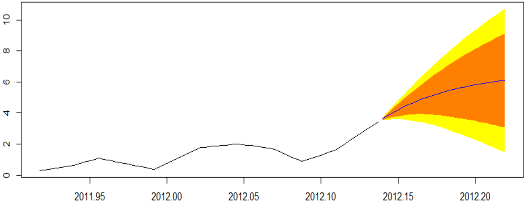


Time series forecasting task — find function f_T:

$$y_{T+d} \approx f_T\left(y_T, \dots, y_1, d\right) \equiv \hat{y}_{T+d|T},$$

where $d \in \{1, \dots, D\}$ — delay, D — horizon.

Forecasting interval, confidence of the forecast

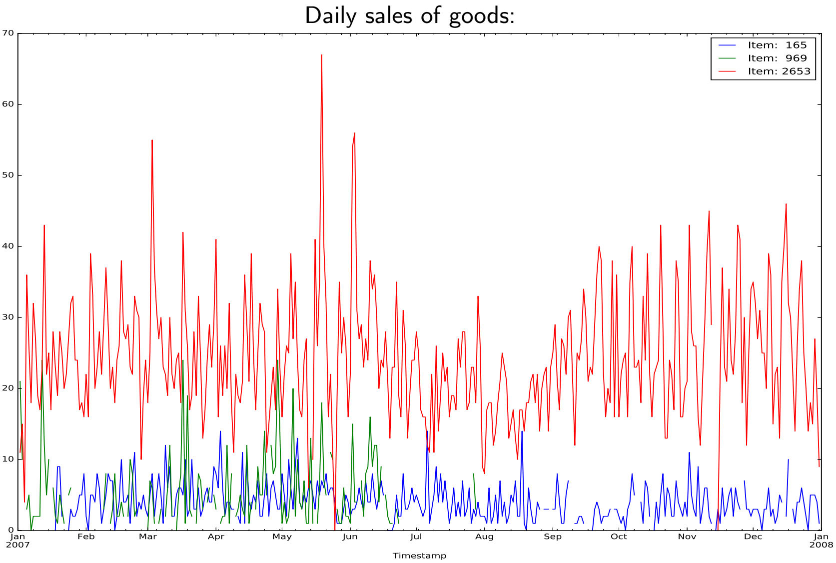


Example: April 1997, Grabd-Forks, ND, unexpected flood:
`https://www.youtube.com/watch?v=0iJUgddua-g`
The city was protected by dam of 51 feet;
The National Weather Service (NWS) had forecast that the river would crest at 49 feet
The river crested at 54.35 feet.

50000 citizens were evacuated, 75% buildings were damaged or destroyed,
Property damage $\$3.5$ billion

The forecasting interval was ± 9 feet

Time series in Retail



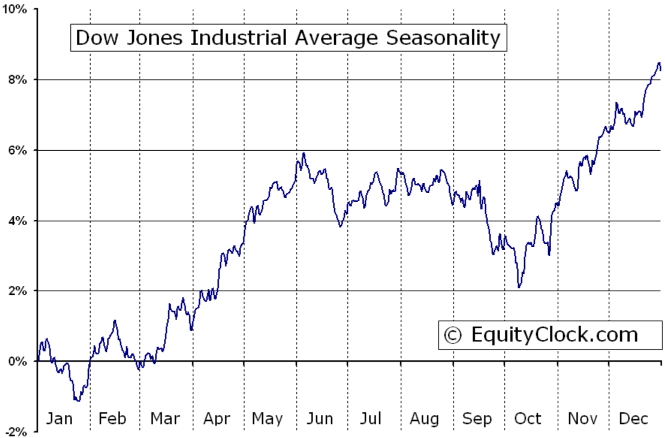
Time series in Retail

Specification of the TS in Retail

- $10^6 - 10^8$ TS to forecast at once
- missing in data
- out-of-stock (no sales by non-zero demand)
- dependence on promo-events, changes in price
- complex loss-function

Time series in Finance

Index Daw-Jones:



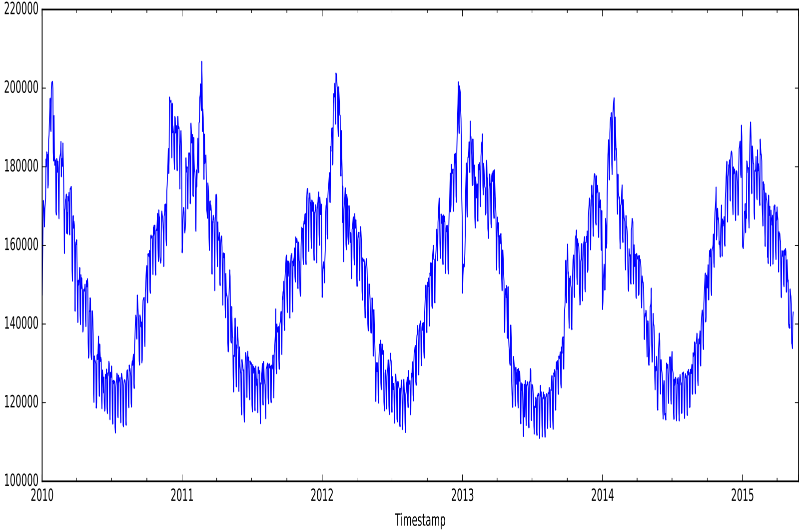
Time series in Finance

Specification of the TSA in Finance:

- high level of noise
- correlation with other financial index
- highly dependence on external events (big market events, politics)

Time series in Industry: Consumption of Electricity

Volume of Consumption of Electricity:



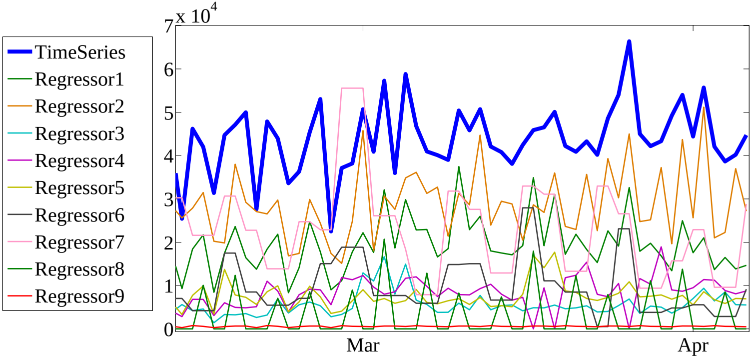
Time series in Industry: Consumption of Electricity

Specification of the TSA

- complex structure (yearly, weekly, daily – seasonality)
- dependence on temperature, price, calendar-events
- nonlinear dependence

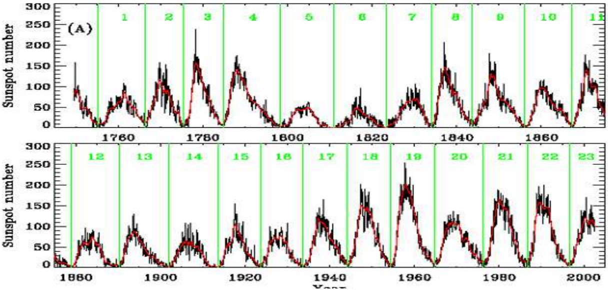
Time series in Industry: Manufacturing

Total man-hours in Warehouse:



- depends mainly on external factors
- can be described by clear physical model

Time series in Physics



- a-periodical changes
- complex physical model of dependence (Newton's laws, Kepler's laws, . . .)

Components of Time Series

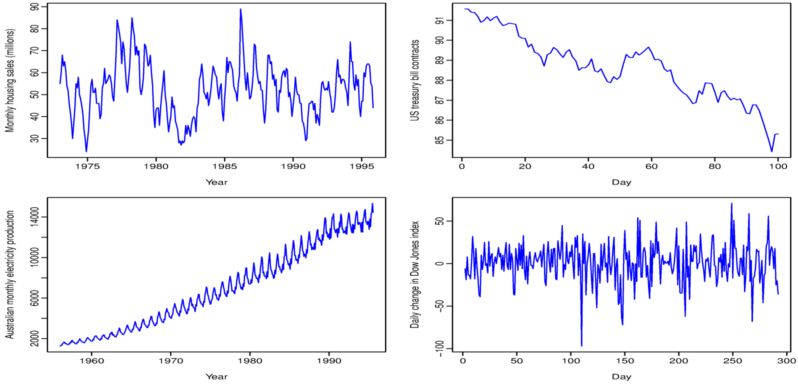
Level — average level of values;
Trend — monotonic long-term changes of Level;

Seasonality — periodical changes of values with constant period;

Cycle — changes in time series values (economical cycles, solar activity cycles).

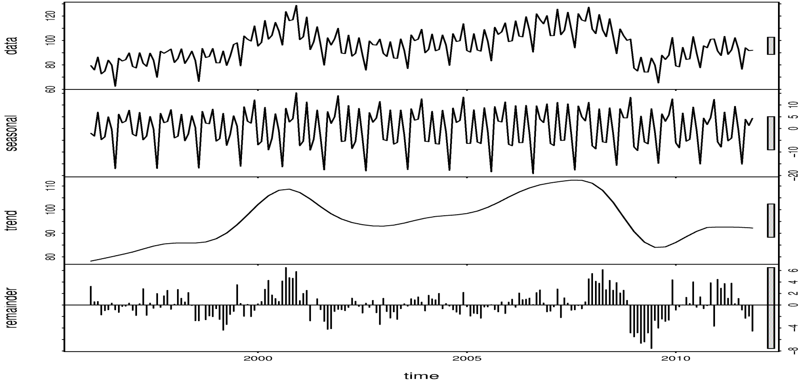
Error — random (unbiased) component of time series.

Components of Time Series



Components of Time Series

STL-decomposition:



Overview of statistical models for TS forecasting

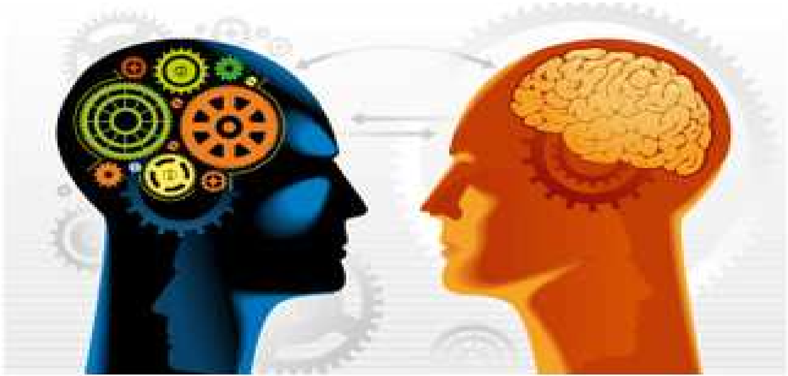
- Exponential Smoothing Models
- ARMA, ARIMA, ARIMAX, SARIMAX
- Dynamic Autoregressive models
- ARCH, GARCH, EGARCH ...
- VAR (vector autoregression)
- Gaussian State Space Models (UCM)
- GAS, GASX (generalized autoregression)
- Quantile Regression
- Adaptive Composition
- Adaptive Selection
- Гусеница [Голяндина, 2003]
- Aggregating Algorithm
- Neural Nets

Jonathan D. Cryer, Kung-Sik Chan Time Series Analysis With Applications in R. Second Edition. Springer, 2008

Магнус Я.Р. и др. Эконометрика. Начальный Курс М.: Дело, 2007

Python package for TS `http://www.pyflux.com/docs/index.html`

Stat Models VS Classic Machine Learning



Conclusion

- time series differs in business regions
- time series forecasting problem has features

Conclusion

- time series differs in business regions
- time series forecasting problem has features

Pro Stats-models:

- forecast of statistical models can be easily interpreted
- stats models are equivalent to ML algorithms in terms of accuracy of forecast
- ensembling of statisitcal models and ML algorithm usually leads to improving forecast accuracy
- a statistical model can build forecast for any horizen (provided all causal variables are given)

Conclusion

- time series differs in business regions
- time series forecasting problem has features

Pro Stats-models:

- forecast of statistical models can be easily interpreted
- stats models are equivalent to ML algorithms in terms of accuracy of forecast
- ensembling of statisitcal models and ML algorithm usually leads to improving forecast accuracy
- a statistical model can build forecast for any horizen (provided all causal variables are given)

Cons Stats-models:

- high risk of overfitting
- are not stable
- need additional review

Your fitback

Leave your fitback here `https://goo.gl/forms/TpY4aaojXLszGPQy2`

If you have questions or suggestions you can write me: alexromsput@gmail.com

Literature

Hyndman R.J., Athanasopoulos G. Forecasting: principles and practice. — OTexts, https://www.otexts.org/book/fpp

Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования временных рядов. Финансы и статистика, 2003, http://www.arshinov74.ru/files/files/3.pdf