Machine Learning

Time Series

Jeff Abrahamson

9, 16 décembre 2016 13 janvier, 9 mars 2017

This is hard, but it depends on your goals. And on context.

Introduction to time series

Definition (discrete time series):

$$\{s_t \mid t \in \mathbb{R}^+ \wedge s \in \mathbb{R}\}$$

(though s in any vector space is fine)

Introduction to time series

Examples domains:

- Weather
- Economics
- Industry (e.g., factories)
- Medicine
- Web
- Biological processes

Introduction to time series

Why?

- Predict
- Control
- Understand
- Describe

Introduction to time series

Some strategies:

- Clustering
- Hidden Markov Models (HMM)
- Recurrent neural networks (RNN)
- Autoregressive integrated moving average

Introduction to time series

One model:

$$s_t = g(t) + \phi_t$$

where

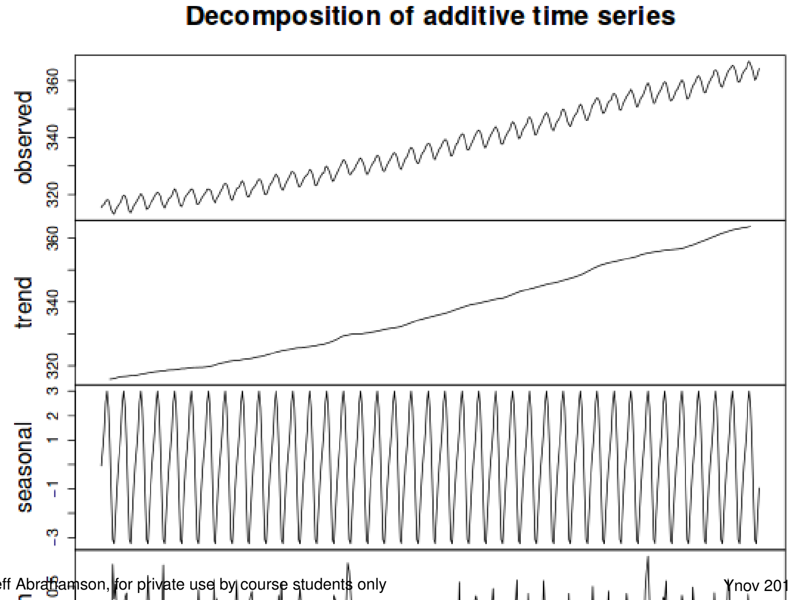
$g(t)$ is deterministic: signal (or trend)

ϕ_t is stochastic noise

Introduction to time series

Variation types:

- Trend (g)
- Seasonal effect (g)
- Irregular fluctuation (residuals: ϕ)


Decomposition of additive time series
observed
trend
seasonal
360
340
320
3
2
1
−1
−3

Introduction to time series

Some easy things to try

- Introduce features to break out seasonality
- Introduce lags as features
- Some domain-specific transformation

HMM

“simplest dynamic Bayesian network”

Markov Chains

A **Discrete time Markov chain (DTMC)** is a random process that undergoes state transitions.

Markov Chains

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & xnn \end{bmatrix} \begin{bmatrix} v_1^{(i)} \\ v_2^{(i)} \\ \vdots \\ v_n^{(i)} \end{bmatrix} = \begin{bmatrix} v_1^{(i+1)} \\ v_2^{(i+1)} \\ \vdots \\ v_n^{(i+1)} \end{bmatrix}$$

Markov Chains

$$Xv_i = v_{i+1}$$

Markov Chains

Examples:

- Random walks
- Weather (first approximation in many places)
- Thermodynamics
- Queuing theory (so also telecommunications)
- Spam

Markov Chains

Properties:

- Stochastic process
- Memoryless ("Markov property")

HMM's

- State is not visible
- Output of state is visible

Examples: noisy sensor, medical diagnosis

HMM's

What we have:

- State space $S = \{s_1, \ldots, s_n\}$
- Observation space $O = \{o_1, \ldots, o_k\}$
- Transition matrix A of size $n \times n$
- Emission matrix B of size $n \times k$
- Initial state probabilities $\pi = \{\pi_1, \ldots, \pi_n\}$
- A sequence of observations $X = \{x_1, \ldots x_T\}$

Here

- $y_t = i \iff$ observation at time t is o_i
- $\Pr(x_1 = s_i) = \pi_i$

We want the sequence of states $X = \{x_1, \ldots, x_T\}$.

HMM's

Some pointers to learn more about HMM:

- Forward-Backward Algorith
- Viterbi Decoding
- Baum-Welch Algorithm


questions?