Probabilistic Inference and Learning
Lecture 08
Gaussian Process Regression: An Extensive Example

Philipp Hennig

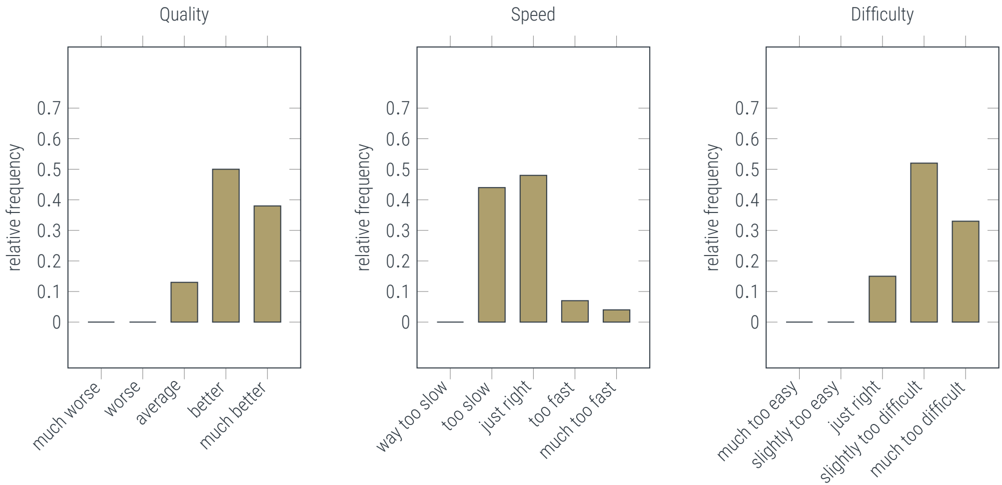
12 November 2018

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Faculty of Science
Department of Computer Science
Chair for the Methods of Machine Learning

Last Lecture: Debrief

Feedback dashboard



Last Lecture: Debrief

Detailed Feedback

Things you did not like:

- **the math**
- not enough math!
- "please upload slides before lecture"
- "too fast toward the end"

Things you did not understand:

- **the math**
- the RKHS bit
- the final third

Things you enjoyed:

- **the math**
- **explanations at beginning**
- frequent repetition of "simple" concepts for added clarity
- that the slides were uploaded before the lecture
- links to background
- the RKHS explanations

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Overview of Lectures so far:

0. Introduction to Reasoning under Uncertainty
1. Probabilistic Reasoning
2. Probabilities over Continuous Variables
3. Gaussian Probability Distributions
4. Gaussian Parametric Regression
5. More on Parametric Regression — Connections to Deep Learning
6. Gaussian Processes
7. More on Kernels & GPs

Today:

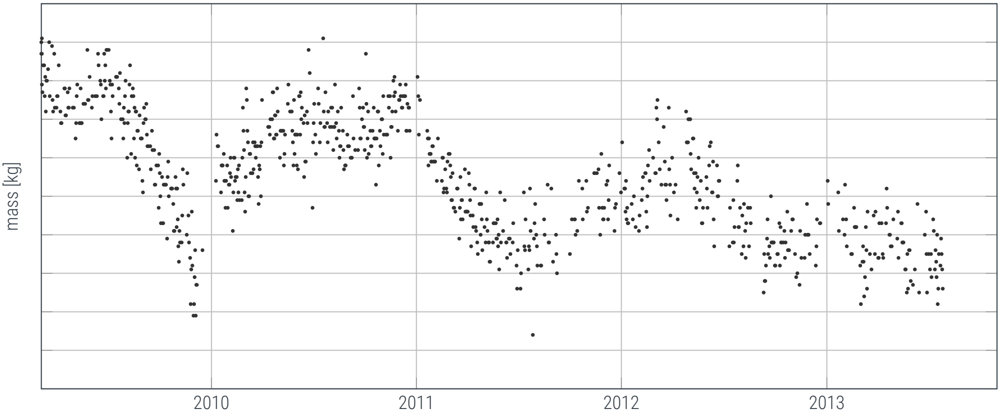
- a concrete example of GP regression

today's lecture:

- getting our hands dirty with a simple but realistic problem
- mostly code
- a simple (1D) dataset, but with significant structure

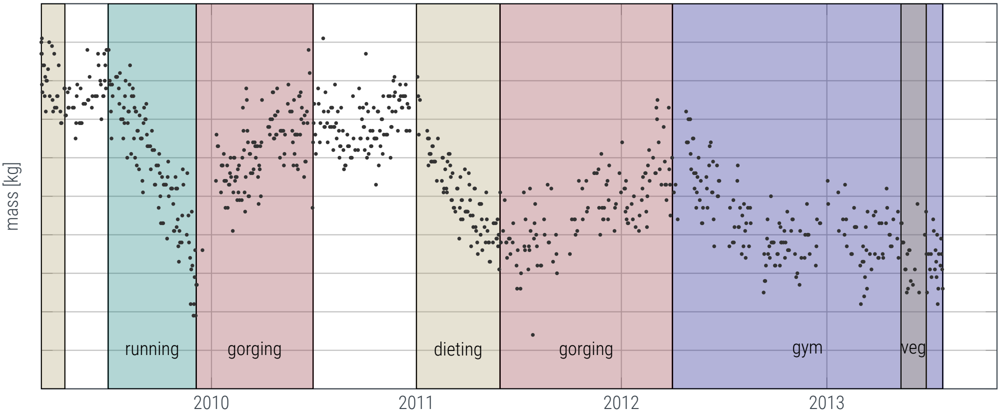
A Dataset

(c) P. Hennig, 2007–2013



A Dataset

(c) P. Hennig, 2007–2013



http://papers.nips.cc/paper/6313-bayesian-intermittent-demand-forecasting-for-large-inventories

Bayesian Intermittent Demand Forecasting for Large Inventories

Matthias Seeger, David Salinas, Valentin Flunkert
Amazon Development Center Germany
Krausenstrasse 38
10115 Berlin
matthis@amazon.de, dsalina@amazon.de, flunkert@amazon.de

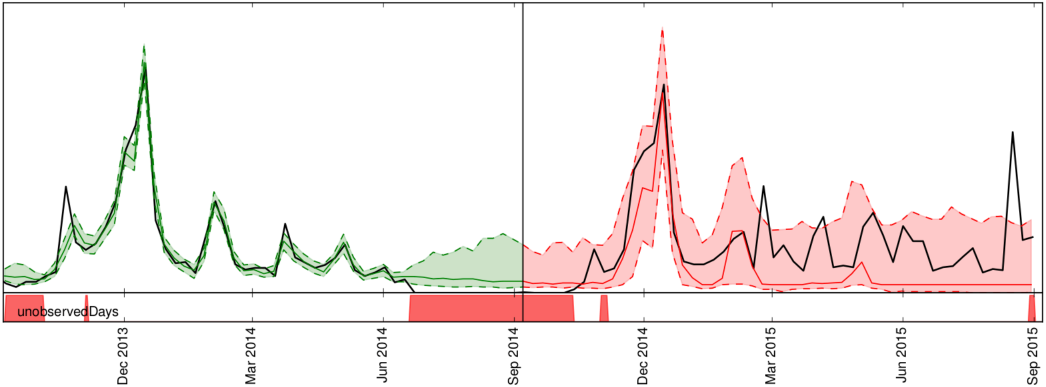
Abstract

We present a scalable and robust Bayesian method for demand forecasting in the context of a large e-commerce platform, paying special attention to intermittent and bursty target statistics. Inference is approximated by the Newton-Raphson algorithm, reduced to linear-time Kalman smoothing, which allows us to operate on several orders of magnitude larger problems than previous related work. In a study on large real-world sales datasets, our method outperforms competing approaches on fast and medium moving items.

Matthias Seeger
Principal Applied Scientist, Amazon
MPI Tübingen, 2006–2011

A Realistic Version

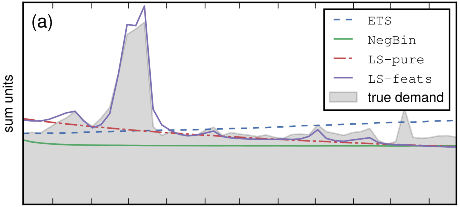
http://papers.nips.cc/paper/6313-bayesian-intermittent-demand-forecasting-for-large-inventories

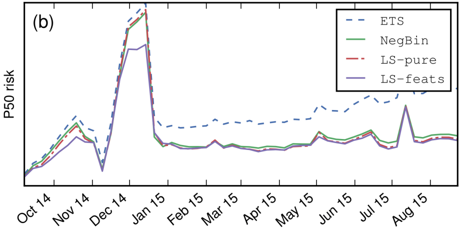


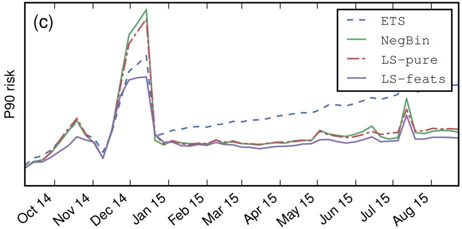
A Realistic Version

http://papers.nips.cc/paper/6313-bayesian-intermittent-demand-forecasting-for-large-inventories

	Parts	EC-sub	EC-all
# items	19874	39700	534884
Unit t	month	day	day
Median CV^2	2.4	5.8	9.7
Freq. $z_t = 0$	54%	46%	83%
In-stock ratio	100%	73%	71%
Avg. size series	33	329	293
# item-days	656K	13M	157M







Summary:

- An unstructured kernel regression model can only do so much. **Extrapolation** and extracting **structural knowledge** require prior knowledge about **causal process**
- Linear models with elaborate features can be quite expressive, while remaining interpretable (try doing this example with a deep network!)
- Physical processes have **units**
- Complicated processes require complicated (and questionable!) prior assumptions
- analogous process in business environments
  - demand and supply **forecasting**
  - financial engineering
  - ad placement (with minor variations)

The ability to build structured predictive models is a **key skill**. Everyone can run a TensorFlow script! Masters of structured probabilistic inference are highly sought after.