Probabilistic Machine Learning
Lecture 11
Gaussian Process Regression:
An Extensive Example

Philipp Hennig

25 May 2020

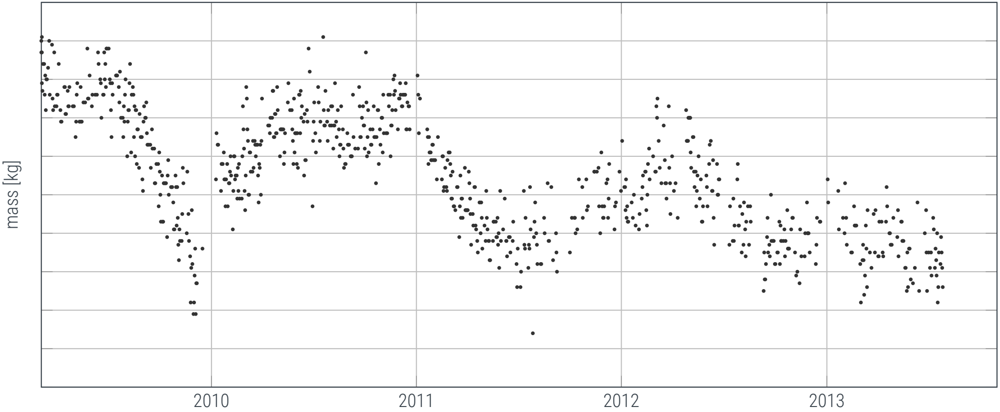
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

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

#	date	content	Ex	#	date	content	Ex
1	20.04.	Introduction	1	14	09.06.	Logistic Regression	8
2	21.04.	Reasoning under Uncertainty		15	15.06.	Exponential Families	
3	27.04.	Continuous Variables	2	16	16.06.	Graphical Models	9
4	28.04.	Monte Carlo		17	22.06.	Factor Graphs	
5	04.05.	Markov Chain Monte Carlo	3	18	23.06.	The Sum-Product Algorithm	10
6	05.05.	Gaussian Distributions		19	29.06.	Example: Topic Models	
7	11.05.	Parametric Regression	4	20	30.06.	Mixture Models	11
8	12.05.	Learning Representations		21	06.07.	EM	
9	18.05.	Gaussian Processes	5	22	07.07.	Variational Inference	12
10	19.05.	Understanding Kernels		23	13.07.	Example: Topic Models	
11	25.05.	An Example for GP Regression	6	24	14.07.	Example: Inferring Topics	13
12	26.05.	Gauss-Markov Models		25	20.07.	Example: Kernel Topic Models	
13	08.06.	GP Classification	7	26	21.07.	Revision	

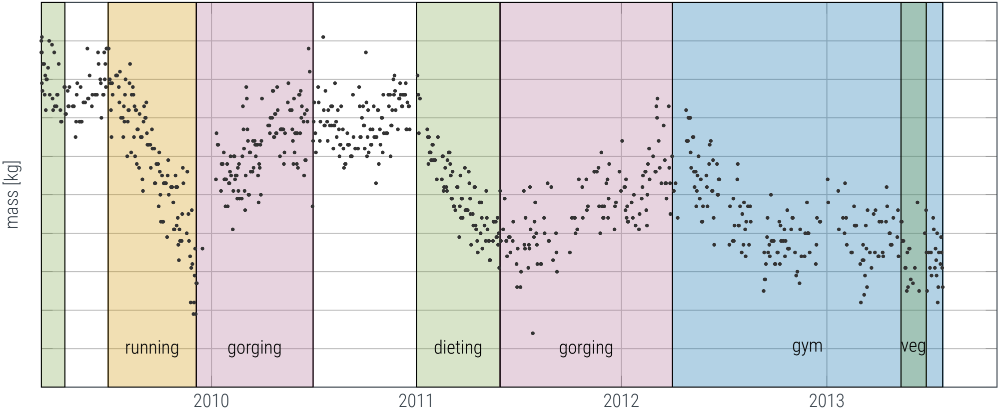
A Dataset

(c) P. Hennig, 2007–2013



A Dataset

(c) P. Hennig, 2007–2013



Seeger et al. NeurIPS 2016

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 ML Scientist, Amazon
MPI Tübingen, 2006–2011

Summary:

- An unstructured kernel regression model can only do so much. **Extrapolation** and extracting **structural knowledge** require prior knowledge about the **causal structure**.
- 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.