TV Ads Attribution and Gaussian Processes

Adrin Jalali

November 16, 2016

Problem Definition

- Website

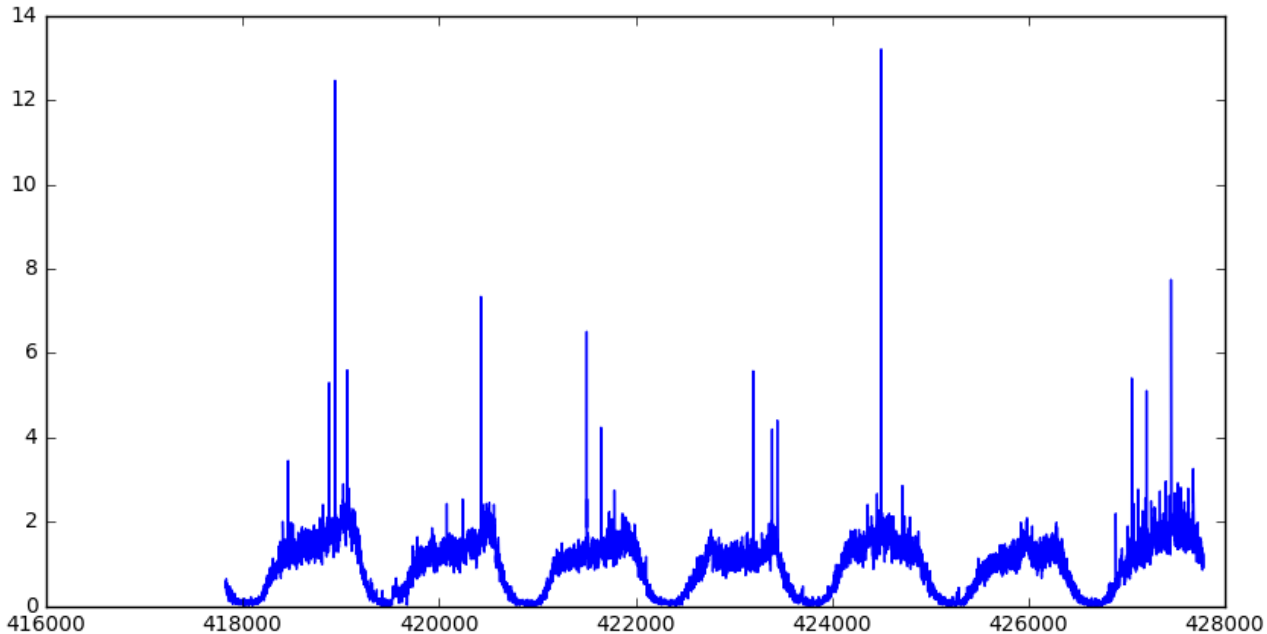
Problem Definition

- Website
- Sources of traffic
  - TV Ads
  - Google Ads
  - ...

Problem Definition

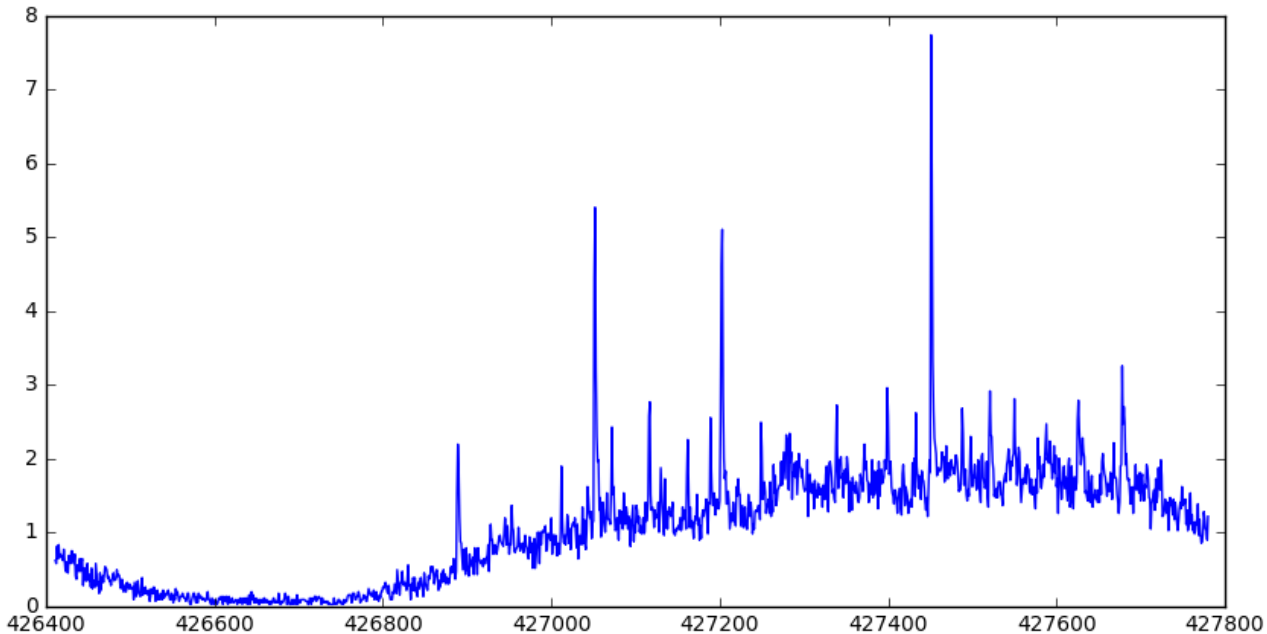
- Website
- Sources of traffic
  - TV Ads
  - Google Ads
  - ...
- How much those campaigns influence the website's traffic?

Data



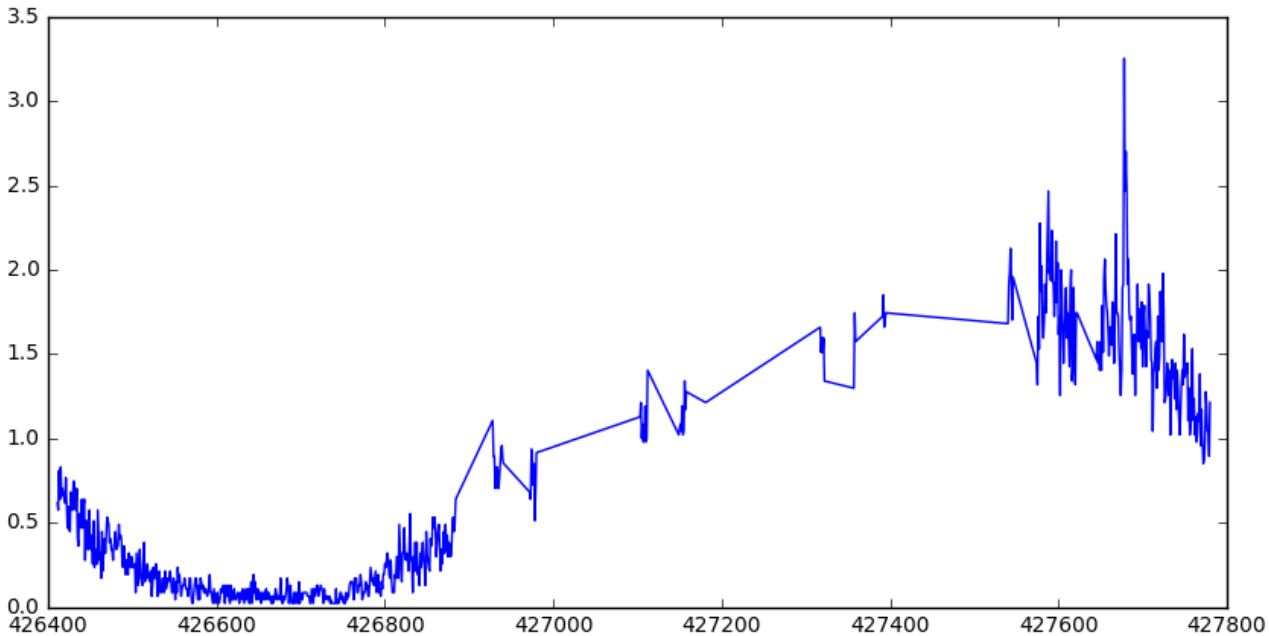
Normalized session count for a week

Data



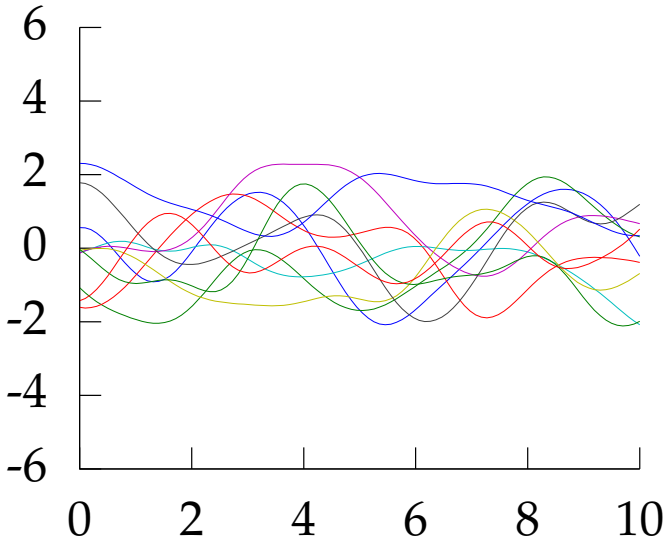
Normalized session count for a day

Training Data



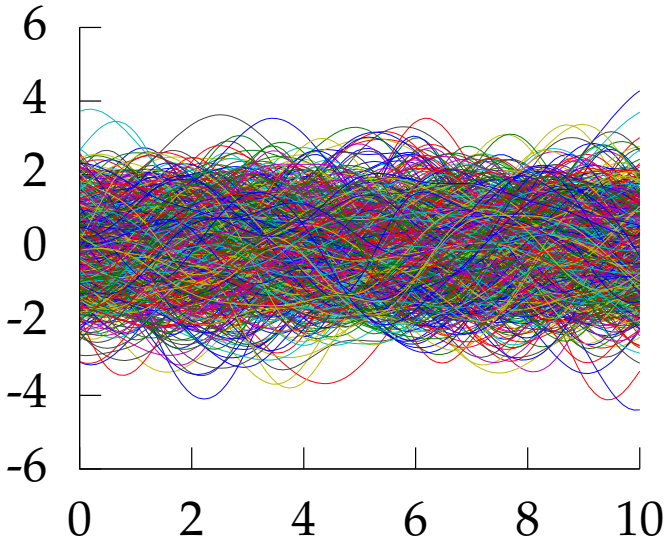
Normalized session count for a day, after removing data around reported events

Gaussian Processes: Extremely Short Overview



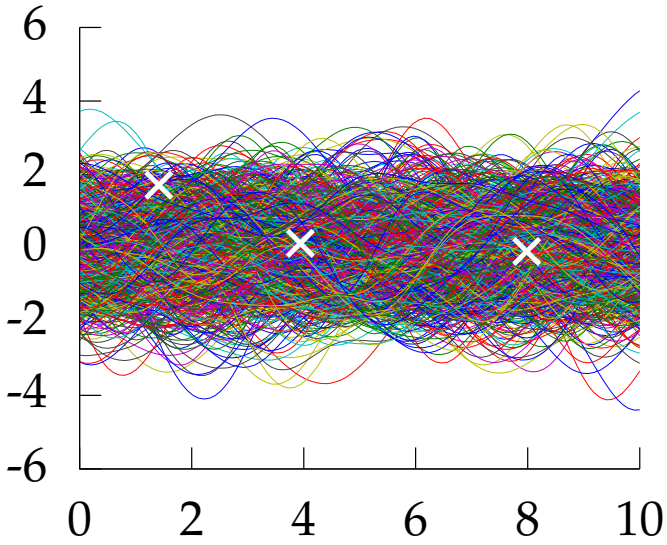
http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Processes: Extremely Short Overview



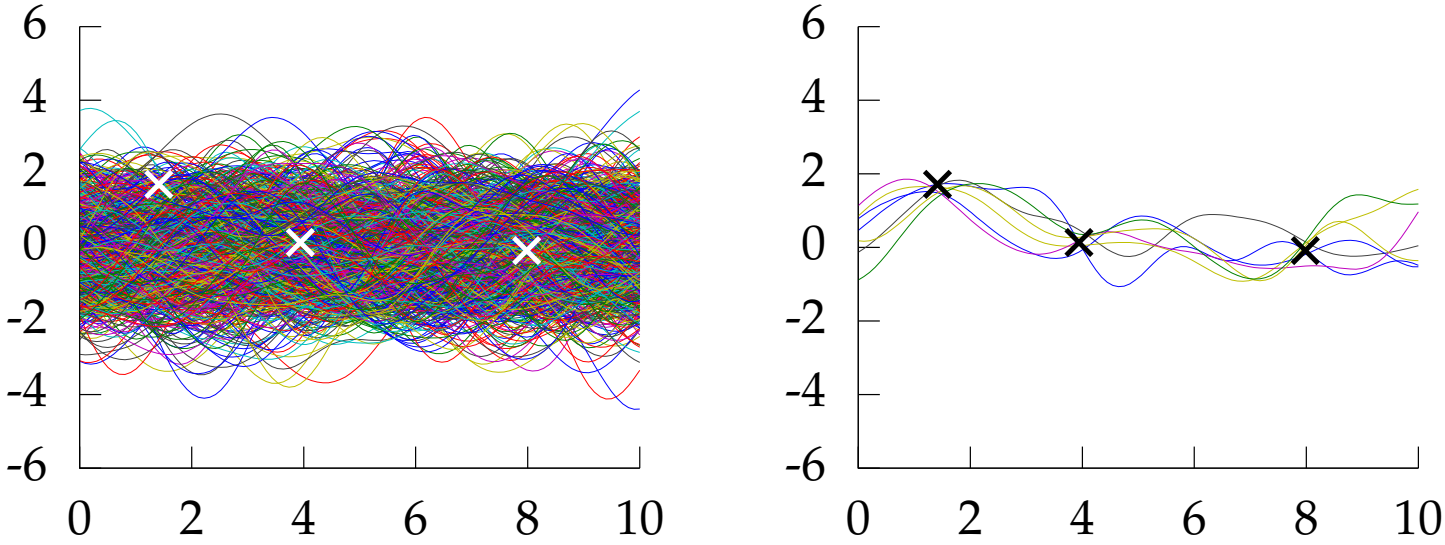
http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Processes: Extremely Short Overview



http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Processes: Extremely Short Overview



http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Process Regression

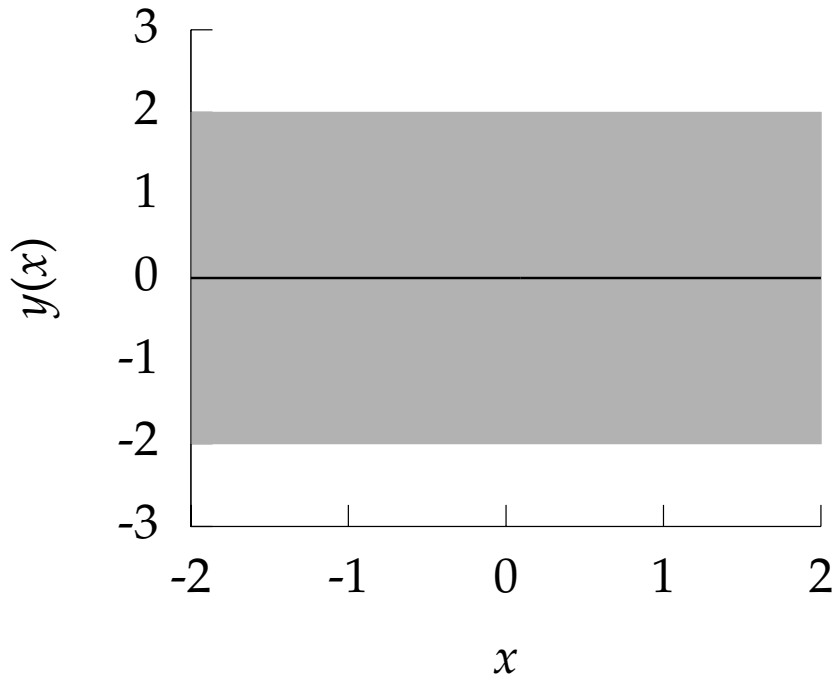


Figure: Examples include WiFi localization, C14 callibration curve.

http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Process Regression

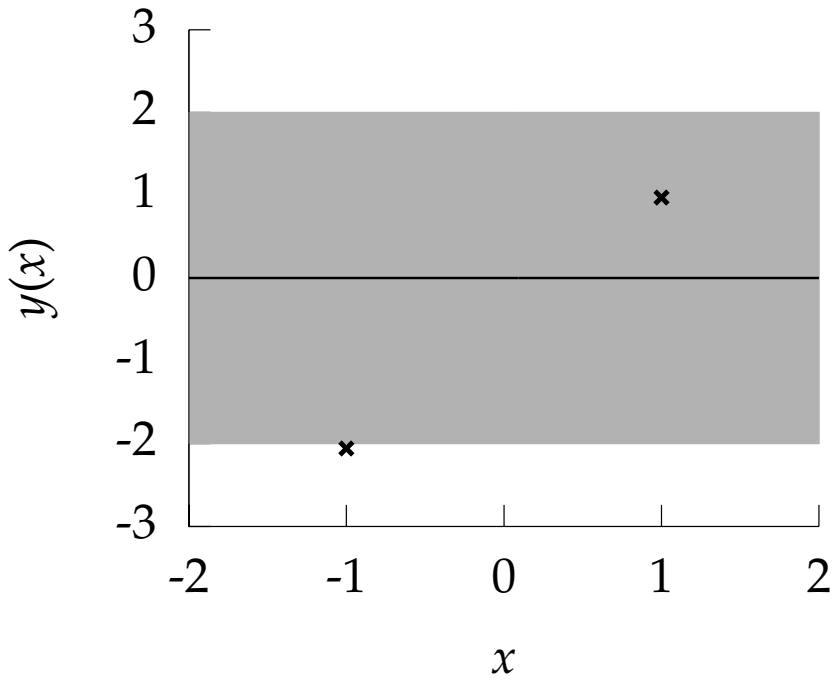


Figure: Examples include WiFi localization, C14 callibration curve.

http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Process Regression

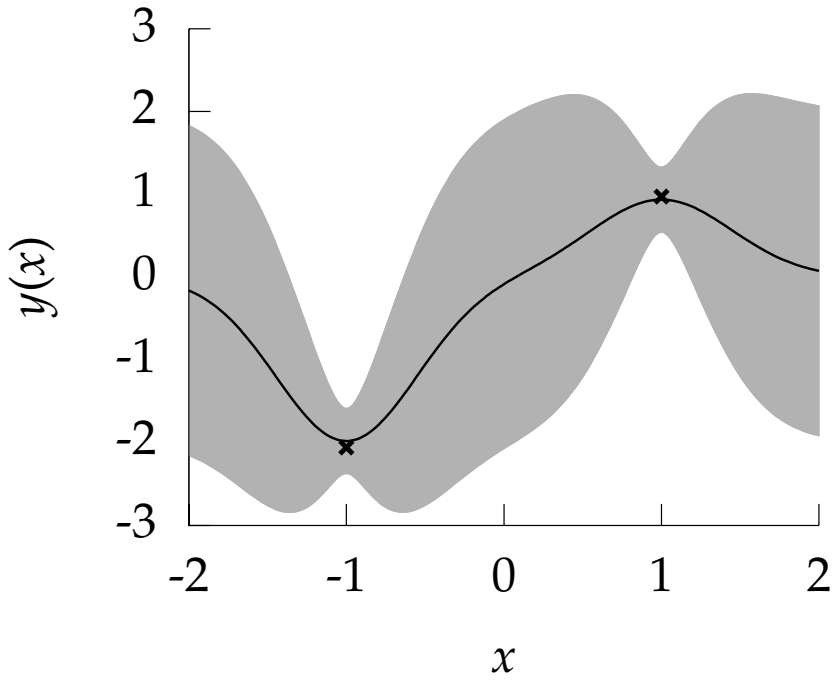


Figure: Examples include WiFi localization, C14 callibration curve.

http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Process Regression

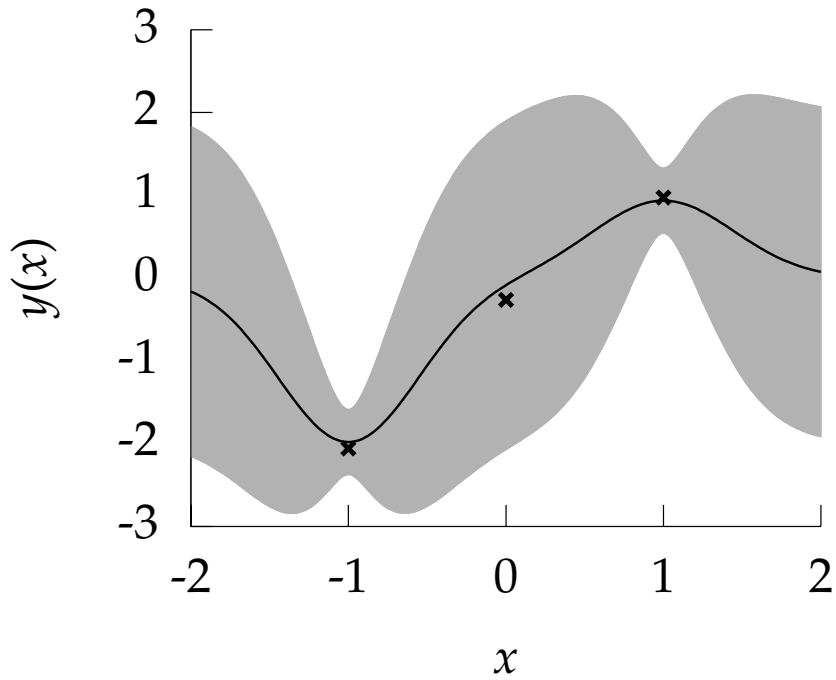


Figure: Examples include WiFi localization, C14 callibration curve.

http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Process Regression

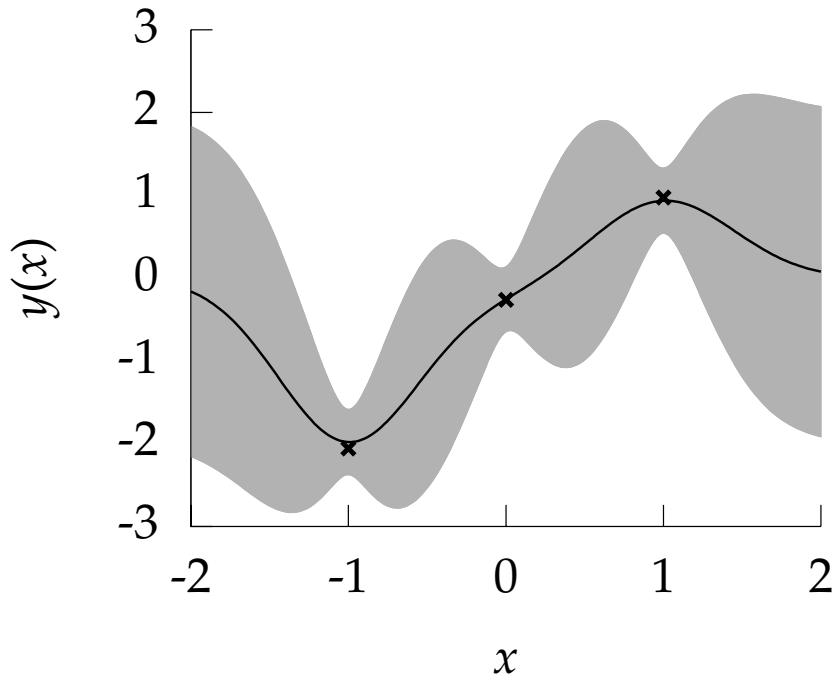


Figure: Examples include WiFi localization, C14 callibration curve.

http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Process Regression

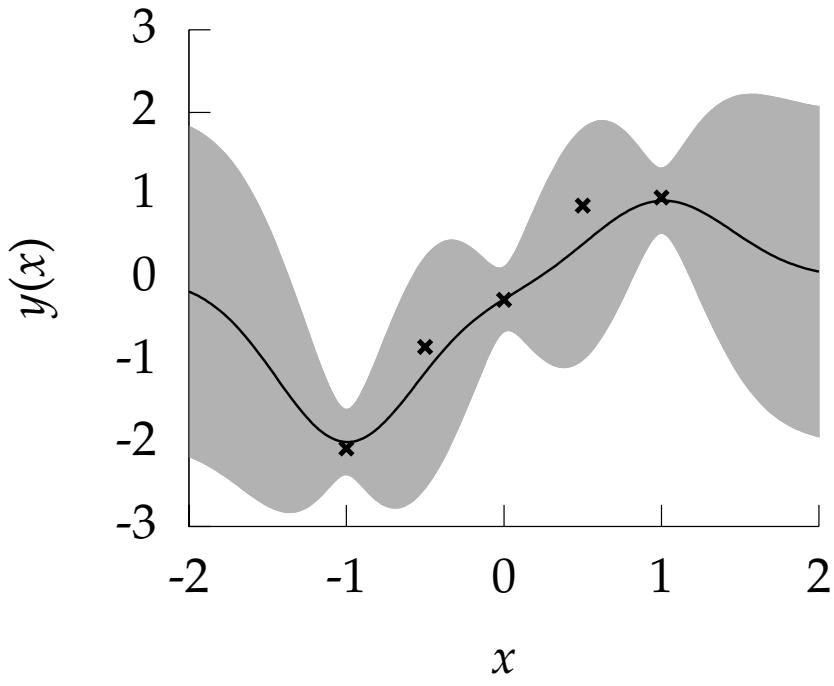


Figure: Examples include WiFi localization, C14 callibration curve.

http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Process Regression

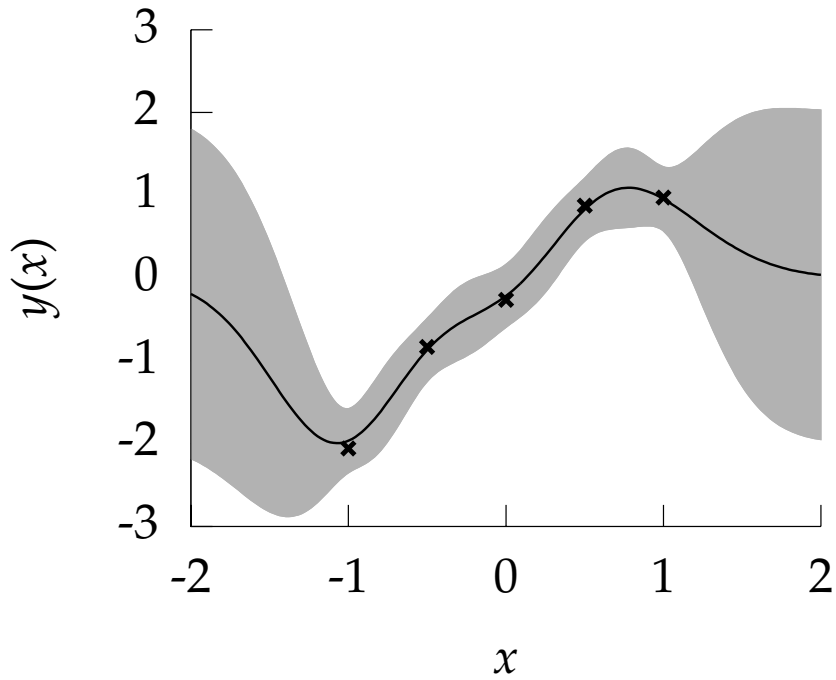


Figure: Examples include WiFi localization, C14 callibration curve.

http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Process Regression

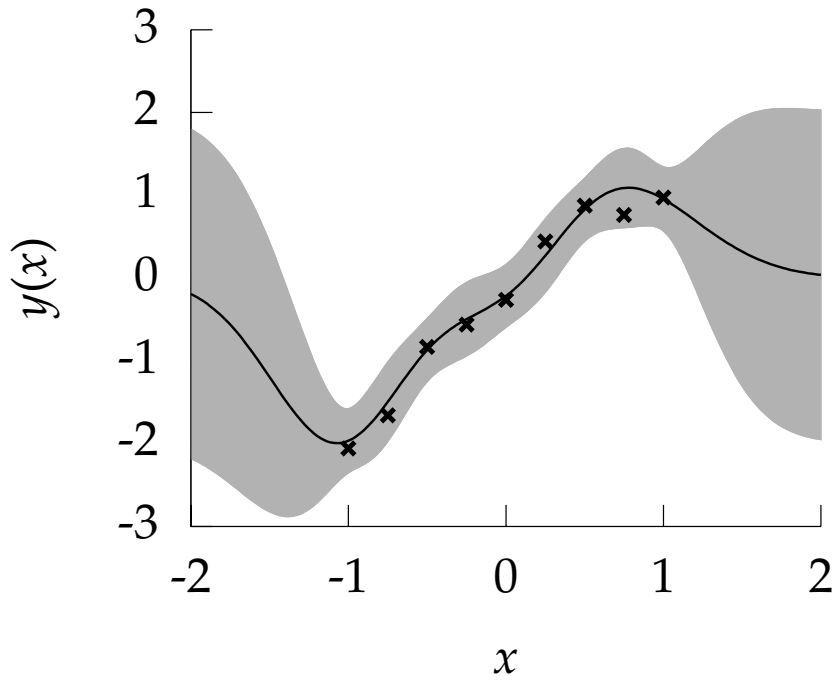


Figure: Examples include WiFi localization, C14 callibration curve.

http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Gaussian Process Regression

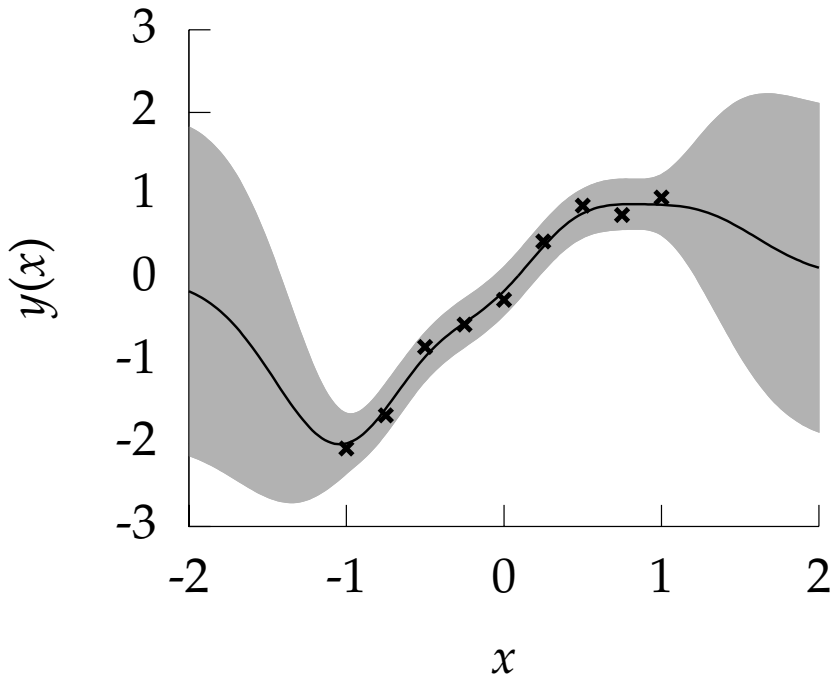
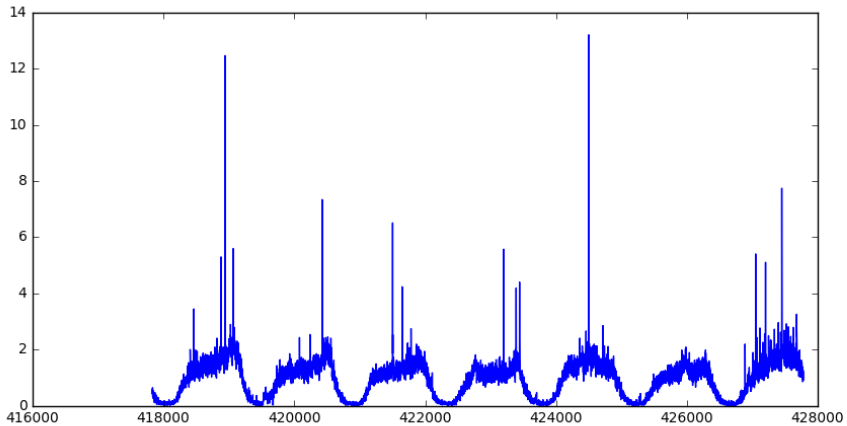


Figure: Examples include WiFi localization, C14 callibration curve.

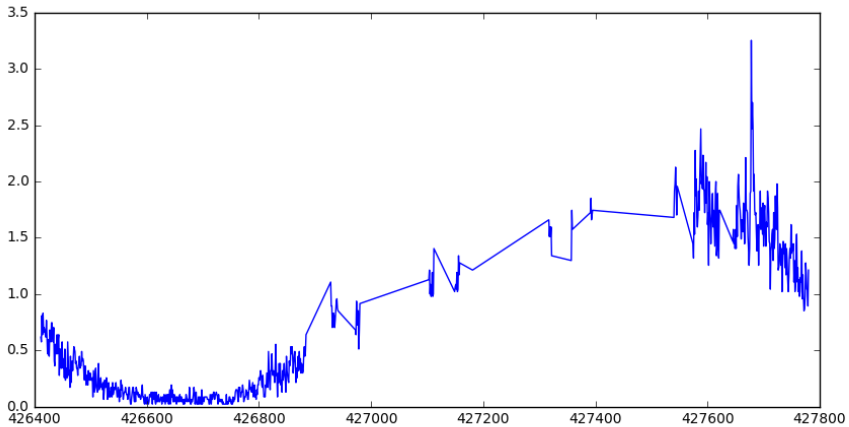
http://mlss.tuebingen.mpg.de/2015/slides/lawrence/lawrence.pdf

Model



- A periodic kernel to handle periodicity
- A Gaussian (RBF) kernel to handle the non-periodic part of the data
- A white noise kernel to handle fluctuations seen in the data

Model

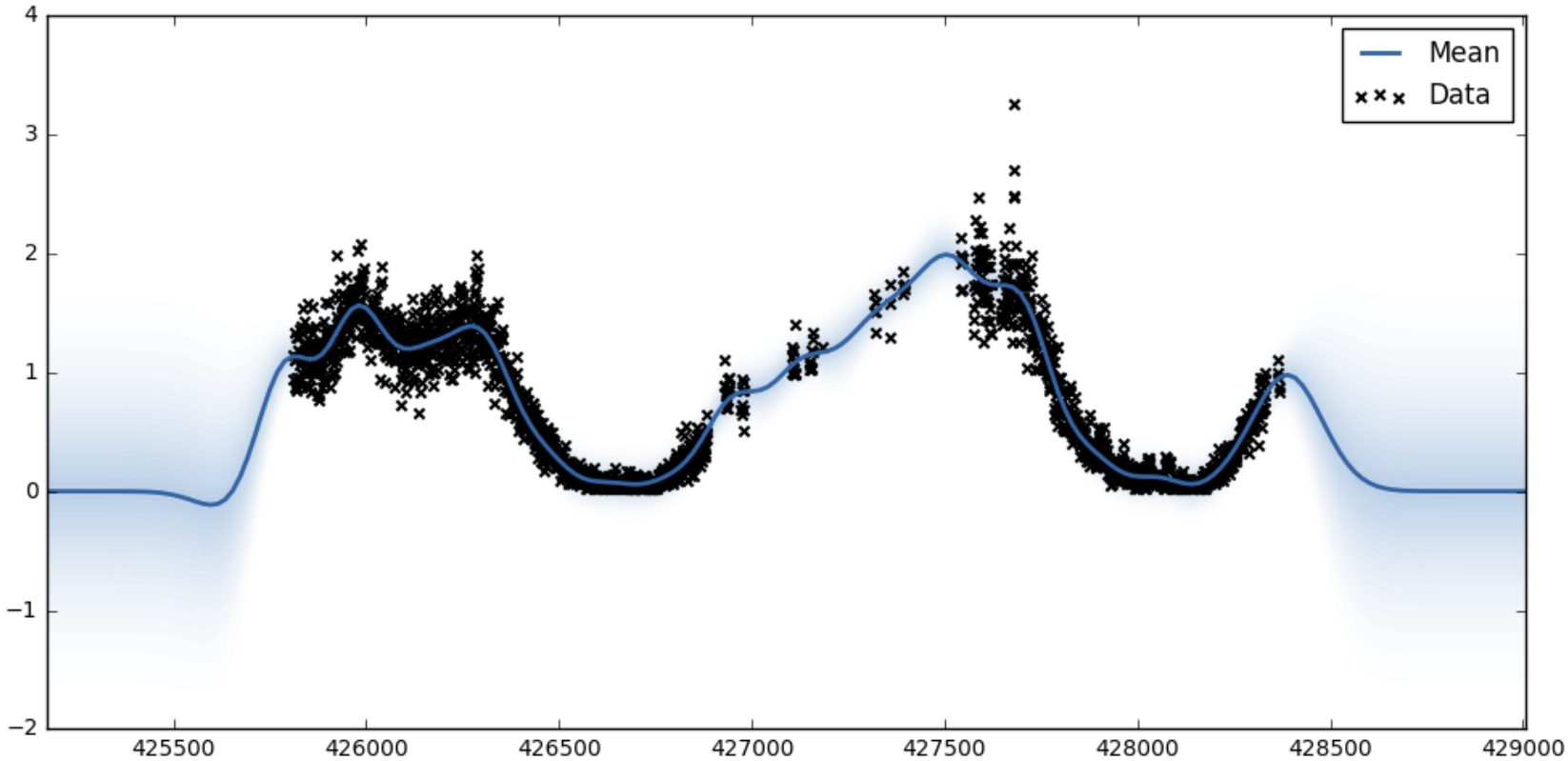


- A Gaussian (RBF) kernel to handle the non-periodic part of the data
- A white noise kernel to handle fluctuations seen in the data

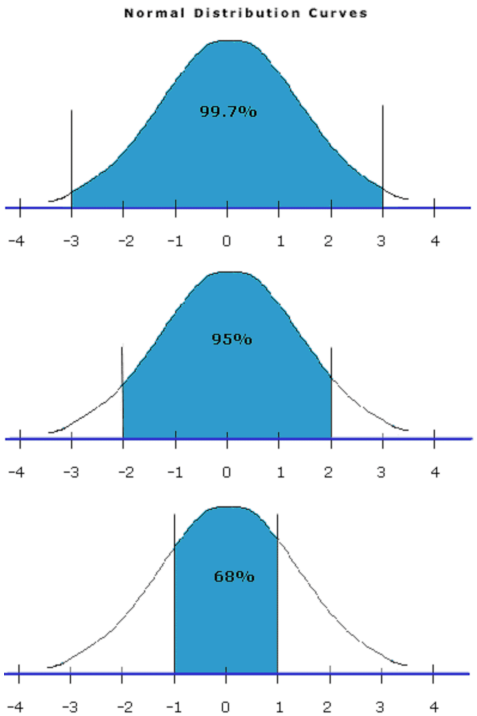
Fit the model, get expected mean and variance

```
import GPy

kernel = GPy.kern.RBF(input_dim=1) +
         GPy.kern.White(input_dim=1)

m = GPy.models.GPRegression(Xtr.reshape(-1,1),
                            ytr.reshape(-1,1),
                            kernel)
m.optimize()

mean, var = m.predict(Xte.reshape(-1,1),
                      full_cov=False,
                      include_likelihood=True)

fig = m.plot(plot_density=True)
```

Fitted Model



Significance




```
import math
import scipy

def phi(x):
    return 0.5 + 0.5 * scipy.special.erf(x / math.sqrt(2))

def score(x):
    return 1 - abs(phi(x) - phi(-x))

y_score = score((y - expected_mean) / expected_std)
```

Result - not so good ads

Observed	Expected Mean	Expected Variance	Score	Portion	Is Significant	TV-ad
1.11	0.98	0.03	0.55	0.11		
0.89	0.98	0.03	0.42	-0.1		TV
1.3	0.99	0.03	0.94	0.24	*	
1.13	0.99	0.03	0.59	0.12		
1.19	0.99	0.03	0.77	0.17		TV
1.28	1	0.03	0.91	0.22	*	
1.04	1	0.03	0.2	0.04		
1.53	1	0.03	1	0.34	*	
1.26	1.01	0.03	0.87	0.2		
1.11	1.01	0.03	0.44	0.09		
1.34	1.01	0.03	0.96	0.24	*	
1.26	1.02	0.03	0.86	0.19		
1.4	1.09	0.02	0.96	0.22	*	TV
2.57	1.09	0.02	1	0.58	*	
2.77	1.1	0.02	1	0.6	*	TV
1.51	1.1	0.02	0.99	0.27	*	
1.3	1.1	0.02	0.8	0.15		
1.34	1.1	0.02	0.87	0.18		
1.3	1.1	0.02	0.79	0.15		

Result - much better ads

Observed	Expected Mean	Expected Variance	Score	Portion	Is Significant	TV-ad
0.77	0.88	0.03	0.48	-0.14		
1.02	0.88	0.03	0.59	0.14		TV
1.62	0.88	0.03	1	0.46	*	TV
1.47	0.88	0.03	1	0.4	*	TV
1.26	0.89	0.03	0.97	0.29	*	
1.19	0.89	0.03	0.92	0.26	*	
1.28	0.89	0.03	0.97	0.3	*	
0.91	0.89	0.03	0.11	0.03		
1.13	0.89	0.03	0.82	0.21		
1.15	0.9	0.03	0.86	0.22		TV
4.45	0.9	0.03	1	0.8	*	TV
5.4	0.9	0.03	1	0.83	*	
3.21	0.9	0.03	1	0.72	*	
2.3	0.91	0.03	1	0.61	*	
1.96	0.91	0.03	1	0.54	*	
1.98	0.91	0.03	1	0.54	*	
1.3	0.91	0.03	0.97	0.3	*	
1.47	0.92	0.03	1	0.38	*	

Acknowledgments

- GPy: `https://github.com/SheffieldML/GPy`
- MLSS 2015: `http://mlss.tuebingen.mpg.de/2015/speakers.html`
- GPWS 2014: `http://ml.dcs.shef.ac.uk/gpss/gpws14/`
- This talk: `http://adrin.info/tv-ad-attribution-gaussian-processes.html`

Thank You!
Questions?