GP-Bandits

Thompson Sampling

Emile Mathieu

January 9, 2017

Ecole des Ponts ParisTech

Table of contents

Problem Statement

Problem Statement

Goal

Optimize a sequence sampled from an unknown reward function $f : D \to \mathbb{R}$.

At each round t:

1. Choose a point $\mathbf{x}_t \in D$
2. Observe the function value perturbed by noise: $y_t = f(\mathbf{x}_t) + \epsilon_t$

Aim to perform as well as $\mathbf{x}^\star = \arg\max_{\mathbf{x} \in D} f(\mathbf{x})$.

Minimize cumulative regret $R_T = \sum_t^T r_t$, with $r_t = f(\mathbf{x}^\star) - f(\mathbf{x}_t)$.

Gaussian Processes

Gaussian Processes

$$f \sim \mathcal{GP}(\mu, k)$$

with $\mu : D \to \mathbb{R}$ and $k : D \times D \to \mathbb{R}$

For any finite combination of dimensions $A = \{\mathbf{x}_1, \ldots, \mathbf{x}_T\}$, and by denoting $\mathbf{K} = [k(\mathbf{x}, \mathbf{x}')]_{\mathbf{x}, \mathbf{x}' \in A}$ and $\mathbf{m} = [\mu(\mathbf{x})]_{\mathbf{x} \in A}$:

$$\mathbf{f_A} \sim \mathcal{N}(\mathbf{m}, \mathbf{K})$$

Common choices of covariance functions:

- Finite dimensional linear: $k(\mathbf{x}, \mathbf{x}') = \mathbf{x}^T \mathbf{x}'$
- Squared Exponential kernel: $k(\mathbf{x}, \mathbf{x}') = \exp\{-(2l^2)^{-1} \|\mathbf{x} - \mathbf{x}'\|^2\}$
- Matern kernel: $k(\mathbf{x}, \mathbf{x}') = (2^{1-\nu}/\Gamma(\nu)) r^\nu B_\nu(r)$,
  $r = (\sqrt{2\nu}/l)\,\|\mathbf{x} - \mathbf{x}'\|$

GP as prior on f

Use a GP$(0_d, k(\cdot,\cdot))$ as a prior distribution over f.

Observe $y_t = f(\mathbf{x}_t) + \epsilon_t$, with i.i.d. $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$

The posterior over f is still a GP distribution which mean and variance are:

- $\mu_t = k_{t-1}(\mathbf{x})^T \left(K_{t-1} + \sigma^2 I_d\right)^{-1} \mathbf{y}_t$
- $k_t = k(\mathbf{x}, \mathbf{x}') - k_{t-1}(\mathbf{x})^T \left(K_{t-1} + \sigma^2 I_d\right)^{-1} k_{t-1}(\mathbf{x}')$
- $\sigma_t^2 = k_t(\mathbf{x}, \mathbf{x})$

with $k_t(\mathbf{x}) = [k(\mathbf{x}_1, \mathbf{x}), \ldots, k(\mathbf{x}_T, \mathbf{x})]^T$ and $\mathbf{K}_T = [k(\mathbf{x}, \mathbf{x}')]_{\mathbf{x}, \mathbf{x}' \in A_T}$.

Algorithms

GP-UCB

Algorithm 1 GP-UCB

Require: k

$\mu \leftarrow 0_d$
for $t \leftarrow 1$ to T **do**
  $\beta_t \leftarrow 2\log(|D|\, t^2\pi^2/6\delta)$
  Choose $\mathbf{x}_t \leftarrow \mathit{arg}\max_i \mu_{t-1} + \sqrt{\beta_t}\sigma_{t-1}$
  Observe $y_t = f(\mathbf{x}_t) + \epsilon_t$
  $\mu_t = k_{t-1}(\mathbf{x})^T \left(K_{t-1} + \sigma^2 I_d\right)^{-1} \mathbf{y}_t$
  $k_t = k(\mathbf{x}, \mathbf{x}') - k_{t-1}(\mathbf{x})^T \left(K_{t-1} + \sigma^2 I_d\right)^{-1} k_{t-1}(\mathbf{x}')$
  $\sigma_t^2 = k_t(\mathbf{x}, \mathbf{x})$
end for

For finite D: $\beta_t = 2\log(|D| t^2\pi^2/6\delta)$, with $\delta \in (0, 1)$

GP-UCB - Exemple

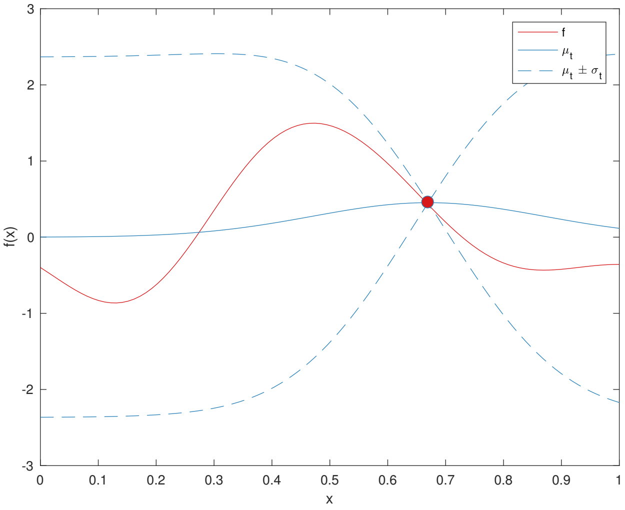


Figure 1: Posterior distribution of f at time step 1

GP-UCB - Exemple

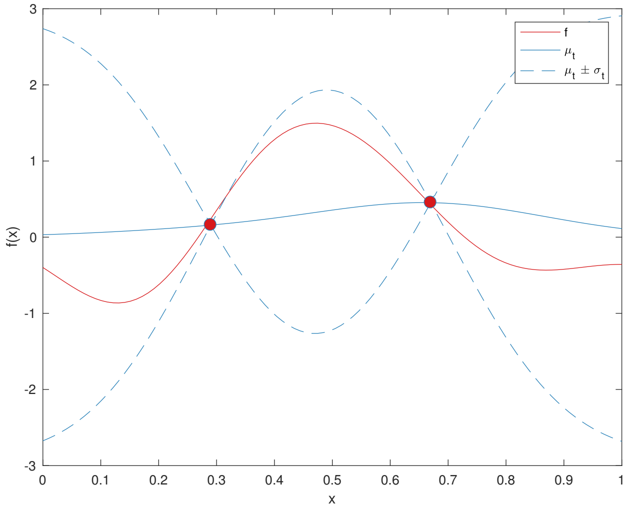


Figure 2: Posterior distribution of f at time step 2

GP-UCB - Exemple

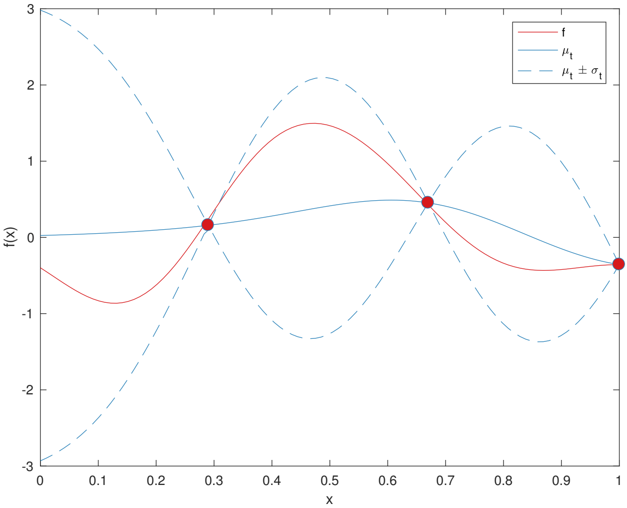


Figure 3: Posterior distribution of f at time step 3

GP-UCB - Exemple

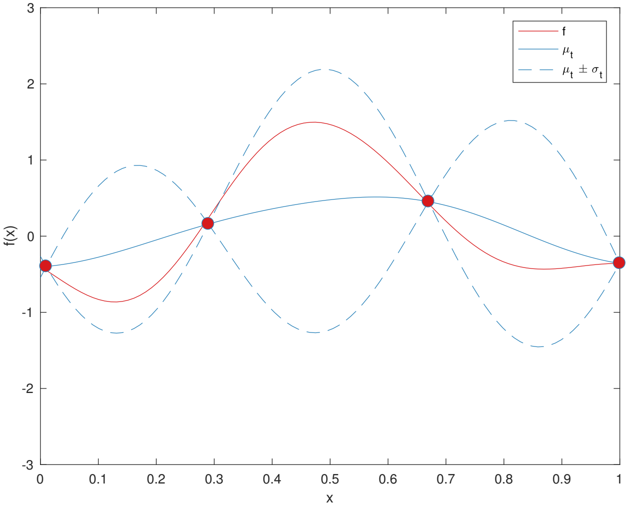


Figure 4: Posterior distribution of f at time step 4

GP-UCB - Exemple

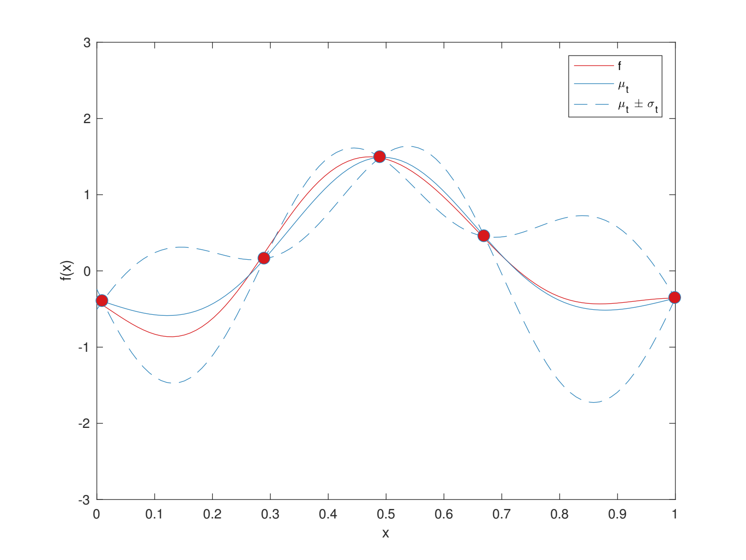


Figure 5: Posterior distribution of f at time step 5

GP-UCB - Exemple

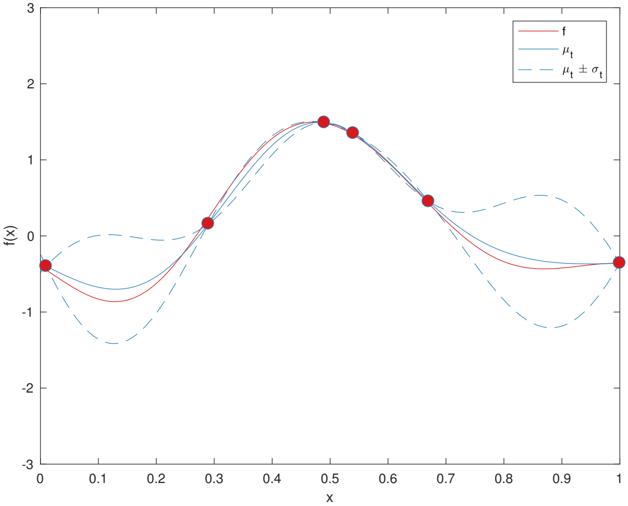


Figure 6: Posterior distribution of f at time step 6

GP-UCB - Exemple

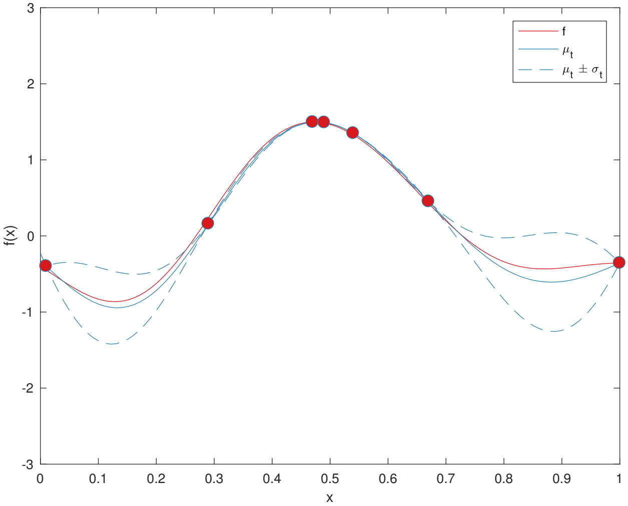


Figure 7: Posterior distribution of f at time step 7

GP-UCB - Exemple

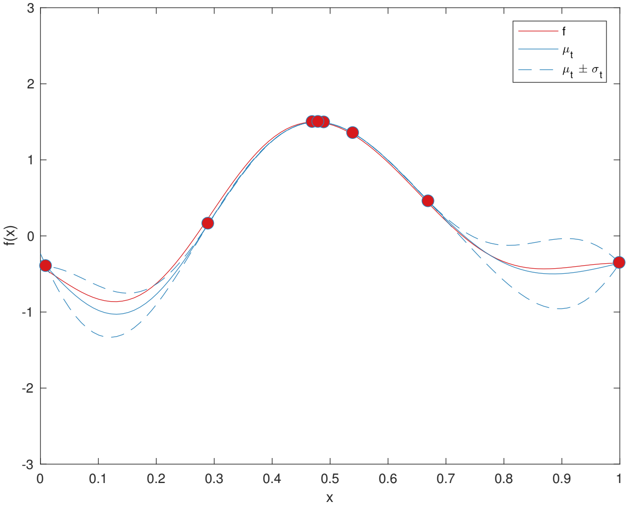


Figure 8: Posterior distribution of f at time step 8

TS-UCB

Algorithm 2 GP-TS

Require: k

1: $\mu \leftarrow 0_d$
2: **for** $t \leftarrow 1$ to T **do**
3: Sample $f_t \sim \mathbf{GP}(\mu_t, k_t)$
4: Choose $\mathbf{x}_t \leftarrow \mathit{arg}\max_{\mathbf{x}} f_t(\mathbf{x})$
5: Observe $y_t = f(\mathbf{x}_t) + \epsilon_t$
6: $\mu_t = k_{t-1}(\mathbf{x})^T (K_{t-1} + \sigma^2 I_d)^{-1} \mathbf{y}_t$
7: $k_t = k(\mathbf{x}, \mathbf{x}') - k_{t-1}(\mathbf{x})^T (K_{t-1} + \sigma^2 I_d)^{-1} k_{t-1}(\mathbf{x}')$
8: $\sigma_t^2 = k_t(\mathbf{x}, \mathbf{x})$
9: **end for**

GP-TS - Exemple

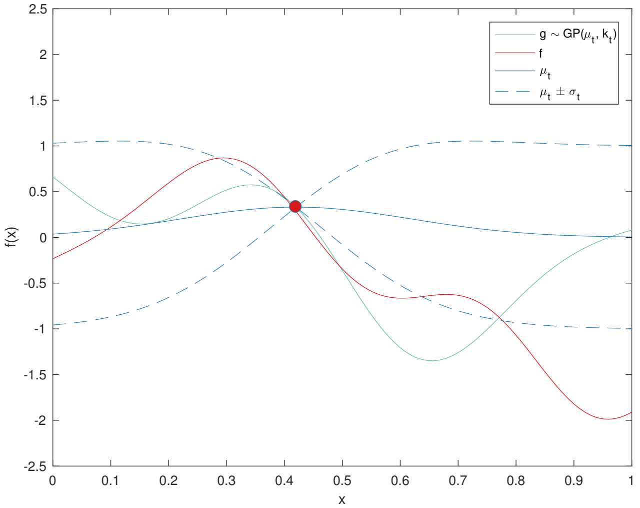


Figure 9: Posterior distribution of f at time step 2

GP-TS - Exemple

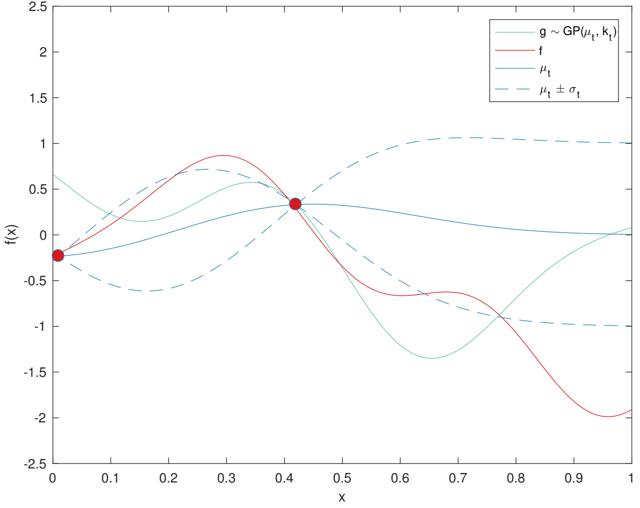


Figure 10: Posterior distribution of f at time step 2

GP-TS - Exemple

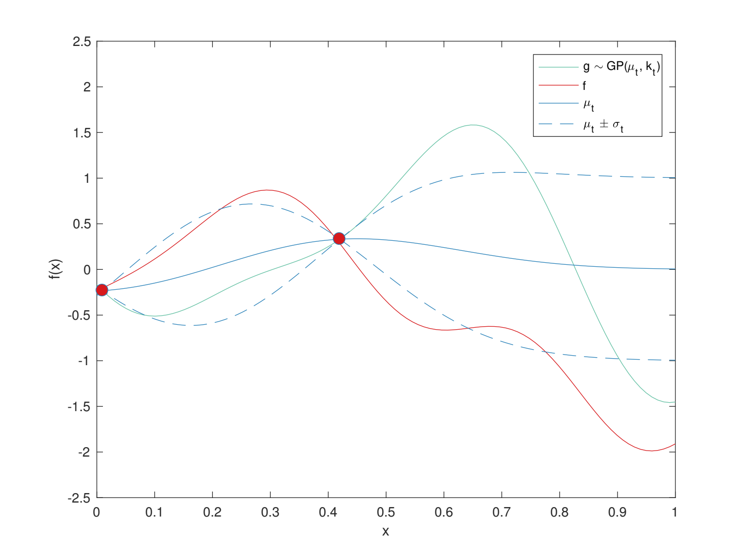


Figure 11: Posterior distribution of f at time step 3

GP-TS - Exemple

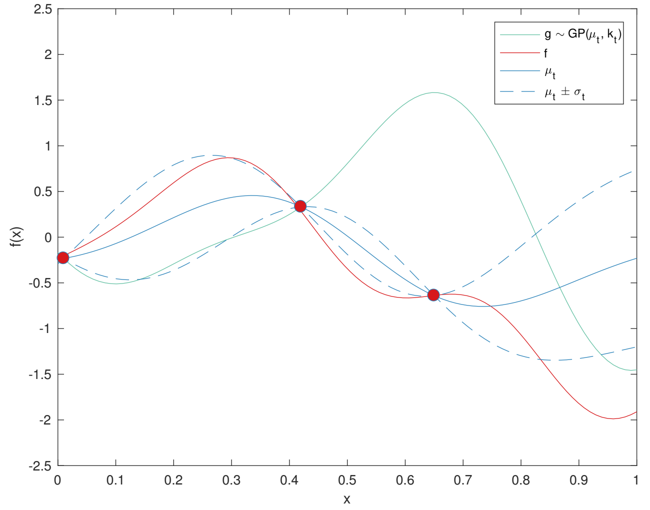


Figure 12: Posterior distribution of f at time step 3

GP-TS - Exemple

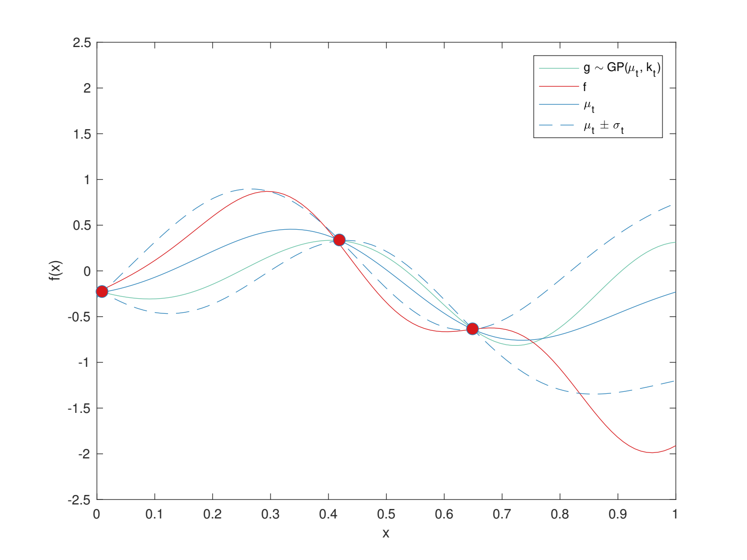


Figure 13: Posterior distribution of f at time step 4

GP-TS - Exemple

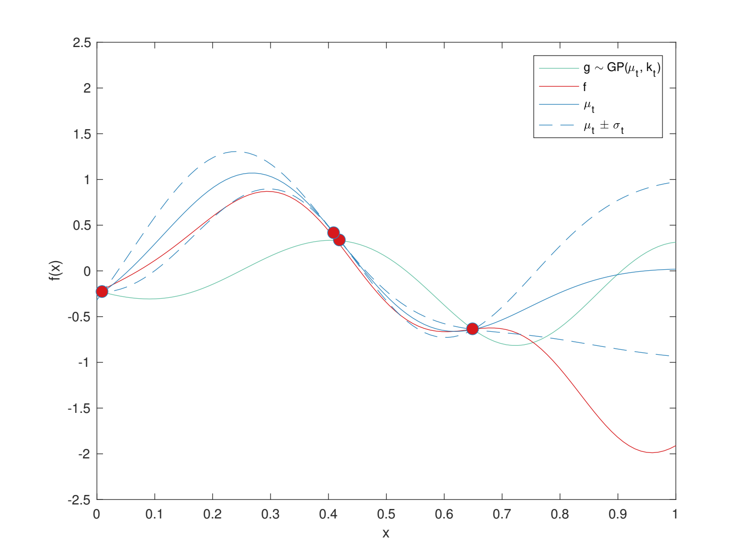


Figure 14: Posterior distribution of f at time step 4

GP-TS - Exemple

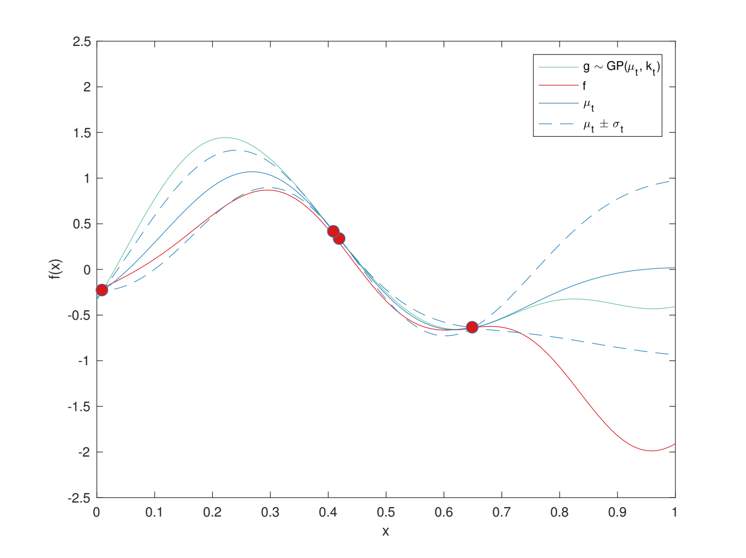


Figure 15: Posterior distribution of f at time step 5

GP-TS - Exemple

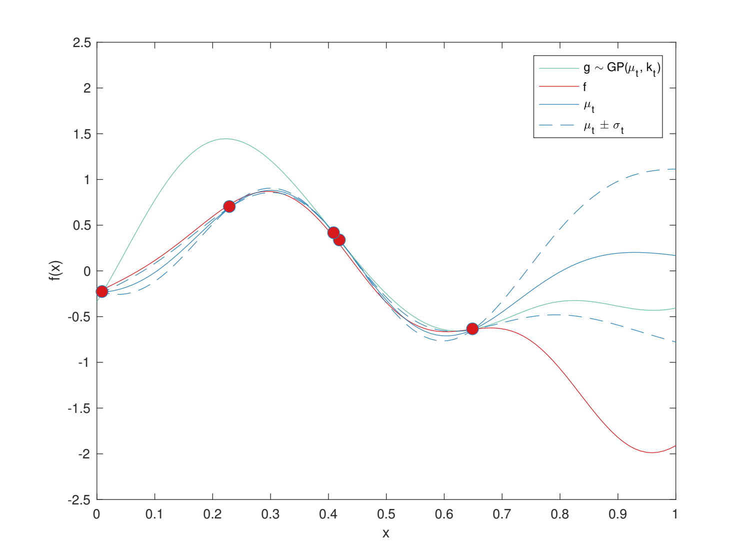


Figure 16: Posterior distribution of f at time step 5

GP-TS - Exemple

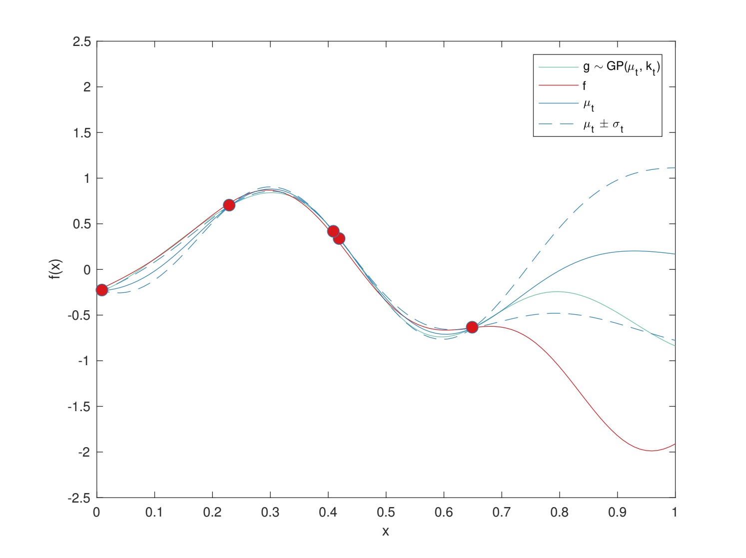


Figure 17: Posterior distribution of f at time step 6

GP-TS - Exemple

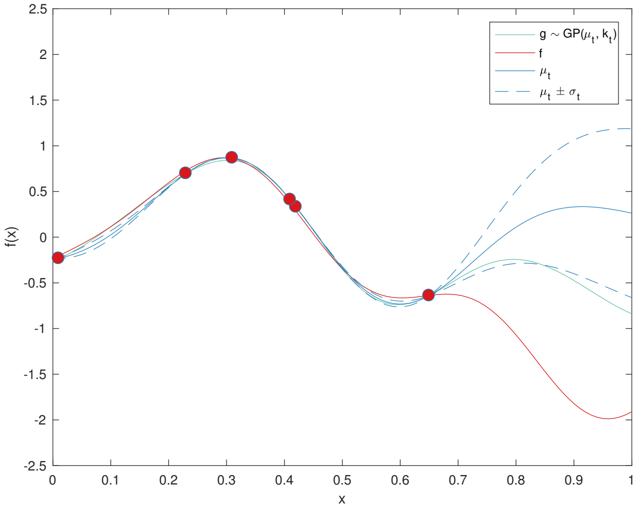


Figure 18: Posterior distribution of f at time step 6

GP-TS - Exemple

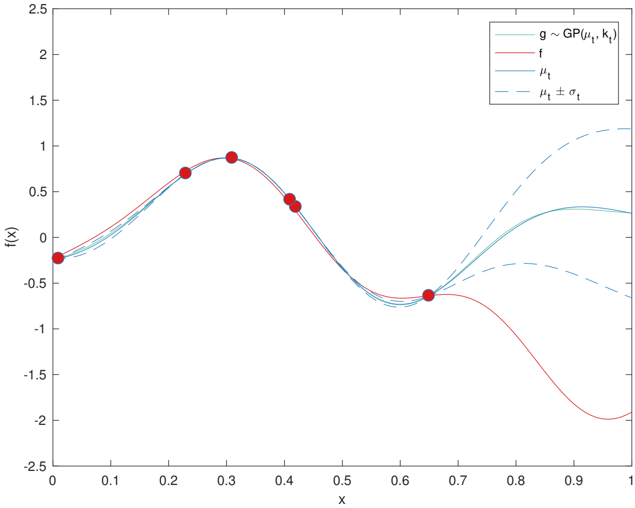


Figure 19: Posterior distribution of f at time step 7

Experiments

Experiments

Methods

GP-UCB, GP-TS and 2 naives methods

Datasets

Synthetic, Temperature, Trafic data

Sources

https://github.com/emilemathieu/Project_GP-bandits

Synthetic Data

- Sample random functions from a GP
- Squared exponential kernel with $l = 0.2$
- Decision set $D = [0, 1]$
- Uniformly discretized into 1000 points
- $\sigma^2 = 0.025$
- $T = 100$
- $\delta = 0.1$
- 150 runs

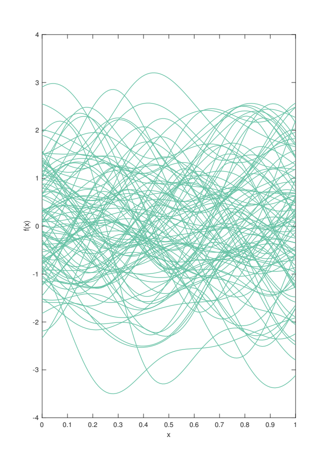


Figure 20: samples of zero mean GP

Synthetic Data - Results

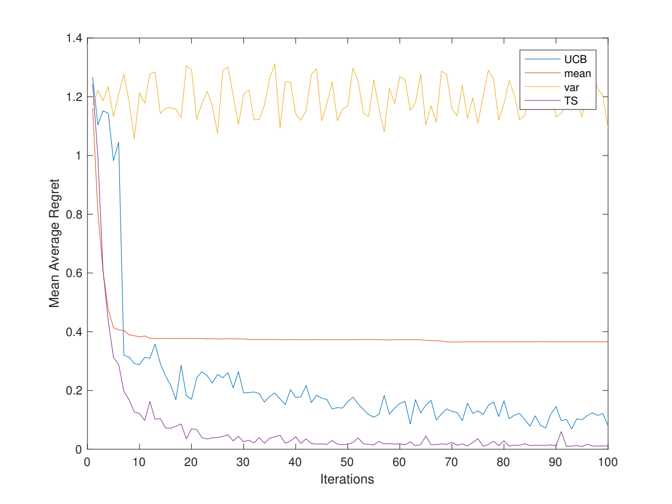


Figure 21: Performances on synthetic data

Temperature Data

- Sensors deployed at Intel Research Berkeley
- Preprocessing with python job
- 45 sensors
- One hour intervals
- A month

- $k = \frac{1}{N-1}\sum_{i=1}^{N} \mathbf{x_i}\mathbf{x_i}^T$
- $T = 45$
- $\sigma^2 = 5.0$
- $\delta = 0.1$
- 187 runs

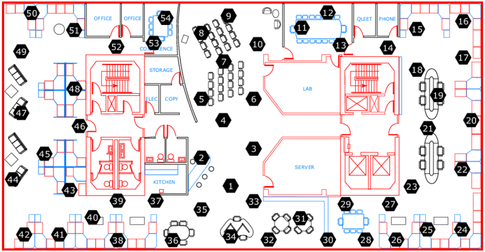


Figure 22: Intel lab sensors' map from
`http://db.csail.mit.edu/labdata/labdata.html`

Trafic Data

- Speed sensors in highway 5 South in San Diego
- Python job scheduled with cron
- 48 sensors
- One minute intervals
- 3 days from 6 AM to 11 AM (local time)
- $k = \frac{1}{N-1} \sum_{i=1}^{N} \mathbf{x_i} \mathbf{x_i}^T$
- $T = 48$
- $\sigma^2 = 4.78$
- $\delta = 0.1$
- 360 runs

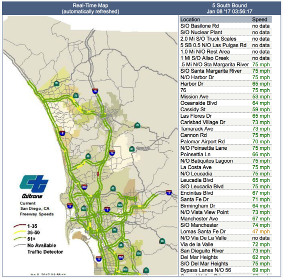
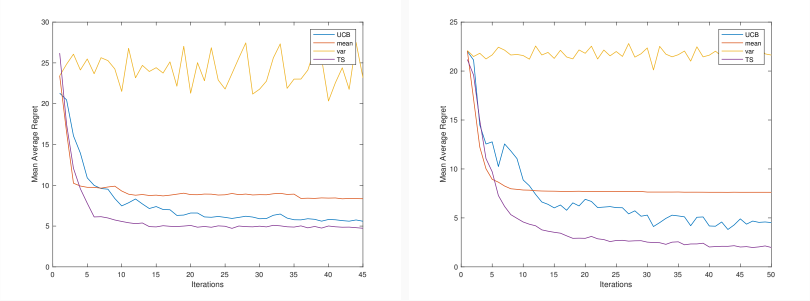


Figure 23: San Diego real-time traffic website from `http://www.dot.ca.gov/dist11/d11tmc/sdmap/showmap.php?route=sb5`

Real Data - Results



(a) Temperature data **(b)** Trafic data

Figure 24: Comparison of performances: GP-UCB, TS-UCB and 2 naive heuristics on temperature data (a) and trafic data (b).

Conclusion

Thank you for your attention !

Questions?

Kernel Learning

Squared exponential kernel:

$$k_l(\mathbf{x}, \mathbf{x}') = \exp\{-(2l^2)^{-1} \|\mathbf{x} - \mathbf{x}'\|^2\}$$

Likelihood of observations $\mathbf{y} = [y_1, \ldots, y_T]^T$:

$$\mathcal{N}(\mathbf{y}|\mathbf{0}, \mathbf{K}_l) = (2\pi)^{-n/2} |\mathbf{K}_l + \sigma^2 I|^{-1/2} \exp\{-\mathbf{y}^T (\mathbf{K}_l + \sigma^2 I)^{-1} \mathbf{y}\}$$

with $\mathbf{K_A} = [k_l(\mathbf{x}, \mathbf{x}')]_{\mathbf{x}, \mathbf{x}' \in \mathbf{A}}$ and $\mathbf{A} = \{\mathbf{x}_1, \ldots, \mathbf{x}_T\}$.

Convex optimization problem:

$$\underset{l}{\text{minimize}} \quad \frac{1}{2} \log |\mathbf{K}_l + \sigma^2 I| + \frac{\mathbf{y}^T (\mathbf{K}_l + \sigma^2 I)^{-1} \mathbf{y}}{2} \propto -\log \mathcal{N}(\mathbf{y}|0, \mathbf{K}_l)$$

Synthetic Data - Kernel Learning

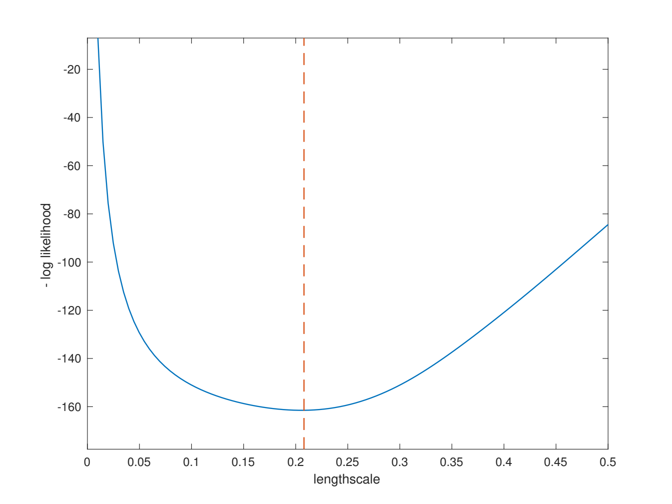


Figure 25: Negative log likelihood versus kernel lengthscale