AML, 17 Sep 2019

Gaussian Process Regression

Scikit-learn implement Rasmussen's (original) algo

Target y_i, Input $\underset{\sim}{x}_i = (x_1, \ldots, x_p)$ $i = 1, \ldots, n$

$y = f(\underset{\sim}{x})$ — find y

$$f(\underset{\sim}{x}) \sim N_n\left(\mu(\underset{\sim}{x}), \Sigma(\underset{\sim}{x}, \underset{\sim}{x}')\right)$$

Standard linear regression

$$\mu(\underset{\sim}{x}) = b + w\underset{\sim}{x}$$

$$\Sigma(\underset{\sim}{x}, x') = \sigma^2 I_n \quad \text{i.i.d.}$$

Instead (Rasmussen)

$$\mu(\underset{\sim}{x}) = 0$$

some distance function

$$\Sigma(\underset{\sim}{x}, \underset{\sim}{x}') = \sigma^2 \exp(-s \cdot |\underset{\sim}{x} - x'|)$$

Computation similar to standard regression

Scikit-learn

Take care of singularities

$$y = f(\underset{\sim}{x}) + \varepsilon$$

$$f(\underset{\sim}{x}) \sim N_n\left(\mu(\underset{\sim}{x}), \Sigma(\underset{\sim}{x}, \underset{\sim}{x}')\right)$$

$$\varepsilon \sim N_n(0, \sigma^2 I_n)$$

Instead of f, use y with noise

Alternative package

GPy → GPflow on Tensorflow

$$D = \left\{ \begin{array}{c|cc} ID & & \\ \hline 1 & y_1 & x_1 \\ 2 & y_2 & x_2 \\ \vdots & \vdots & \vdots \\ n & y_n & x_n \end{array} \right.$$

n very large

Choose $m << n$

1. Draw n id^* from ID by SRSWR
2. $D_m^* = D[id^*]$
3. Fit GP $\hat{f}_i$ with D^*
4. Estimate y^* for from $\hat{f}_i$

Iteration i

Each iteration is independent - parallelize!

M iterations $\longrightarrow \hat{f}_1, \dots, \hat{f}_M$

Final $\hat{f}$ is mean $\longrightarrow \frac{1}{M} \sum_{i=1}^{M} \hat{f}_i$