Lecture 22: Gaussian process regression

Professor Ilias Bilionis

Multivariate regression and automatic relevance determination

PREDICTIVE SCIENCE LABORATORY

Automatic relevance determination

$\underline{d > 1}$ $\quad \underline{x} = (x_1, \dots, x_d)$

$$\underline{x}_{1:n} = (\underline{x}_1, \dots, \underline{x}_n)$$

observations are vectors

$$\underline{x}_1 = (x_{11}, \dots, x_{1d})$$

SE Naive:

$$c(\underline{x}, \underline{x}') = s^2 \exp\left\{-\frac{\|\underline{x}-\underline{x}'\|^2}{2\ell^2}\right\} = s^2 \exp\left\{-\sum_{i=1}^{d} \frac{(x_i - x_i')^2}{2\ell^2}\right\}$$

same lengthscale for all inputs (big assumption)

Smart ARD

$$c(x, x') = s^2 \exp\left\{-\sum_{i=1}^{d} \frac{(x_i - x_i')^2}{2\ell_i^2}\right\}$$

different lengthscale for each input/dimension

$$\theta = (s, \ell_1, \ell_2, \dots, \ell_d) \quad (\underline{d+1} \text{ parameters})$$

MAP: $\max_{\sigma, s, \ell_1, \dots, \ell_d} \log p(\theta, \sigma \mid x_{1:n}, y_{1:n}) \Rightarrow$ Some of the ℓ_i may become very big. corresponding x_i components can be removed from the analysis