Gaussian Processes

A) Plan: apply Lemma 1 $\rightarrow p\left(\begin{pmatrix} y \\ f \end{pmatrix}\right) \rightarrow$ appy conditional normal dist. from Exercise 1

$$y \mid f \sim N(If, \sigma^2 I)$$

$$f = \begin{bmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{bmatrix} \qquad y \mid \theta \sim N(R\theta, \Sigma), \theta \sim N(m, V)$$

$$f \sim N(0, C)$$

$R = I,\ \theta = f,\ \Sigma = \sigma^2 I,\ m = D,\ V = C$ Solution: Lemma 1: $\begin{pmatrix} y \\ \theta \end{pmatrix}$ $N\left(\begin{bmatrix} Rm \\ m \end{bmatrix} \begin{bmatrix} RVR^T + \Sigma & RV \\ VR^T & V \end{bmatrix}\right)$

$$\rightarrow p(f \mid y) = N(C(C + \sigma^2 I)^{-1} y, C - ((C + \sigma^2 I)^{-1} C)$$

Nonparametric Regression

B) $\begin{pmatrix} y \\ f^* \end{pmatrix} \sim N\left(0, \begin{bmatrix} C + I\sigma^2 & (C_{*1} \ldots C) \\ (C_{*1} \ldots C & C_{**} \end{bmatrix}\right)$

$$f^* \mid f, y \sim N(0 + (C_{*1} \ldots C_{*N}) C_{N \times N}^{-1} \begin{pmatrix} f(x_1) \\ f(x_n) \end{pmatrix} \ldots \ldots$$

Solution:

$$E(f^* \mid y_1, \ldots, y_N) = (C_{*1}, C_{*2}, \ldots, C_{*n}) C_{N \times N}^{-1} \cdot \left(\frac{I}{\sigma^2} + C_{N \times N}^{-1}\right) \frac{y}{\sigma^2}$$

$$I = C_{N \times N}^{-1} \cdot C_{N \times N}$$

D) from A) $\begin{pmatrix} y \\ f \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} C + I\sigma^2 & C \\ C & C \end{bmatrix}\right)$

$$y \sim N(0, C(\tau_1^2, \tau_2^2, b) + \sigma^2 I)$$

marginal max. likelihood / empirical bayes / evidence $\rightarrow$ way to choose hyperparams (but difficult)

$$\log p(y) = -\frac{1}{2} y^T [\underbrace{C(\tau_1^2, \tau_2^2, b)}_{} + \sigma^2 I]^{-1} y - \frac{1}{2} \log(\det |C(\tau_1^2, \tau_2^2, b) + \sigma^2 I|)$$

$$\int p(y, f \mid \text{stuff})\, df$$

$\rightarrow$ dense covariance matrix $\rightarrow$ difficult to invert

How can we make the cov. matrix sparse?

$$C_{ij}^{(x)} = \tau_1^2 \exp\left\{-\frac{1}{2}\left(\frac{x_i - x_j}{b}\right)^2\right\}$$

completely supported