1(a) Write down the log posterior $\log p(f_x | y, X)$ for the representers.
$f_x = [f(x_1), f(x_2), \ldots, f(x_N)]$

① prior over latent f : $p(f) = GP(f; m, k)$

② $y_i \in \{-1, 1\}$ input output pairs := $\{(x_i, y_i)\}_{i=1}^{N}$

③ $p(y | f_x) = \prod_{i=1}^{N} \frac{1}{1 + \exp(-y_i f(x_i))}$

$\sigma(y_i f(x_i)) = \frac{1}{1 + \exp(-y_i f(x_i))}$

$$\log p(f_x | y) = \log p(y | f_x) + \log p(f_x) - \log(p(y))$$

$$= \sum_{i=1}^{n} \log \sigma(y_i f(x_i)) - \frac{1}{2} (f_x - m_x)^T K_{xx}^{-1} (f_x - m_x) + \text{constant term.}$$

$$\cong -\sum_{i=1}^{n} \log(1 + \exp(-y_i f(x_i))) - \frac{1}{2} (f_x - m_x)^T K_{xx}^{-1} (f_x - m_x)$$

(b) Compute the gradient of this log posterior

$$\nabla_{f_x} \log(p(f_x | y, X))$$

Answer: ① prior over latent f : $p(f) = GP(f; m, \kappa)$

② $y_i \in \{-1, 1\}$ input output pairs $:= \{(x_i, y_i)\}_{i=1}^{N}$

③ $p(y | f_x) = \prod_{i=1}^{N} \frac{1}{1 + \exp(-y_i f(x_i))}$

$$\log p(f_x | y) = \log p(y | f_x) + \log p(f_x) - \log(p(y))$$

$$= \sum_{i=1}^{n} \log \sigma(y_i f(x_i)) - \frac{1}{2} (f_x - m_x)^T K_{xx}^{-1} (f_x - m_x) + \text{constant term.}$$

$$\nabla \log(p(f_x | y)) = \sum_{i=1}^{n} \nabla_{f_x} \log(\sigma(y_i f_{x_i})) - K_{xx}^{-1} (f_x - m_x)$$

$$= \sum_{i=1}^{n} - \nabla_{f_x} \underbrace{\log(1 + \exp(-y_i f_{x_i}))} - K_{xx}^{-1} (f_x - m_x)$$

$$-\frac{\partial}{\partial f_{x_j}} \log(1 + \exp(-y_i f_{x_i}))$$

$$= \delta_{ij} \left(\frac{y_i + 1}{2} - \sigma(f_{x_i}) \right)$$

here as $y_i \in \{-1, 1\}$ $\quad \sigma(f_{x_i}) = \frac{1}{1 + \exp(-f_{x_i})}$

$\frac{y_i + 1}{2} \in \{0, 1\}$

Consider $-\dfrac{\partial}{\partial f_{x_i}} \log(1+\exp(-y_i f_{x_i})) = \dfrac{-1}{1+\exp(-y_i f_{x_i})} \, -y_i \exp(-y_i f_{x_i})$

$$= \frac{y_i \exp(-y_i f_{x_i})}{1+\exp(-y_i f_{x_i})}$$

① $y_i = -1 \Rightarrow \dfrac{-\exp(f_{x_i})}{1+\exp(f_{x_i})} = \dfrac{-1}{1+\exp(-f_{x_i})} = -\sigma(f_{x_i})$

② $y_i = 1 \Rightarrow \dfrac{\exp(-f_{x_i})}{1+\exp(-f_{x_i})} = 1 - \dfrac{1}{1+\exp(-f_{x_i})} = 1 - \sigma(f_{x_i})$

$$\Rightarrow \frac{y_i+1}{2} - \sigma(f_{x_i})$$

$$\nabla \log(p(f_x \mid y)) = \begin{bmatrix} \frac{y_1+1}{2} - \sigma(f(x_1)) \\ \frac{y_2+1}{2} - \sigma(f(x_2)) \\ \vdots \\ \frac{y_n+1}{2} - \sigma(f(x_n)) \end{bmatrix} - K_{xx}^{-1}(f_x - m_x)$$