- Let $D_t = \{(x_1, f(x_1)), \ldots, (x_n, f(x_n))\}$ be a set of collected training data
- Let the covariance matrix $Q(D_t, \vartheta)$ be given by

$$Q(\vartheta) = \begin{bmatrix} k(x_1, x_1, \vartheta) & k(x_1, x_2, \vartheta) & k(x_1, x_n, \vartheta) \\ \vdots & \ddots^{\sigma} & \vdots \\ k(x_n, x_1, \vartheta) & \cdots & k(x_n, x_n, \vartheta) \end{bmatrix} + \sigma^2 \, Id_n$$

where Id_n is the identity matrix, and σ is eventual noise of observation contained in $f(x)$

- Find the ϑ which maximizes the likelihood (optimization)

$$\log[p(D_t | \vartheta)] = -\frac{1}{2}\left(Y - \overline{F}\right)^t Q^{-1}(\vartheta)\left(Y - \overline{F}\right) - \frac{1}{2}\log|Q(\vartheta)| - \frac{n}{2}\log(2\pi)$$