Sparse Gaussian

January 2, 2017

1 mathematical notations

We first give brief description of mathematical notations will be used through out the project.

The original data set will be denoted as $\mathcal{D}$ which consists of N d-dimensional vectors $\mathbf{X} = \{\mathbf{x}^{(i)} = (x_1, \ldots, x_d) \,|\, i = 1, \ldots, N\}$. Let the new input data be $\mathbf{x}^* = (x_1^*, \ldots, x_d^*)$. The pseudo input data set is denoted as $\bar{\mathcal{D}}$ consists of $\bar{\mathbf{X}} = \{\bar{\mathbf{x}}^{(i)} = (x_1, \ldots, x_d) \,|\, i = 1, \ldots, M\}$. $\mathbf{X}$ is paired with target $\mathbf{Y} = (y^{(1)}, \ldots, y^{(N)})$, notice that $y^{(i)}$ are scalars. $\mathbf{x}^*$ is paired with new target y^*. The underlining latent function is denoted as $\mathbf{f}(\mathbf{x}) = \mathbf{y}$ and the pseudo one is $\bar{\mathbf{f}}$. A Gaussian distribution is denoted as $\mathcal{N}(\mathbf{f}|\mathbf{m}, \mathbf{V})$ with mean $\mathbf{m}$ and variance $\mathbf{V}$.

2 sparse Gaussian process

We first give a zero mean Gaussian prior over the underlining latent function: $p(\mathbf{f}|\mathbf{X}) = \mathcal{N}(\mathbf{f}|\mathbf{0}, \mathbf{K}_N)$ where $\mathbf{K}_N$ is our kernel matrix with elements given by, $[\mathbf{K}_N]_{nn'} = K(\mathbf{x}, \mathbf{x}')$:

$$K(\mathbf{x}, \mathbf{x}') = c \exp[-\frac{1}{2} \sum_{i=1}^{D} b_i (x_i^{(n)} - x_i^{(n')})^2], \quad \theta \equiv \{c, \mathbf{b}\}. \tag{1}$$