CO-496: Mathematics for Inference and Machine Learning

Problem Sheet for Tutorial 6

Problem 1

Whitened SVMs. Assume the centralised data $\mathbf{x}_i$, $i = 1, \ldots, n$. Assume further that each datum comes with a label $y_i = -1, 1$. Let $\mathbf{S}_t$ is the covariance matrix of the data and D a positive constant between 0 and 1. Then, formulate the dual of the following whitened SVM optimisation problem

$$\begin{aligned} \min_{\mathbf{w},b,\xi_i} & \frac{1-D}{2}\mathbf{w}^T\mathbf{w} + \frac{D}{2}\mathbf{w}^T\mathbf{S}_t\mathbf{w} + C\sum_{i=1}^{n}\xi_i \\ \text{s.t.} & y_i\left(\mathbf{w}^T\mathbf{x}_i + b\right) \geq 1 - \xi_i, \xi_i \geq 0. \end{aligned} \tag{1}$$

Problem 2

Kernel Discriminant Analysis.

Assume the data samples $\mathbf{x}_1, \ldots, \mathbf{x}_n$, split in C different classes. Assume you are given a positive definite kernel k which defines an implicit Hilbert space on the vectors $\phi(\mathbf{x}_i) \in \mathcal{H}$. Assume the Kernel Discriminant Analysis (KDA) optimisation problem

$$\begin{aligned} \max_{\mathbf{W}_\Phi} & \ \mathrm{tr}(\mathbf{W}_\Phi^T\mathbf{S}_b^\Phi\mathbf{W}_\Phi) \\ \text{subject to} & \ \mathbf{W}_\Phi^T\mathbf{S}_w^\Phi\mathbf{W}_\Phi = \mathbf{I} \end{aligned} \tag{2}$$

where $\mathbf{S}_b^\Phi$ and $\mathbf{S}_w^\Phi$ are the between and within class scatter matrices, respectively, defined in the Hilbert space of the vectors $\phi(\mathbf{x}_i) \in \mathcal{H}$. Find the optimal $\mathbf{W}_\Phi$ and extract the features from a test vector $\phi(\mathbf{y})$.

Problem 3

SVMs are very powerful methods for classification. Nevertheless, they are sensitive to affine transformations of the data and to directions with large data spread. Maximum margin solutions may be misled by the spread of data and preferentially separate classes along large spread directions. An alternative is the Relative Margin Machines (RMM) which creates a relative margin, controlled by a parameter $B > 1$. Assume the centralised data $\mathbf{x}_i$, $i = 1, \ldots, n$. Assume further that each datum comes with a label $y_i = -1, 1$. Then, the RMM optimisation problem

$$
\begin{aligned}
&\min_{\mathbf{w},b,\xi_i} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{n}\xi_i \qquad (3)\\
&\text{s.t. } y_i\left(\mathbf{w}^T\mathbf{x}_i + b\right) \geq 1 - \xi_i, \xi_i \geq 0\\
&\quad \mathbf{w}^T\mathbf{x}_i + b \leq B\\
&\quad -\mathbf{w}^T\mathbf{x}_i - b \leq B.
\end{aligned}
$$

Formulate the dual of the above optimisation problem.