

Exercise Sheet 8

Exercise 1: Dual formulation of the Soft-Margin SVM (5 + 20 + 10 + 5 P)

The primal program for the linear soft-margin SVM is

$$\min_{\mathbf{w}, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i$$

subject to

$$\forall_{i=1}^N : y_i \cdot (\mathbf{w}^\top \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \quad \text{and} \quad \xi_i \geq 0$$

where $\|\cdot\|$ denotes the Euclidean norm, ϕ is a feature map, $\mathbf{w} \in \mathbb{R}^d, b \in \mathbb{R}$ are the parameter to optimize, and $\mathbf{x}_i \in \mathbb{R}^d, y_i \in \{-1, 1\}$ are the labeled data points regarded as fixed constants. Once the hard-margin SVM has been learned, prediction for any data point $\mathbf{x} \in \mathbb{R}^d$ is given by the function

$$f(\mathbf{x}) = \text{sign}(\mathbf{w}^\top \phi(\mathbf{x}) + b).$$

- (a) *State* the conditions on the data under which a solution to this program can be found from the Lagrange dual formulation (*Hint: verify the Slater's conditions*).
- (b) *Derive* the Lagrange dual and show that it reduces to a constrained quadratic optimization problem. State both the objective function and the constraints of this optimization problem.
- (c) *Describe* how the solution (\mathbf{w}, b) of the primal program can be obtained from a solution of the dual program.
- (d) *Write* a kernelized version of the dual program and of the learned decision function.

Exercise 2: SVMs and Quadratic Programming (10 P)

We consider the CVXOPT Python software for convex optimization. The method `cvxopt.solvers.qp` solves quadratic optimization problems given in the matrix form:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \frac{1}{2} \mathbf{x}^\top P \mathbf{x} + \mathbf{q}^\top \mathbf{x} \\ \text{subject to} \quad & G \mathbf{x} \preceq \mathbf{h} \\ \text{and} \quad & A \mathbf{x} = \mathbf{b}. \end{aligned}$$

Here, \preceq denotes the element-wise inequality: $(\mathbf{h} \preceq \mathbf{h}') \Leftrightarrow (\forall_i : h_i \leq h'_i)$. Note that the meaning of the variables \mathbf{x} and \mathbf{b} is different from that of the same variables in the previous exercise.

- (a) *Express* the matrices and vectors $P, \mathbf{q}, G, \mathbf{h}, A, \mathbf{b}$ in terms of the variables of Exercise 1, such that this quadratic minimization problem corresponds to the kernel dual SVM derived above.

Exercise 3: Programming (50 P)

Download the programming files on ISIS and follow the instructions.