Practical Session 6

Constrained Optimization and SVM

1 Duality of finding maxima of a set

Problem 1: Given a set of variables $x_1, ..., x_N \in \mathbb{R}$, define an equation that finds the largest value in the set via minimization. Then, use the Lagrange dual function to derive a second, equivalent maximization problem.

Problem 2: Given a set of variables $x_1, ..., x_N \in \mathbb{R}$, define an equation that calculates the sum of the k largest values via maximization. Then, use the Lagrange dual function to derive a second, equivalent minimization problem.

2 Constrained Optimization Toy Problem

Suppose we have 40 pieces of raw material. Toy A can be made of one piece material with 3 EUR machining fee. A larger toy B can be made from two pieces of material with 5 EUR machining fee.

Because distribution costs decrease with larger quantities, we can sell x pieces of toy A for $20 - x$ EUR each, and y pieces of toy B for $40 - y$ EUR each. From our experience, toy B is more popular than toy A; therefore, we will produce not more of toy A than of toy B.

To get the maximum profit, we want to calculate the amount of toy A and toy B that we should produce.

Problem 3: Write down the constrained optimization problem and the associated Lagrangian.

Problem 4: Write down the Karush–Kuhn–Tucker (KKT) conditions for the above optimization problem.

Problem 5: Obtain the solution to the constrained optimization problem by solving the KKT conditions. Do not worry about non-integer production quantities.

3 Concrete SVM Example

You are given a data set with data from a single feature x in $\mathbb{R}$ and corresponding labels $y \in \{+1, -1\}$. Data points for $+1$ are at $-3, -2, 3$ and data points for -1 are at $-1, 0, 1$.

Problem 6: Can this data set in its current feature space be separated using a linear separator? Why/why not?

Now, we define a simple feature map $\vec{\phi}(x) = (x, x^2)$ that transforms points in $\mathbb{R}$ to points in $\mathbb{R}^2$.

Problem 7: After applying $\vec{\phi}$ to the data, can it now be separated using a linear separator? Why/why not? (Plotting the data may help you with your answer.)

Problem 8: *Construct* a maximum-margin separating hyperplane (i.e. you do not need to solve a quadratic program). Clearly mark the support vectors. Also draw the resulting decision boundary in the feature space $\vec{\phi}(x) = (x, x^2)$. Is it possible to add another point to the training set in such a way, that the hyperplane *does not* change? Why/why not?

Problem 9: For this specific training set write down the SVM optimization problem, the Lagrangian, the Lagrange dual function and the dual problem.

Problem 10: Write down the KKT conditions for this training set explicitly and verify that the maximum-margin hyperplane you constructed satisfies them.