1 Optimization of SVM

A: Convex Optimization: multiplicity of solution in SVM SVM is based on solving a convex optimization problem, where the objective function $\|w\|^2$ is strictly convex. As discussed in the lecture, while the convex problem admits a single global optimum and hence leads to a unique vector $w\ in \Re^N$, there can be multiple ways in which w is constructed. Indeed, w is constructed as a linear combination of support vectors. If one has at disposal a set of K linearly independent support vectors with $K > N$, then there exists more than one combination of scalars $\alpha_i, i = 1...K$, such that $w = \sum_i = 1^K \alpha_i x^i$ is not unique.

Convince yourself that this is the case when considering the linear SVM case, assuming that $N = 2$ and that you have at your disposal 3 non-zero and not-collinear vector point $x^i, i = 1, 2, 3$ that satisfy the constraint $y_i(w^T x^i + b) = 1, \forall i$. Show that there exist another combination of points that can construct w.

B: Margin The constraints of the SVM problem specify that all support vectors should lie on either of the two hyperplanes parallel to the separating hyperplane with equations $w^T x + b = \pm 1$. Show that the constant 1 is arbitrary and does not affect the solution.

C: Convexity of the relaxed problem: Is $f(w, \xi) = \|w\|^2 + \sum_i \xi$, $\xi > 0 \forall i$ convex?

D: Optimum of the relaxed problem: The introduction of slack variables in the SVM optimization allows to find a solution to a problem that would otherwise been deemed infeasible. The drawback is that the slack leads to solutions that are "suboptimal". Note that the problem remains convex, but the slacks shift the optimum to a value different from the true optimum.

Prove first that the problem remains convex. Recall the conditions for convexity and strict convexity: a convex function f is such that $f(\lambda x + (1 - \lambda)y \leq f(\lambda x) + f((1 - \lambda)y)$. Strict convexity arise when the inequality is replaced by a strict inequality ($<$ in place of $\leq$).

Prove that the optimum in the relaxed problem is identical to the original problem only under certain conditions for the linear SVM problem.