1 Projection formulation

Starting from the following formalization:

$$\min_{\beta_i} \quad \frac{1}{2} \sum_i \sum_j \beta_i \beta_j K(x_i, x_j)$$

$$+ \varepsilon \sum_i |\beta_i|$$

$$- \sum_i y_i \beta_i \tag{1}$$

$$With\ the\ constraints \quad \begin{cases} \sum_i \beta_i = 0 \\ \beta_i \in [-C,\ C] \end{cases}$$

The optimization problem can be solved by the following steps, including the deflection as specified by the task:

Algorithm 1: Compute β

---
1 **while** *optima not found* **do**
2 $\quad v \longleftarrow f(\beta)$
3 $\quad g \longleftarrow \nabla f(\beta)$
4 $\quad$ `check if current solution is best so far (to update)`
5 $\quad d \longleftarrow \gamma g + (1 - \gamma) dprev$
6 $\quad stepsize \longleftarrow \frac{\psi(v - f_{ref} + \delta)}{\|d\|^2}$
7 $\quad \beta \longleftarrow \beta - stepsize \cdot d$
8 $\quad$ `project` β `by solving the KP`
9 **end**

Projection of line 8 is where a simple formulation can be exploited to solve efficiently a Convex Separable Knapsack Problem. The main objective of the formulation is keeping the resulting β_{proj} vector as similar as possible to the one produced by line 7 while at the same time producing a β_{proj} vector satisfying the constraints. Therefore:

$$\min_{\beta_{proj}} \quad \frac{1}{2} \|\beta - \beta_{proj}\|^2$$

$$With\ the\ constraints \quad \begin{cases} \sum_i \beta_{proj}^i = 0 \\ \beta_{proj}^i \in [-C,\ C] \end{cases} \tag{2}$$

Which by lagrangian relaxation leads to:

$$\min_{\beta_{proj}} \quad \frac{1}{2} \sum_i ((\beta^i - \beta^i_{proj})^2 + \mu \beta^i_{proj})$$

$$With\ the\ constraints \qquad \left\{ \ \beta^i_{proj} \in [-C,\ C] \right. \tag{3}$$

2 Alternative formulation

A different concept came up while discussing with each other, probably slightly more intensive on the computational side (because of the more complex constraints), but carrying the same underlying thought process. Instead of focusing on ensuring the constraints exploiting as reference the β vector from line 7 of [1] we could ensure them by exploiting the d vector from line 5. This would imply projecting the deflected direction trying to keep as small as possible the variation to it, which in the language of mathematics results in:

$$\min_{s} \quad \frac{1}{2} \|d - s\|^2 \tag{4}$$

Obviously the constraints are different in this occasion. The only constraints are on β and since the supposedly new β is constructed using the chosen direction (i.e. $\beta = \beta - stepsize \cdot s$)) we can test and ensure the constraints by deriving the ones relative to s:

$$\left\{ \begin{array}{l} s_i \in [\frac{\beta_i - C}{stepsize},\ \frac{\beta_i + C}{stepsize}] \\ \sum_i s_i = \frac{\sum_i \beta_i}{stepsize} \end{array} \right. \tag{5}$$

Which, wrapping up by lagrangian relaxation, leads to:

$$\min_{s} \quad \frac{1}{2} \sum_i (d_i - s_i)^2 + \mu(s_i - \frac{\sum_j \beta_j}{stepsize})$$

$$With\ the\ constraints \qquad \left\{ \ s_i \in [\frac{\beta_i - C}{stepsize},\ \frac{\beta_i + C}{stepsize}] \right. \tag{6}$$

$stepsize$ is obtained using original direction d so to avoid a recursive constraint on s. This is not a dramatic approximation since the objective of the minimization of (4) is keeping s as close as possible to d.

2

3 Analysis of the projection problems

Let's consider the two possible projection problems and try to correlate one with the other.

$$
\begin{aligned}
\textit{Project the direction} &\longrightarrow \min \|d - s\| \\
\textit{Project the point} &\longrightarrow \min \|\beta - \beta_{proj}\|
\end{aligned}
\tag{7}
$$

Consider as starting point the second one.
We can develop the two β by its definition, as we can see at step 7 of 1.

$$
\|\beta - \beta_{proj}\| = \|\beta_{old} - stepsize \cdot d - (\beta_{old} - stepsize' \cdot d')\|
$$

$$
simplifying
\tag{8}
$$

$$
\|\beta - \beta_{proj}\| = \|-stepsize \cdot d + stepsize' \cdot d'\|
$$

Where $stepsize'$ and d' represent the stepsize and direction projected into the feasible region.

As we previously explained for (6) we can approximate $stepsize'$ with $stepsize$ without loss of generality.

$$
\|-stepsize \cdot d + stepsize' \cdot d'\| \approx \|stepsize(d' - d)\|
$$

$$
\tag{9}
$$

$$
stepsize \|d' - d\| \approx \|d' - d\|
$$

So we can conclude that with reasonable assumption and with a small margin of approximation the two problems at the end of the day lead to the same solution.
Obviously this differences will play a role at each step of the iteration, so the analysis of both should be interesting to maybe derive some properties or behaviours wrt solving one projection or the other.