

Projection formulations

There are two different types of projection, the first one is necessary for admissibility and regards the projection of points in the *feasible region*. This problem can be formulated as following:

$$\begin{aligned} \min_{\beta_{proj}} \quad & \frac{1}{2} \|\beta - \beta_{proj}\|^2 \\ \text{With the constraints} \quad & \begin{cases} \sum_i \beta_{proj}^i = 0 \\ \beta_{proj}^i \in [-C, C] \end{cases} \end{aligned} \tag{1}$$

A further not strictly necessary projection (as stated in [1]) regards the one on direction in the *tangent cone*. It is important to **highlight** that this projection is computed before (1).

This projection is useful to maintain coherence between the directions among the iterations. This ensures that at each step the previous d is valid and coherent with the current β value. Even further, this helps avoiding zig-zagging behaviours and may help reaching a better convergence.

$$\begin{aligned} \min_s \quad & \frac{1}{2} \|d - s\|^2 \\ \text{With the constraints} \quad & \left\{ (\beta_i - s_i) \in [-C, C] \right. \end{aligned} \tag{2}$$

First of all no linear constraints are imposed on (2) since the objective of this projection is simply to revise the deflected direction to respect the box constraints defining the tangent cone. This problem's solution does not give us the final β for the iteration, which will be obtained later by (1) ensuring the missing constraint. (Is it correct? Is it safe to ignore this in the first projection since it is ensured in the second one?).

One last fact to remark is that applying only the projection on the direction is not sufficient to bring admissible β . The simple reason stands in the fact that finding a correct direction is not enough: a following incorrect stepsize may lead us to a β not part of the *feasible region*.

So, wrapping up, the abstract procedure will be:

Algorithm 1: Compute β

```

1 while optima not found do
2    $v \leftarrow f(\beta)$ 
3    $g \leftarrow \nabla f(\beta)$ 
4    $d \leftarrow \gamma g + (1 - \gamma)d_{prev}$ 
5   [ optionally project  $d$  as in (2) ]
6    $stepsize \leftarrow \frac{\psi(v - f_{ref} + \delta)}{\|d\|^2}$ 
7    $\beta \leftarrow \beta - stepsize \cdot d$ 
8   project  $\beta$  as in (1)
9 end
```

References

1. Convergence Analysis Of Deflected Conditional Approximate Subgradient Methods - Giacomo D'Antonio and Antonio Frangioni