Conservative OCO

Conventions:
- $\Theta \subseteq \mathbb{R}^d$ is a convex domain.
- $f_t : \Theta \to [\varepsilon_l, \varepsilon_u]$ are convex, differentiable functions. These are loss functions. Notice that each f_t is bounded.
- $T \to$ Time horizon.

OCO Framework: At each $1 \le t \le T$, pick $\theta_t \in \Theta$. loss $f_t(\theta_t)$ is revealed at each step.

- Try to minimize regret

$$R_T(U) := \sum_{t=1}^{T} f_t(\theta_t) - \inf_{\theta} \sum_{t=1}^{T} f_t(\theta)$$

"Conservative OCO": Do the above + do atleast as good as default strategy $\tilde{\theta} \to$ fixed beforehand.

$\rightarrow$

Conservativeness Constraint: Let α be a number, known as the conservativeness level. For all t, we want the following to hold.

$$(*) \qquad L_t \leq (1+\alpha)\tilde{L}_t \quad \forall t \leq T$$

where

$$\tilde{L}_t = \sum_{t=1}^{T} f_t(\tilde{\theta}) \quad \text{(the loss of the default strategy)}$$

Def. Let $\mathcal{U}$ be our algorithm. Define:

$$Z_t[\mathcal{U}] := (1+\alpha)\tilde{L}_t - L_t$$

This quantity is called the budget.

Assumption: $\exists\, \mu > \varepsilon_e$ s.t. $\tilde{L}_t \geq \mu t$ $\forall t \leq T$, i.e the default strategy $\tilde{\theta}$ gives suboptimal regret. We will come up with an algorithm giving $o(t)$ regret ($O(\sqrt{t})$ to be specific)

Remark: Suppose $\mathcal{U}$ is an algorithm $\rightarrow$

Providing regret $R_t(\mu) \le \xi\sqrt{t}$. Then, it has the property that:

$$L_t - \tilde{L}_t \le R_t(\mathcal{U}) \le \xi\sqrt{t}$$

So, if t is s.t

$$\xi\sqrt{t} \le (1+\alpha)\mu t \le (1+\alpha)\tilde{L}_t$$

then clearly

$$L_t \le (1+\alpha)\tilde{L}_t$$

i.e (*) holds. $t \ge \left[\frac{\xi}{\alpha\mu}\right]^2$ satisfy the above. Having said this, we want (*) to actually hold $\forall t \le T$.

THM: High grade of conservativeness is not viable. Formally, $\nexists$ algorithm $\mathcal{U}$ (in the OLO setting) which obtains $L_t \le \tilde{L}_t$ (i.e $\alpha < 0$) unless $\theta_t = \tilde{\theta}\ \forall t$.

Proof: Let k be the first round in which $\mathcal{U}$ plays a choice other than $\tilde{\theta}$, i.e $\theta_k \neq \tilde{\theta}$. Suppose the loss function revealed is the following.

$$\underbrace{f_k(x) = f_k(\hat{\theta})}_{\text{some constant}} + \|\tilde{\theta} - x\|$$

Clearly, in that case, $f_k(\dot{\theta}_t) > f_k(\tilde{\theta})$, and hence $L_k > \tilde{L}_k$.

Rem: In simple words, there is no algorithm that can always do better than a fixed strategy, until each choice made is that strategy.