

# Conservative OCO

Conventions: •  $\Theta \subseteq \mathbb{R}^d$  is a convex domain.

•  $f_t: \Theta \rightarrow [\ell_L, \ell_U]$  are convex, differentiable functions. These are **loss functions**.

Notice that each  $f_t$  is bounded.

•  $T \rightarrow$  **Time horizon**.

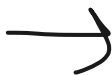
OCO Framework: At each  $1 \leq t \leq T$ , pick  $\theta_t \in \Theta$ . loss  $f_t(\theta_t)$  is revealed at each step.

• Try to minimize **regret**

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

"Conservative OCO": Do the above + do at least as good as

default strategy  $\tilde{\Theta} \rightarrow$  fixed beforehand.



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

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

where

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

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:  $\exists \mu > \epsilon$  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  $R_t(p) \leq \xi \sqrt{t}$ . Then, it has the property that:

$$L_t - \tilde{L}_t \leq R_t(u) \leq \xi \sqrt{t}$$

$$\text{So, if } t \leq s \cdot t \quad \xi \sqrt{t} \leq (1+\alpha) \sqrt{t} \leq (1+\alpha) \tilde{L}_t$$

then clearly

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

i.e (\*) holds.  $t \geq \left[ \frac{\xi}{\alpha p} \right]^2$  satisfy

the above. Having said this, we want (\*) to actually hold  $\forall t \leq T$ .

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

Proof: Let  $k$  be the **first round** in which  $U$  plays a choice other than  $\Theta$ , i.e  $\Theta_k \neq \Theta$ . Suppose the loss function

revealed is the following.

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

Clearly, in that case,  $f_k(\hat{\theta}_t) > f_k(\tilde{\theta})$ ,  
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.