Exercices
Online linear optimization
Université Paris–Saclay

Let $d \geqslant 1$ be an integer and $\mathcal{X}$ a nonempty closed convex subset of $\mathbb{R}^d$.

Exercice 1. — Let $h : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a regularizer, $y, u \in \mathbb{R}^d$ and $x = \nabla h^*(y)$. Prove that

$$\nabla h^*(y+u) = \underset{x' \in \mathbb{R}^d}{\arg\max} \left\{ \langle u, x' \rangle - \mathrm{D}_h(x', x;\ y) \right\}.$$

Exercice 2 (*Dual averaging with time-dependent parameters*). — Prove properties (iii) and (iv) in Proposition 3.2.6 from the lecture notes.

Exercice 3 (*Exponential weights algorithm as mirror descent*). — Let $\mathrm{H}_{\text{ent}} : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be defined as

$$\mathrm{H}_{\text{ent}}(x) = \begin{cases} \sum_{i=1}^d x_i \log x_i & \text{if } x \in (\mathbb{R}_+)^d \\ +\infty & \text{otherwise.} \end{cases}$$

1) Prove that H_{ent} is a mirror map compatible with all nonempty closed convex subsets of $\mathbb{R}_+^d$.

2) Let $(u_t)_{t \geqslant 1}$ be a sequence in $\mathbb{R}^d$, and for $t \geqslant 1$,

$$x_t = \nabla h_{\text{ent}}^* \left(\sum_{s=1}^{t-1} u_s \right),$$

where h_{ent} is the entropic regularizer on the simplex. Prove that $(x_t)_{t\geqslant 1}$ is a sequence of online mirror descent iterates on Δ_d associated with constant mirror map $\mathrm{H}_{\mathrm{ent}}$ and dual increments $(u_t)_{t\geqslant 1}$.

Exercice 4 (*Iterates based on squared Mahalanobis norms*). — Let $(\mathrm{A}_t)_{t\geqslant 1}$ be a sequence of symmetric positive definite matrices of size $d \times d$, and $(u_t)_{t\geqslant 1}$ a sequence in $\mathbb{R}^d$. For each iterates definition below, prove that they are UMD iterates and derive bounds on the regret $\sum_{t=1}^{\mathrm{T}} \langle u_t, x - x_t \rangle$ for $\mathrm{T} \geqslant 1$ and $x \in \mathscr{X}$.

1) Let $x_1 \in \mathscr{X}$ and

$$x_{t+1} = \operatorname*{arg\,min}_{x \in \mathscr{X}} \left\|(x_t + \mathrm{A}_t^{-1} u_t) - x\right\|_{\mathrm{A}_t}, \qquad t \geqslant 1.$$

2) Let $x_1 \in \mathscr{X}$ and

$$x_{t+1} = \operatorname*{arg\,min}_{x \in \mathscr{X}} \left\{ - \langle \mathrm{A}_t x_t + u_t, x \rangle - \frac{1}{2} x^\top \mathrm{A}_{t+1} x \right\}, \qquad t \geqslant 1.$$

3) Let $y_1 \in \mathbb{R}^d$ and

$$x_t = \operatorname*{arg\,min}_{x \in \mathscr{X}} \left\{ - \left\langle y_1 + \sum_{s=1}^{t-1} u_s, x \right\rangle + \frac{1}{2} \langle x, \mathrm{A}_t x \rangle \right\}, \qquad t \geqslant 1.$$

Exercice 5 (*Sparse payoff vectors*). — Consider online linear optimization on $\mathscr{X} = \Delta_d$. Let $1 \leqslant s \leqslant d$ be an integer, assume that payoff vectors $(u_t)_{t\geqslant 1}$ are in $[0,1]^d$ and that for all $t \geqslant 1$, u_t has at most s nonzero components.

1) Using a constant ℓ_p regularizer (or mirror map) for a well-chosen value p, derive the best possible regret bound. *Recall that for* $1 < p < 2$, $x \mapsto \frac{1}{2}\|x\|_p^2$ *is* $(p-1)$*-strongly convex with respect to* $\|\cdot\|_p$.

2) Using time-dependent regularizers (or mirror maps), derive the best possible *horizon-free* regret bound.

Exercice 6 (*A hybrid of mirror descent and dual averaging*). — Let H be a mirror map compatible with $\mathcal{X}$, $(u_t)_{t \geqslant 1}$ be a sequence in $\mathbb{R}^d$, $(\eta_t)_{t \geqslant 1}$ a positive sequence, and $x_1 \in \mathcal{X} \cap \operatorname{dom} H$. Then define

$$x_{t+1} = \operatorname*{arg\,max}_{x \in \mathcal{X}} \left\{ \langle \nabla H(x_t) + \eta_t u_t, x \rangle - \frac{\eta_t}{\eta_{t+1}} H(x) \right\}, \qquad t \geqslant 1.$$

One special case of interest is when $(\eta_t)_{t \geqslant 1}$ is nonincreasing.

1) Prove that the above can be seen as UMD iterates.

2) Is there cases where the above are MD with step-sizes? MD with parameters?

3) Derive regret bounds.

Exercice 7 (*Exponential weights with step-sizes*). — Let $(u_t)_{t \geqslant 1}$ be a sequence in $\mathbb{R}^d$ and $(\gamma_t)_{t \geqslant 1}$ a nonincreasing sequence in $\mathbb{R}^d$. Consider:

$$x_t = \left(\frac{\exp\left(\left(\sum_{s=1}^{t-1} \gamma_s u_{s,i} \right) \right)}{\sum_{j=1}^{d} \exp\left(\sum_{s=1}^{t-1} \gamma_s u_{s,j} \right)} \right)_{1 \leqslant i \leqslant d}. \tag{1}$$

1) Let $T \geqslant 1$ and $x \in \Delta_d$. Derive a general bound on $\sum_{t=1}^{T} \langle u_t, x - x_t \rangle$.

2) Derive a regret bound in the case where there exists $L > 0$ such that $\|u_t\|_\infty \leqslant L$ for all $t \geqslant 1$.

3) In the multi-armed bandit problem, consider the variant of EXP3 based on (1), and derive a guarantee with similar assumptions as for EXP3 in the course.