# Regret bounds

*and their applications to online learning, optimization and games*

Joon Kwon

March 20, 2025

# Contents

# Foreword

As of 2025, this document contains lecture notes from a course given in Master 2 in *Université Paris–Saclay*. These are highly incomplete and constantly updated as the lectures are given.

The course proposes a unified presentation of regret minimization for adversarial online learning problems, and its application to various problems such as Blackwell's approachability, optimization algorithms (GD, Nesterov, SGD, AdaGrad), variational inequalities with monotone operators (Extragradient, Mirror-Prox, Dual Extrapolation), fixed-point iterations (Krasnoselskii-Mann), and games. The presentation aims at being modular, so that introduced tools and techniques could easily be used to define and analyze new algorithms.

The central notion of this presentation is the *regret*, which will be analyzed using the Legendre–Fenchel transform and Bregman divergences. An excellent recent monograph on the topic of online learning is the following:

- Francesco Orabona. A modern introduction to online learning. *arXiv:1912.13213*, 2023.

Additional notable references on the topic include:

- H. Brendan McMahan. A survey of algorithms and analysis for adaptive online learning. *The Journal of Machine Learning Research*, 18(1):3117–3166, 2017,

- Elad Hazan. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016,

- Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2011,

- Sébastien Bubeck. *Introduction to Online Optimization: Lecture Notes.* Princeton University, 2011,

- Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games.* Cambridge University Press, 2006.

Regarding convex analysis, we refer to the following classical book:

- R. Tyrrell Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

## Acknowledgements

These notes highly benefited from the help of Anatoli Juditsky, discussions with Sylvain Sorin, Gilles Stoltz, Marius Potfer, Bruno Ziliotto, Jaouad Mourtada, Côme Fiegel, Vianney Perchet, and the encouragements from Liliane Bel.

Typos have been found by Pablo Hueso Merino and Yanyu Zhou.

# Chapter 1

# Convexity tools

We present the basic convexity notions and tools that will be used in the subsequent chapters. Most of them are classical and are given without proof.

## 1.1 Preliminaries

Let $d \geqslant 1$. Throughout the chapter, we consider Euclidean space $\mathbb{R}^d$ equipped with its canonical inner product denoted $\langle \cdot \, , \cdot \rangle$. For $1 \leqslant p < +\infty$, the $\ell^p$ norm is defined as

$$\|x\|_p = \left( \sum_{i=1}^{d} x_i^p \right)^{1/p}, \quad x \in \mathbb{R}^d,$$

and the $\ell^\infty$ norm as

$$\|x\|_\infty = \max_{1 \leqslant i \leqslant d} |x_i| \,, \quad x \in \mathbb{R}^d.$$

We denote as $\mathcal{S}_d^{+}(\mathbb{R})$ (resp. $\mathcal{S}_d^{++}(\mathbb{R})$) the set of symmetric semi-definite positive matrices (resp. symmetric definite positive matrices). For $A \in \mathcal{S}_d^{++}(\mathbb{R})$, the Mahalanobis norm associtad with $A$ is defined as

$$\|x\|_A = \sqrt{\langle x, Ax \rangle}, \quad x \in \mathbb{R}^d.$$

For a set $\mathcal{X} \subset \mathbb{R}^d$, $\operatorname{int} \mathcal{X}$ and $\operatorname{cl} \mathcal{X}$ denote its interior and closure, respectively.

**Definition 1.1.1** (Domain of a function)**.** The *domain* of a function $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is the set

$$\operatorname{dom} f = \left\{ x \in \mathbb{R}^d, \; f(x) < +\infty \right\}.$$

$f$ is said to be *proper* if its domain is nonempty.

**Definition 1.1.2** (Dual norm)**.** Let $\|\cdot\|$ be a norm in $\mathbb{R}^d$. Its *dual norm* is defined as

$$\|y\|_* = \max_{\|x\|\leqslant 1} \langle y, x\rangle , \quad y \in \mathbb{R}^d.$$

*Remark* 1.1.3. The above maximum is indeed attained because for a given $y \in \mathbb{R}^d$, function $x \mapsto \langle y, x\rangle$ is continuous on the closed unit ball, which is compact.

**Proposition 1.1.4.** *The dual norm is ineed a norm.*

**Proposition 1.1.5.** *Let $\|\cdot\|$ be a norm in $\mathbb{R}^d$. Then, $\|\cdot\|_{**} = \|\cdot\|$.*

**Example 1.1.6** (Common dual norms)**.** In $\mathbb{R}^d$, $\ell_2$ is its own dual norm, $\ell_p$ and $\ell_q$ (with $p, q \geqslant 1$ such that $1/p + 1/q = 1$) are dual of each other, and $\ell_1$ and $\ell_\infty$ are dual of each other. If $A \in \mathcal{S}_d^{++}(\mathbb{R})$, the dual norm of the associated Mahalanobis norm is the Mahalanobis norm associated with $A^{-1}$.

*Remark* 1.1.7. It follows from the definition of the dual norm that for all $x, y \in \mathbb{R}^d$, $\langle y, x\rangle \leqslant \|y\|_* \|x\|$, which, together with the above examples recovers Cauchy-Schwarz and Hölder's inequalities.

For a function $f : \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ and $x \in \operatorname{int}\operatorname{dom} f$, if $f$ is differentiable in $x$ (resp. twice differentiable), we denote $\nabla f(x)$ its gradient at $x$ (resp. $\nabla^2 f(x)$ its Hessian matrix at $x$).

## 1.2 Convexity

**Definition 1.2.1.** A set $\mathcal{X} \subset \mathbb{R}^d$ is *convex* if for all $x, x' \in \mathcal{X}$ and $\lambda \in [0, 1]$, $\lambda x + (1 - \lambda)x' \in \mathcal{X}$.

**Example 1.2.2** (Unit balls)**.** For all norms, as an immediate consequence of the triangle inequality, the unit ball is convex.

**Example 1.2.3** (Simplex)**.** Denote $\Delta_d$ the simplex in $\mathbb{R}^d$:

$$\Delta_d = \left\{ x \in \mathbb{R}_+^d,\ \sum_{i=1}^d x_i = 1 \right\},$$

which is a closed convex set. Note that it is contained in a hyperplane and therefore has empty interior.

**Proposition 1.2.4** (Euclidean projection on a closed convex set)**.** *Let $\mathcal{X} \subset \mathbb{R}^d$ be a nonempty closed convex set and $x \in \mathbb{R}^d$. Then, the Euclidean projection of $x$ onto $\mathcal{X}$ exists an is unique. In other words, $x' \mapsto \|x' - x\|_2^2$ admits a unique minimizer on $\mathcal{X}$.*

**Definition 1.2.5** (Convex functions)**.** A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is *convex* if for all $x, x' \in \mathbb{R}^d$ and $\lambda \in [0,1]$,

$$f(\lambda x + (1-\lambda)x') \leqslant \lambda f(x) + (1-\lambda) f(x').$$

$f$ is *strictly convex* if the above inequality is strict for $\lambda \in (0,1)$.

*Remark* 1.2.6. The convexity of a function $f$ is closely related to the convexity of sets, as the former can be equivalently defined as the epigraph

$$\operatorname{epi} f = \left\{ (x,a) \in \mathbb{R}^d \times \mathbb{R},\ a \geqslant f(x) \right\}$$

being convex.

*Remark* 1.2.7. The domain of a convex function is convex.

**Example 1.2.8.** The following functions are convex: linear functions, quadratic functions of the form $x \mapsto \langle x, Ax \rangle$ where $A$ is a positive semi-definite matrix, the exponential, the negative logarithm, convex combinations of convex functions, the point-wise supremum of convex functions. Let $\|\cdot\|$ be a norm in $\mathbb{R}^d$ and $a \geqslant 1$. Function $x \mapsto \|x\|^a$ is convex.

**Proposition 1.2.9** (Jensen's inequality)**.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a convex function, and $X$ a random variable with values in $\operatorname{dom} f$ so that $\mathbb{E}[X]$ exists. Then, $f$ is measurable, $\mathbb{E}[f(X)]$ exists in $\mathbb{R} \cup \{+\infty\}$ and*

$$f(\mathbb{E}[X]) \leqslant \mathbb{E}[f(X)].$$

*In particular, for $n \geqslant 1$, $x_1, \dots, x_n \in \mathbb{R}^n$ and $\lambda_1, \dots, \lambda_n \geqslant 0$ such that $\lambda_1 + \cdots + \lambda_n > 0$,*

$$f\left( \frac{\sum_{i=1}^n \lambda_i x_i}{\sum_{i=1}^n \lambda_i} \right) \leqslant \frac{\sum_{i=1}^n \lambda_i f(x_i)}{\sum_{i=1}^n \lambda_i}.$$

**Proposition 1.2.10** (First and second order characterizations of convexity)**.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ a function with open domain.*

*(i) If $f$ is differentiable on its domain, $f$ is convex if, and only if, for all $x, x' \in \operatorname{dom} f$,*

$$f(x') \geqslant f(x) + \left\langle \nabla f(x), x' - x \right\rangle.$$

*(ii) If $f$ is twice differentiable on its domain, $f$ is convex if, and only if, for all $x \in \operatorname{dom} f$, $\nabla^2 f(x)$ is positive semi-definite.*

**Proposition 1.2.11.** *Let $\mathcal{X} \subset \mathbb{R}^d$ be a nonempty convex set, and $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a convex function, differentiable on an open set containing $\mathcal{X}$. Then, $x_* \in \mathcal{X}$ is a minimizer of $f$ on $\mathcal{X}$ if, and only if,*

$$\forall x \in \mathcal{X},\ \langle \nabla f(x_*), x - x_* \rangle \geqslant 0.$$

**Definition 1.2.12** (Lower semicontinuity)**.** A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous if for all $a \in \mathbb{R}$, the set $\left\{x \in \mathbb{R}^d,\ f(x) \leqslant a\right\}$ is closed.

*Remark* 1.2.13. A continuous function is lower semicontinuous. The sum of two lower semicontinuous functions is also lower semicontinuous.

**Example 1.2.14.** Let $\mathcal{X} \subset \mathbb{R}^d$ be a convex set. The *convex indicator* of $\mathcal{X}$ is the convex function defined as

$$I_{\mathcal{X}}(x) = \begin{cases} 0 & \text{if } x \in \mathcal{X} \\ +\infty & \text{otherwise,} \end{cases}$$

which is lower semicontinuous if, and only if $\mathcal{X}$ is closed.

**Example 1.2.15** (Negative entropy on the simplex)**.** The function $h_{\mathrm{ent}} : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ defined as

$$h_{\mathrm{ent}}(x) = \begin{cases} \sum_{i=1}^d x_i \log x_i & \text{if } x \in \Delta_d \\ +\infty & \text{otherwise,} \end{cases}$$

with convention $0 \log 0 = 0$, is a lower semicontinuous convex function. It will also be called the *entropic regularizer* and studied in Chapter 3.

**Proposition 1.2.16.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function and $\mathcal{X}_0 \subset \mathbb{R}^d$ a compact set such that $\operatorname{dom} f \cap \mathcal{X}_0 \neq \varnothing$. Then, $f$ attains a minimum on $\mathcal{X}_0$.*

## 1.3 Subgradients

**Definition 1.3.1** (Subgradients)**.** Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ and $x, y \in \mathbb{R}^d$. $y$ is a *subgradient* of $f$ at $x$ if for all $x' \in \mathbb{R}^d$,

$$f(x') \geqslant f(x) + \left\langle y, x' - x \right\rangle .$$

The set of all subgradients of $f$ at $x$ is called the *subdifferential* of $f$ at $x$ and is denoted as $\partial f(x)$.

**Example 1.3.2** (Absolute value)**.** For $f : x \mapsto |x|$ defined on $\mathbb{R}$, the subdifferential is given by

$$\partial f(x) = \begin{cases} \{-1\} & \text{if } x < 0 \\ [-1, 1] & \text{if } x = 0 \\ \{1\} & \text{if } x > 0. \end{cases}$$

**Proposition 1.3.3.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$. A point $x_* \in \mathbb{R}^d$ is a global minimizer of $f$ if, and only if $0 \in \partial f(x_*)$.*

*Proof.* $x_*$ being a global minimizer can be written

$$\forall x \in \mathbb{R}^d, \quad f(x) \geqslant f(x_*) + \langle 0, x - x_* \rangle \,,$$

in other words $0 \in \partial f(x_*)$. $\square$

**Proposition 1.3.4.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a convex function and $x \in \operatorname{int}\operatorname{dom} f$. $f$ is differentiable in $x$ if, and only if, $\partial f(x)$ is a singleton. When this is the case, $\partial f(x) = \{\nabla f(x)\}$.*

*Remark* 1.3.5. Even in the case of a point in the domain of a convex function, the subdifferential may be empty. Consider for instance $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ defined as:

$$f(x) = \begin{cases} -\sqrt{x} & \text{if } x \geqslant 0 \\ +\infty & \text{if } x < 0, \end{cases}$$

which is a proper lower semicontinuous convex function. 0 belongs to the domain of $f$ and yet, $\partial f(0) = \varnothing$.

**Proposition 1.3.6** (see e.g. Theorem 23.4 in [Roc70])**.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function. If $x \notin \operatorname{dom} f(x)$, then $\partial f(x) = \varnothing$ and if $x \in \operatorname{int}\operatorname{dom} f$, then $\partial f(x) \neq \varnothing$.*

**Proposition 1.3.7.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function, $\|\cdot\|$ a norm on $\mathbb{R}^d$, and $L > 0$. Then, $f$ is $L$-Lipschitz continuous on* $\operatorname{int}\operatorname{dom} f$ *with respect to $\|\cdot\|$ if, and only if:*

$$\forall x \in \operatorname{int}\operatorname{dom} f, \ \forall y \in \partial f(x), \quad \|y\|_* \leqslant L.$$

## 1.4 Legendre–Fenchel transform

**Definition 1.4.1.** Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper function. The *Legendre–Fenchel transform* (or *convex conjugate*) of $f$ is a function $f^* : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ defined as

$$f^*(y) = \sup_{x \in \mathbb{R}^d} \left\{ \langle y, x \rangle - f(x) \right\}, \quad y \in \mathbb{R}^d.$$

*Remark* 1.4.2. In the above definition, for a given $y \in \mathbb{R}^d$, because $f$ is assumed proper, quantity $\langle y, x \rangle - f(x)$ is not $-\infty$ for at least some point $x \in \mathbb{R}^d$, and therefore, the supremum is indeed a value in $\mathbb{R} \cup \{+\infty\}$.

*Remark* 1.4.3 (Fenchel's inequality). It follows from the above definition that for all $x, y \in \mathbb{R}^d$, $\langle y, x \rangle \leqslant f(x) + f^*(y)$.

*Remark* 1.4.4 (Legendre–Fenchel transformation is order-reversing). If $f, g : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ are two functions such that $f \leqslant g$, then $g^* \geqslant f^*$ as an immediate consequence of the definition of the convex conjugation.

**Proposition 1.4.5.** *If $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is a proper function, $f^*$ is lower semicontinuous and convex.*

**Theorem 1.4.6** (Fenchel–Moreau)**.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper function. Then, $f$ is lower semicontinuous and convex if, and only if $f = f^{**}$. In this case, $f^*$ is proper.*

**Proposition 1.4.7.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous convex function and $x, y \in \mathbb{R}^d$. The following statements are equivalent:*

*(i)* $x \in \partial f^*(y)$,

*(ii)* $y \in \partial f(x)$,

*(iii)* $\langle y, x\rangle = f(x) + f^*(y)$,

*(iv)* $x \in \operatorname{Arg\,max}_{x' \in \mathbb{R}^d} \{\langle y, x'\rangle - f(x')\}$,

*(v)* $y \in \operatorname{Arg\,max}_{y' \in \mathbb{R}^d} \{\langle y', x\rangle - f^*(y')\}$.

$$\begin{tikzcd} x \arrow[r, bend left, "\partial f"] & y \arrow[l, bend left, "\partial f^*"] \end{tikzcd}$$

Figure 1.1: If $f$ is proper lower semicontinus and convex, the set-valued mappings $\partial f$ and $\partial f^*$ are inverses of each other.

**Example 1.4.8** (Norms and squared norms)**.** Let $\|\cdot\|$ be a norm in $\mathbb{R}^d$ and denote $B$ its closed unit ball. Then, $I_B$ is then a proper lower semicontinuous convex function and $I_B^* = \|\cdot\|_*$. Therefore, the involutional property of the Legendre-Fenchel transform is an extension of the involutional property for dual norms. Besides, if $f : x \mapsto \frac{1}{2}\|x\|^2$, then for $y \in \mathbb{R}^d$, $f^*(y) = \frac{1}{2}\|y\|_*^2$.

## 1.5 Bregman divergences

We now define a large class of similarity measures in $\mathbb{R}^d$ called Bregman divergences, which in general are not distances because they may fail to be symmetric. They contain the squared Euclidean norm and the Kullback–Leibler divergence as special cases. Bregman divergences are used as an alternative geometry to the Euclidean one when it comes to e.g. defining and analyzing iterative algorithms. We present the classical definition which involves a gradient.

**Definition 1.5.1.** Let $\mathcal{X} \subset \mathbb{R}^d$, $f : \mathcal{X} \to \mathbb{R}$, $x \in \operatorname{int} \mathcal{X}$ and $x' \in \mathcal{X}$ such that $f$ is differentiable in $x$. Then, the *Bregman divergence* from $x$ to $x'$ is defined as

$$D_f(x', x) = f(x') - f(x) - \left\langle \nabla f(x), x' - x \right\rangle .$$

*Remark* 1.5.2. $D_f(x, x')$ is the remainder of the first order Taylor's expansion from $x$ to $x'$, and is a measure of the curvature of $f$ between those two points. The Bregman divergence is nonnegative as soon as $f$ is convex. In the case of a linear function $f$, the Bregman divergence is zero, which corresponds to the linear function having no curvature.

The above definition which requires the differentiability at starting point $x$ is the most common. For our purposes however, we also consider the following generalization (proposed in [JKM23]) which involves a subgradient instead of a gradient.

**Definition 1.5.3** (Generalized Bregman divergences)**.** Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a function, $x, x', y \in \mathbb{R}^d$ such that $x \in \operatorname{dom} f$ and $y \in \partial f(x)$. The *Bregman divergence* from $x$ to $x'$ with subgradient $y$ is then defined as

$$D_f(x', x;\ y) = f(x') - f(x) - \left\langle y, x' - x \right\rangle .$$

*Remark* 1.5.4. The above generalized Bregman divergence may not exist even when $x$ belongs to the domain of $f$, as the subdifferential may be empty. When it exists, because $x \in \operatorname{dom} f$, it belongs to $\mathbb{R} \cup \{+\infty\}$ and is nonnegative when $f$ is convex. When $f$ is convex and differentiable at point $x$, the only subgradient at $x$ is $\nabla f(x)$ according to Proposition 1.3.4, and the two previous definitions coincide.

**Proposition 1.5.5.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous convex function and $x, x', y, y' \in \mathbb{R}^d$ such that $y \in \partial f(x)$ and $y' \in \partial f(x')$. Then,*

$$D_f(x', x;\ y) = D_{f^*}(y, y';\ x').$$

*Proof.*

$$\begin{aligned} D_f(x', x;\ y) - D_{f^*}(y, y';\ x') &= f(x') - f(x) - \left\langle y, x' - x \right\rangle \\ &\quad - f^*(y) + f^*(y') + \left\langle x', y -' y \right\rangle \\ &= \left\langle x', y' \right\rangle - \langle x, y \rangle - \left\langle y, x' - x \right\rangle + \left\langle x', y - y' \right\rangle \\ &= 0, \end{aligned}$$

where for the second equality, we applied Fenchel's identity from property (iii) in Proposition 1.4.7. □

**Example 1.5.6** (Squared Euclidean norm)**.** Consider the squared Euclidean norm $H_2 : x \mapsto \frac{1}{2} \|x\|_2^2$, which is differentiable in $\mathbb{R}^d$. Then for all $x, x' \in \mathbb{R}^d$, $D_{H_2}(x', x) = \frac{1}{2} \|x' - x\|_2^2$.

**Example 1.5.7** (Squared Mahalanobis norm)**.** Let $A$ be a symmetric positive definite matrix of size $d$. Consider the squared Mahalanobis norm $H_A : x \mapsto \frac{1}{2} \|x\|_A^2$, which is differentiable in $\mathbb{R}^d$. Then for all $x, x' \in \mathbb{R}^d$, $D_{H_A}(x', x) = \frac{1}{2} \|x' - x\|_A^2$.

## 1.6 Strong convexity and smoothness

We now introduce strongly convexity and smoothness which intuitively correspond to the curvature of a function being respectively bounded from below (by a positive number), and bounded from above in absolute value.

**Definition 1.6.1** (Strong convexity)**.** Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, $\|\cdot\|$ a norm in $\mathbb{R}^d$ and $K > 0$. $f$ is *$K$-strongly convex* with respect to $\|\cdot\|$ if for all $x, x' \in \mathbb{R}^d$ and $\lambda \in [0, 1]$,

$$f(\lambda x + (1 - \lambda)x') \leqslant \lambda f(x) + (1 - \lambda) f(x') - \frac{K\lambda(1 - \lambda)}{2} \left\|x' - x\right\|^2.$$

**Definition 1.6.2** (Smoothness)**.** Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function, $\|\cdot\|$ a norm in $\mathbb{R}^d$ and $L > 0$. $f$ is *$L$-smooth* with respect to $\|\cdot\|$ if for all $x, x' \in \mathbb{R}^d$, $|D_f(x', x)| \leqslant \frac{L}{2} \|x' - x\|^2$, in other words,

$$\left| f(x') - f(x) - \left\langle \nabla f(x), x' - x \right\rangle \right| \leqslant \frac{L}{2} \left\|x' - x\right\|^2.$$

*Remark* 1.6.3. If $f$ is convex, the above definition reduces to $D_f(x', x) \leqslant \frac{L}{2} \|x' - x\|_2^2$ $(x, x' \in \mathbb{R}^d)$.

**Proposition 1.6.4** (Duality between strong convexity and smoothness)**.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper, lower semicontinuous convex function, $\|\cdot\|$ a norm in $\mathbb{R}^d$ and $K > 0$. The following statements are equivalent.*

*(i) $f$ is $K$-strongly convex with respect to $\|\cdot\|$.*

*(ii) For all $x, x', y \in \mathbb{R}^d$ such that $y \in \partial f(x)$, $D_f(x', x;\ y) \geqslant \frac{K}{2} \|x' - x\|^2$, in other words*

$$f(x') \geqslant f(x) + \left\langle y, x' - x \right\rangle + \frac{K}{2} \left\|x' - x\right\|^2.$$

*(iii) For all $x, x', y, y' \in \mathbb{R}^d$ such that $y \in \partial f(x)$ and $y' \in \partial f(x')$,*

$$\left\langle y' - y, x' - x \right\rangle \geqslant K \left\|x' - x\right\|^2.$$

*(iv) For all $x, x', y, y' \in \mathbb{R}^d$ such that $y \in \partial f(x)$ and $y' \in \partial f(x')$,*

$$\left\langle y' - y, x' - x \right\rangle \leqslant \frac{1}{K} \left\|y' - y\right\|_*^2.$$

(v) *For all* $x, x', y, y' \in \mathbb{R}^d$ *such that* $y \in \partial f(x)$ *and* $y' \in \partial f(x')$, $D_f(x', x;\ y) \leqslant \frac{1}{2K} \left\| y' - y \right\|_*^2$, *in other words*

$$f(x') \leqslant f(x) + \langle y, x' - x \rangle + \frac{1}{2K} \left\| y' - y \right\|_*^2 .$$

(vi) $f^*$ *is differentiable on* $\mathbb{R}^d$ *and* $1/K$*-smooth with respect to* $\| \cdot \|_*$.

**Corollary 1.6.5.** *Let* $K > 0$, $\| \cdot \|$ *a norm on* $\mathbb{R}^d$ *and* $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ *a proper lower semicontinuous function which we assume* $K$*-strongly convex with respect to* $\| \cdot \|$. *Then, for all* $y, y' \in \mathbb{R}^d$,

$$D_{f^*}(y', y) \leqslant \frac{1}{2K} \left\| y' - y \right\|_*^2 .$$

**Proposition 1.6.6** (Second order sufficient condition for strong convexity)**.** *Let* $f : \mathbb{R}^d \to \mathbb{R}$ *be a twice differentiable function,* $\| \cdot \|$ *a norm in* $\mathbb{R}^d$ *and* $K > 0$ *such that*

$$\forall x \in \mathbb{R}^d,\ \forall u \in \mathbb{R}^d, \quad \langle u, \nabla^2 f(x) u \rangle \geqslant K \left\| u \right\|^2 .$$

*Then,* $f$ *is* $K$*-strongly convex with respect to* $\| \cdot \|$.

**Corollary 1.6.7.** *The squared Euclidean norm* $h_2 : x \mapsto \frac{1}{2} \left\| x \right\|_2^2$ *is* $1$*-strongly convex and* $1$*-smooth with respect to* $\| \cdot \|_2$.

**Corollary 1.6.8.** *Let* $A$ *be a symmetric positive definite matrix of size* $d$. *The associatd squared Mahalanobis norm* $x \mapsto \frac{1}{2} \left\| x \right\|_A^2$ *is* $1$*-strongly convex for* $\| \cdot \|_A$.

**Proposition 1.6.9.** *The negative entropy* $h_{ent}$ *is* $1$*-strongly convex with respect to* $\ell_1$.

**Proposition 1.6.10.** *A proper lower semicontinuous strongly convex function* $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ *admits a unique minimizer on* $\mathbb{R}^d$.

# Chapter 2

# UMD theory

Throughout the chapter, $\mathcal{X}$ is a nonempty closed convex set of $\mathbb{R}^d$.

## 2.1 Introduction

The goal of this chapter is to introduce a general scheme for defining a sequence of iterates $(x_t)_{t\geqslant 1}$ in $\mathcal{X}$ based on another sequence $(u_t)_{t\geqslant 1}$ in $\mathbb{R}^d$, where for each $t \geqslant 1$, vector $u_t$ is used in the update from $x_t$ to $x_{t+1}$. General properties are then established. All algorithms and guarantees from the following chapters will be derived using this general approach, called UMD for *unified mirror descent*. To get some taste and intuition, we first examine a few simple special cases before presenting our general theory.

The simplest update is given by

$$x_{t+1} = x_t + u_t, \quad t \geqslant 1,$$

and is already of great interest, as it contains as special cases gradient descent (where $u_t = -\gamma_t \nabla f(x_t)$ is then a step in the opposite direction of the gradient of some objective function), as well as its stochastic counterpart (SGD). Such a sequence satisfies the following elementary result.

**Proposition 2.1.1.** *For all $t \geqslant 1$ and $x \in \mathbb{R}^d$,*

$$\langle u_t, x - x_t \rangle = \frac{1}{2} \left\| x - x_t \right\|_2^2 - \frac{1}{2} \left\| x - x_{t+1} \right\|_2^2 + \frac{1}{2} \left\| u_t \right\|_2^2 .$$

*Consequently, for all $T \geqslant 1$,*

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle = \frac{1}{2} \left\| x - x_1 \right\|_2^2 - \frac{1}{2} \left\| x - x_{T+1} \right\|_2^2 + \frac{1}{2} \sum_{t=1}^{T} \left\| u_t \right\|_2^2 .$$

*Proof.* Let $t \geqslant 1$. Using the definition of $x_{t+1}$,

$$\left\| x_{t+1} - x \right\|_2^2 = \left\| x_t + u_t - x \right\|_2^2 = \left\| x_t - x \right\|_2^2 + 2 \langle u_t, x_t - x \rangle + \left\| u_t \right\|_2^2 ,$$

and the result follows. □

For instance, the classical convergence guarantees about (stochastic) gradient descent in various settings, are consequences of the above identity. The quantity $\sum_{t=1}^{T} \langle u_t, x - x_t \rangle$ is called the *regret*, but the corresponding interpretation will be presented in the next chapter only. Some intuition about the above can be obtained through a continuous-time counterpart:

$$\text{if} \quad \frac{d\tilde{x}_t}{dt} = \tilde{u}_t, \quad \text{then} \quad \frac{d}{dt}\left( \frac{1}{2} \left\| x - \tilde{x}_t \right\|_2^2 \right) = \langle \tilde{u}_t, \tilde{x}_t - x \rangle .$$

Therefore, going back to discrete-time, one can interpret the difference

$$\frac{1}{2} \left\| x - x_t \right\|_2^2 - \frac{1}{2} \left\| x - x_{t+1} \right\|_2^2$$

as a discrete-time derivative, and the term $\frac{1}{2} \left\| u_t \right\|_2^2$—which does not appear in continuous-time—as a discretization error.

The quantity $\frac{1}{2} \left\| x - x_t \right\|_2^2$ also appears in the following alternative expression which will inspire generalizations.

**Proposition 2.1.2.** *For $t \geqslant 1$, $x_{t+1} = x_t + u_t$ if, and only if,*

$$x_{t+1} = \operatorname*{arg\,min}_{x \in \mathbb{R}^d} \left\{ - \langle u_t, x \rangle + \frac{1}{2} \left\| x - x_t \right\|_2^2 \right\} .$$

The above expression is an *incremental* point of view, in the sense that the next iterate $x_{t+1}$ is written as a function of previous iterate $x_t$ and vector $u_t$ only. Another equivalent formulation is the *cumulative* one:

$$x_{t+1} = x_1 + \sum_{s=1}^{t} u_s,$$

where the next iterate $x_{t+1}$ is only a function of the sum of the previous vectors $u_s$ ($1 \leqslant s \leqslant t$) (*plus* the initial point $x_1$). These two points of view will yield different extensions below.

We now turn to a constrained setting were we need the iterates to all lie in a given nonempty closed convex set $\mathcal{X} \subset \mathbb{R}^d$. Then, the definition of the iterates can be adapted by adding a projection step onto $\mathcal{X}$ with respect to the Euclidean distance. The projected gradient descent algorithm is a special case. Then, the iterates satisfy the following *regret bound*, where the comparison point $x$ must lie in $\mathcal{X}$.

**Proposition 2.1.3.** *For $t \geqslant 1$, $x_{t+1} = \operatorname{arg\,min}_{x \in \mathcal{X}} \left\| x_t + u_t - x \right\|_2^2$ if, and only if,*

$$x_{t+1} = \operatorname*{arg\,min}_{x \in \mathcal{X}} \left\{ - \langle u_t, x \rangle + \frac{1}{2} \left\| x - x_t \right\|_2^2 \right\} .$$

*In that case, for all $x \in \mathcal{X}$,*

$$\langle u_t, x - x_t \rangle \leqslant \frac{1}{2} \left\| x - x_t \right\|_2^2 - \frac{1}{2} \left\| x - x_{t+1} \right\|_2^2 + \frac{1}{2} \left\| u_t \right\|_2^2 .$$

Another possibility for constraining the iterates in $\mathcal{X}$ is the following, where the vector $u_t$ is not added to the current iterate $x_t$ as above, but to the point before the projection onto $\mathcal{X}$.

**Proposition 2.1.4.** *If $(x_t)_{t\geqslant 1}$ and $(y_t)_{t\geqslant 1}$ satisfy*

$$y_{t+1} = y_t + u_t \quad \textit{and} \quad x_{t+1} = \operatorname*{arg\,min}_{x\in\mathcal{X}} \left\| y_{t+1} - x \right\|_2^2, \qquad t \geqslant 1,$$

*then for all $x \in \mathcal{X}$ and $t \geqslant 1$,*

$$\langle u_t, x - x_t \rangle \leqslant D_t(x) - D_{t+1}(x) + \frac{1}{2} \left\| u_t \right\|_2^2,$$

*where for all $s \geqslant 1$,*

$$D_s(x) = \frac{1}{2} \left\| x \right\|_2^2 - \frac{1}{2} \left\| x_s \right\|_2^2 - \langle y_s, x - x_s \rangle .$$

We now go back to the unconstrained case ($\mathcal{X} = \mathbb{R}^d$) and consider the following generalization.

**Proposition 2.1.5.** *For a symmetric positive definite matrix $A \in \mathbb{R}^{d\times d}$ and $t \geqslant 1$, $x_{t+1} = x_t + A^{-1}u_t$ if, and only if*

$$x_{t+1} = \operatorname*{arg\,min}_{x\in\mathbb{R}^d} \left\{ - \langle u_t, x \rangle + \frac{1}{2} \left\| x - x_t \right\|_A^2 \right\}.$$

*In that case, for all $x \in \mathbb{R}^d$,*

$$\langle u_t, x - x_t \rangle = \frac{1}{2} \left\| x - x_t \right\|_A^2 - \frac{1}{2} \left\| x - x_{t+1} \right\|_A^2 + \frac{1}{2} \left\| u_t \right\|_{A^{-1}}^2 .$$

The last example considers the simplex $\mathcal{X} = \Delta_d$, and appears at first sight to be quite different from the above.

**Proposition 2.1.6.** *If for $t \geqslant 1$,*

$$x_{t+1} = \left( \frac{x_{t,i} \exp\left(u_{t,i}\right)}{\sum_{j=1}^d x_{t,j} \exp\left(u_{t,j}\right)} \right)_{1\leqslant i\leqslant d},$$

*then for all $x \in \Delta_d$,*

$$\langle u_t, x - x_t \rangle = \mathrm{KL}(x, x_t) - \mathrm{KL}(x, x_{t+t}) + \log\left( \sum_{i=1}^d x_{t,i} \exp\left(u_{t,i}\right) - \langle u_t, x_t \rangle \right),$$

*where $\mathrm{KL}(x', x) = \sum_{i=1}^d x'_i \log(x'_i / x_i)$ denotes the Kullback–Leibler divergence.*

## 2.2 Regularizers

**Definition 2.2.1.** A function $h : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is an *pre-regularizer* if it is proper, strictly convex, and lower semicontinuous. Moreover, if $\operatorname{dom} h^* = \mathbb{R}^d$, then $h$ is said to be an *regularizer*.

*Remark* 2.2.2. A regularizer begin proper, convex, and lower semicontinuous, Proposition 1.4.7 applies.

The following proposition gives several sufficient conditions for the condition $\operatorname{dom} h^* = \mathbb{R}^d$ to be satisfied.

**Proposition 2.2.3.** *Let $h$ be an pre-regularizer.*

*(i) If $\operatorname{dom} h$ is bounded, then $h$ is a regularizer.*

*(ii) If $h$ is differentiable on $\mathcal{D}_h := \operatorname{int} \operatorname{dom} h$ and $\nabla h(\mathcal{D}_h) = \mathbb{R}^d$, then $h$ is a regularizer.*

*(iii) If $h$ is strongly convex, then $h$ is a regularizer.*

*Proof.* Let $y \in \mathbb{R}^d$. For each of the three assumptions, let us prove that $h^*(y)$ is finite. This will prove that $\operatorname{dom} h^* = \mathbb{R}^d$.

(i) By definition of a pre-regularizer, we have:

$$h^*(y) = \sup_{x \in \mathbb{R}^d} \{\langle y, x\rangle - h(x)\} = \sup_{x \in \operatorname{cl} \operatorname{dom} h} \{\langle y, x\rangle - h(x)\}.$$

Besides, the function $x \mapsto \langle y, x\rangle - h(x)$ is upper semicontinuous and therefore, according to Proposition 1.2.16, attains a maximum on compact set $\operatorname{cl} \operatorname{dom} h$ because $\operatorname{dom} h$ is assumed to be bounded. Therefore $h^*(y) < +\infty$.

(ii) Because $\nabla h(\mathcal{D}_h) = \mathbb{R}^d$ by assumption, there exists $x \in \mathcal{D}_h$ such that $\nabla h(x) = y$. Then, by Proposition 1.4.7, $h^*(y) = \langle y, x\rangle - h(x) < +\infty$.

(iii) The function $x \mapsto \langle y, x\rangle - h(x)$ is the opposite of a strongly convex lower semicontinous function on $\mathbb{R}^d$ and therefore admits a maximum by Proposition 1.6.10. Therefore, $h^*(y) < +\infty$.

□

**Proposition 2.2.4.** *A regularizer admits a minimum.*

*Proof.* Let $h : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a regularizer. By definition of a regularizer, $\operatorname{dom} h^* = \mathbb{R}^d$. In particular,

$$+\infty > h^*(0) = \sup_{x \in \mathbb{R}^d} \{-h(x)\},$$

in other words $\inf_{x\in\mathbb{R}^d} h(x) > -\infty$. Besides, $h^*$ is real-valued, convex and therefore continuous on $\mathbb{R}^d$, which implies that $h^*$ is bounded on the closed Euclidean unit ball. Let $M > 0$ such that $h^*(y) \leqslant M$ for all $y$ in the closed Euclidean unit ball. Let $y \in \mathbb{R}^d$ such that $\|y\|_2 \leqslant 1$. Let $x \in \mathbb{R}^d$. It holds that

$$\langle y, x\rangle - h(x) \leqslant h^*(y) \leqslant M,$$

and thus

$$\langle y, x\rangle - M \leqslant h(x).$$

In particular, for $y = x/\|x\|_2$,

$$\|x\| - M \leqslant h(x).$$

Let $\mathcal{B}$ be the closed centered ball of radius $M + \inf h + 1$. For $x \notin \mathcal{B}$, $h(x) \geqslant \inf h + 1$. The set $\mathcal{B}$ being compact, $h$ admits a minimizer on $\mathcal{B}$ and therefore on $\mathbb{R}^d$. □

**Proposition 2.2.5** (Differentiability of $h^*$)**.** *Let $h$ be a regularizer. Then, $h^*$ is differentiable on $\mathbb{R}^d$, $\nabla h^*$ takes values in* $\operatorname{dom} h$*, and*

$$\nabla h^*(y) = \operatorname*{arg\,max}_{x\in\mathbb{R}^d} \left\{\langle y, x\rangle - h(x)\right\}.$$

*Proof.* Let $y \in \mathbb{R}^d$. Because $\operatorname{dom} h^* = \mathbb{R}^d$, the subdifferential $\partial h^*(y)$ is nonempty by Proposition 1.3.6, $\partial h^*(y)$ is the set of maximizers of function $x \mapsto \langle y, x\rangle - h(x)$, which is strictly concave. Therefore, the maximizer belongs to $\operatorname{dom} h$, is unique, and thus $h^*$ is differentiable at $y$ by Proposition 1.3.4. □

**Proposition 2.2.6** (Euclidean regularizer)**.** *The* Euclidean regularizer *on $\mathcal{X}$, defined as*

$$h_2(x) = \frac{1}{2}\|x\|_2^2 + I_{\mathcal{X}}(x), \quad x \in \mathbb{R}^d,$$

*is a regularizer, satisfies* $\operatorname{dom} h_2 = \mathcal{X}$ *and $\nabla h_2^*$ is the Euclidean projection onto $\mathcal{X}$, in other words:*

$$\nabla h_2^*(y) = \operatorname*{arg\,min}_{x\in\mathcal{X}} \|y - x\|, \quad y \in \mathbb{R}^\cdot$$

*In particular, in the unconstrained case $\mathcal{X} = \mathbb{R}^d$, $\nabla h_2^*(y) = y$ for all $y \in \mathbb{R}^d$.*

*Proof.* □

**Proposition 2.2.7** (Mahalanobis regularizer)**.** *Let $A \in \mathcal{S}_d^{++}(\mathbb{R})$. The* Mahalanobis regularizer *on $\mathcal{X}$ associated with $A$, defined as*

$$h_A(x) = \frac{1}{2}\|x\|_A^2 + I_{\mathcal{X}}(x), \quad x \in \mathbb{R}^d,$$

*is a regularizer, satisfies* $\operatorname{dom} h_A = \mathcal{X}$ *and*

$$\nabla h_A^*(y) = \operatorname*{arg\,min}_{x \in \mathcal{X}} \left\| A^{-1}y - x \right\|_A .$$

*In particular, in the unconstrained case* $\mathcal{X} = \mathbb{R}^d$, $\nabla h_A^*(y) = A^{-1}y$ *for all* $y \in \mathbb{R}^d$.

**Proposition 2.2.8** (Entropic regularizer)**.** *The* entropic regularizer

$$h_{\mathrm{ent}}(x) = \begin{cases} \sum_{i=1}^d x_i \log x_i & \text{if } x \in \Delta_d \\ +\infty & \text{otherwise,} \end{cases}$$

*is a regularizer, satisfies* $\operatorname{dom} h_{\mathrm{ent}} = \Delta_d$, *and*

$$\nabla h_{\mathrm{ent}}^*(y) = \left( \frac{\exp(y_i)}{\sum_{j=1}^d \exp(y_j)} \right)_{1 \leqslant i \leqslant d}, \quad y \in \mathbb{R}^d.$$

*Proof.* □

## 2.3 UMD iterates

**Definition 2.3.1.** Let $h$ be a regularizer and $(u_t)_{t \geqslant 1}$ a sequence in $\mathbb{R}^d$. A sequence $((x_t, y_t))_{t \geqslant 1}$ in $\mathbb{R}^d \times \mathbb{R}^d$ is a sequence of *UMD iterates* associated with regularizer $h$ and *dual increments* $(u_t)_{t \geqslant 1}$ if for all $t \geqslant 1$,

(i) $y_t \in \partial h(x_t)$,

(ii) $x_{t+1} = \nabla h^*(y_t + u_t)$.

*Remark* 2.3.2. By Proposition 1.4.7, property (ii) is equivalent to $y_t + u_t \in \partial h(x_{t+1})$.

*Remark* 2.3.3. For each $t \geqslant 1$, $x_t \in \operatorname{dom} h$, because otherwise $\partial h(x_t)$ would be empty by Proposition 1.3.6 and could not contain $y_t$.

**Definition 2.3.4.** Let $h$ be a regularizer and $(u_t)_{t \geqslant 1}$ a sequence in $\mathbb{R}^d$. A sequence $((x_t, y_t))_{t \geqslant 1}$ in $\mathbb{R}^d \times \mathbb{R}^d$ is a sequence of *strict UMD iterates* associated with $h$ and $(u_t)_{t \geqslant 1}$ if for all $t \geqslant 1$,

(I) $y_t \in \partial h(x_t)$,

(II) $\forall x \in \operatorname{dom} h,\ \langle y_t + u_t - y_{t+1} | x - x_{t+1} \rangle \leqslant 0$.

**Proposition 2.3.5.** *Let* $((x_t, y_t))_{t \geqslant 1}$ *be a sequence of strict UMD iterates defined as above. Then for all* $t \geqslant 1$, $x_{t+1} = \nabla h^*(y_t + u_t)$ *and thus* $((x_t, y_t))_{t \geqslant 1}$ *are UMD iterates.*

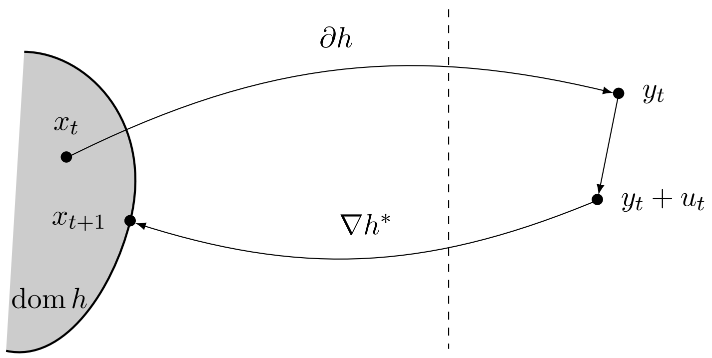


Figure 2.1: UMD iterates

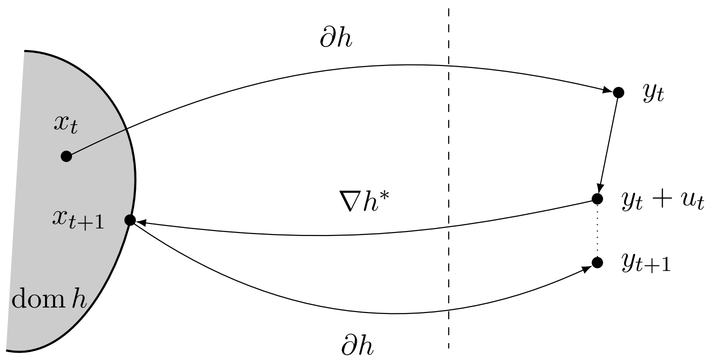


Figure 2.2: Strict UMD iterates. The dotted line represents the variational condition (II) that relates $y_t + u_t$ and $y_{t+1}$.

*Proof.* Let $t \geqslant 1$ and let us equivalently prove that $y_t + u_t \in \partial h(x_{t+1})$. Let $x \in \mathbb{R}^d$. If $x \notin \operatorname{dom} h$,

$$+\infty = h(x) - h(x_{t+1}) \geqslant \langle y_t + u_t, x - x_{t+1} \rangle .$$

If $x \in \operatorname{dom} h$, using assumption (II) and the fact that $y_{t+1} \in h(x_{t+1})$,

$$h(x) - h(x_{t+1}) \geqslant \langle y_{t+1}, x - x_{t+1} \rangle \geqslant \langle y_t + u_t, x - x_{t+1} \rangle .$$

Therefore, $y_t + u_t \in \partial h(x_{t+1})$, in other words $x_{t+1} = \nabla h^*(y_t + u_t)$. $\square$

**Example 2.3.6** (Euclidean regularizer)**.** Denote $\Pi_{\mathcal{X}}$ the Euclidean projection onto $\mathcal{X}$ and consider the Euclidean regularizer on $\mathcal{X}$: $h_2 = \frac{1}{2}\|\cdot\|_2^2 + I_{\mathcal{X}}$ and $x_1 \in \mathcal{X}$. Let $(u_t)_{t\geqslant 1}$ be a sequence in $\mathbb{R}^d$.

- If $x_{t+1} = \Pi_{\mathcal{X}}(x_t + u_t)$ for all $t \geqslant 1$, then $((x_t, x_t))_{t\geqslant 1}$ can be proved to be a sequence of strict UMD iterates associated with $h_2$ and $(u_t)_{t\geqslant 1}$.
- If $y_{t+1} = y_t + u_t$ and $x_{t+1} = \Pi_{\mathcal{X}}(y_{t+1})$ for all $t \geqslant 1$, then $((x_t, y_t))_{t\geqslant 1}$ can also be proved to be a sequence of strict UMD iterates associated with $h_2$ and $(u_t)_{t\geqslant 1}$.

*Remark* 2.3.7 (Non-unicity of strict UMD iterates). As already seen in the above example, an interesting character of (strict) UMD iterates is that for a given sequence $(u_t)_{t\geqslant 1}$ of dual increments and initial points $(x_1, y_1)$ such that $y_1 \in \partial h(x_1)$, there may be several possible strict UMD iterates. Here is a simple and explicit example. Consider $d = 1$, $\mathcal{X} = [0,1]$, $h(x) = \frac{1}{2}x^2 + I_{\mathcal{X}}(x)$, $(x_1, y_1) = (1,1)$ and $u_t = (-1)^t$ for $t \geqslant 1$. Then, one can verify that $((1, \frac{3+(-1)^t}{2}))_{t\geqslant 1}$ is a strict UMD sequence, and so is $((x_t, y_t))_{t\geqslant 1}$ where $x_t = y_t = \frac{1+(-1)^{t+1}}{2}$ for $t \geqslant 1$.

*Remark* 2.3.8 (Existence of strict UMD iterates). As soon as regularizer $h$ and sequence of dual increments $(u_t)_{t\geqslant 1}$ are given, we can see that associated strict UMD iterates always exist. Indeed, from the definition of a regularizer, $\operatorname{dom} h^* = \mathbb{R}^d$ and thus one can choose any $y_1 \in \mathbb{R}^d$ and consider $x_1 := \nabla h^*(y_1)$, which satisfies $y_1 \in \partial h(x_1)$. Then, for $t \geqslant 1$, one can consider $y_{t+1} := y_t + u_t$ which indeed satisfies variational condition (II), and then define $x_{t+1} := \nabla h^*(y_{t+1})$, which ensures $y_{t+1} = \partial h(x_{t+1})$, as required by (i).

*Remark* 2.3.9 (Alternative notation for strict UMD iterates). For a given regularizer $h$, let $\Pi_h : \mathbb{R}^d \rightrightarrows \mathcal{X} \times \mathbb{R}^d$ be a set-valued mapping defined as follows. For $y_1 \in \mathbb{R}^d$, $\Pi_h(y_1)$ is the set of couples $(x, y)$ satisfying

$$x = \nabla h^*(y_1), \quad y \in \partial h(x), \quad \text{and} \quad \forall x' \in \operatorname{dom} h, \quad \left\langle y_1 - y, x' - x \right\rangle \leqslant 0.$$

Then, one can verify that $((x_t, y_t))_{t\geqslant 1}$ is a strict UMD sequence associated with $h$ and given sequence $(u_t)_{t\geqslant 1}$ if, and only if, $y_1 \in \partial h(x_1)$ and

$$(x_{t+1}, y_{t+1}) \in \Pi_h(y_t + u_t), \quad t \geqslant 1.$$

## 2.4 Regret bounds

**Lemma 2.4.1** (UMD lemma)**.** *Let $h$ be a regularizer, $(u_t)_{t\geqslant 1}$ a sequence in $\mathbb{R}^d$, and $((x_t, y_t))_{t\geqslant 1}$ a sequence of UMD iterates associated with regularizer $h$ and* dual increments *$(u_t)_{t\geqslant 1}$ and $x \in \operatorname{dom} h$. Consider notation*

$$D_t = D_h(x, x_t;\ y_t), \quad D_t' = D_h(x_{t+1}, x_t;\ y_t), \quad D_t^* = D_{h^*}(y_t{+}u_t, y_t), \quad t \geqslant 1.$$

*(i) Then for all $t \geqslant 1$,*

$$\langle u_t, x - x_{t+1} \rangle = D_t - D_{t+1} - D_t' + \langle y_t + u_t - y_{t+1}, x - x_{t+1} \rangle\,,$$

*and*

$$\langle u_t, x - x_t \rangle = D_t - D_{t+1} + D_t^* + \langle y_t + u_t - y_{t+1}, x - x_{t+1} \rangle\,.$$

*(ii) Moreover, $((x_t, y_t))_{t\geqslant 1}$ are strict UMD iterates, then for all $t \geqslant 1$,*

$$\langle u_t, x - x_{t+1} \rangle \leqslant D_t - D_{t+1} - D_t',$$

*and*

$$\langle u_t, x - x_t \rangle \leqslant D_t - D_{t+1} + D_t^*.$$

*(iii) Besides, if $h$ is $K$-strongly convex with respect to some norm $\|\cdot\|$ and some $K > 0$, then for all $t \geqslant 1$,*

$$D_t^* \leqslant \frac{1}{2K} \|u_t\|_*^2 .$$

*Proof.* The first identity from (i) can be verified by writing explicitly the difference between both sides and simplifying. The second identity follows from noticing that for all $t \geqslant 1$,

$$\langle u_t, x_{t+1} - x_t \rangle = D_t' + D_h(x_t, x_{t+1};\ y_t + u_t) = D_t' + D_t^*,$$

where the second equality comes from Proposition 1.5.5; and adding to the first equality. Equalities in (ii) are an immediate consequence of (i). (iii) follows from Proposition 1.6.4. □

## 2.5 Time-dependent regularizers

**Definition 2.5.1.** Let $(h_t)_{t \in 1 + \frac{1}{2}\mathbb{N}}$ be a sequence of regularizers and $(u_t)_{t \geqslant 1}$ a sequence in $\mathbb{R}^d$. An associated sequence $((x_t, y_t))_{t \geqslant 1}$ in $\mathbb{R}^d \times \mathbb{R}^d$ of *UMD iterates* satisfy for all $t \geqslant 1$,

(i) $y_t \in \partial h_t(x_t)$,

(ii) $x_{t+1} = \nabla h_{t+1/2}^*(y_t + u_t)$.

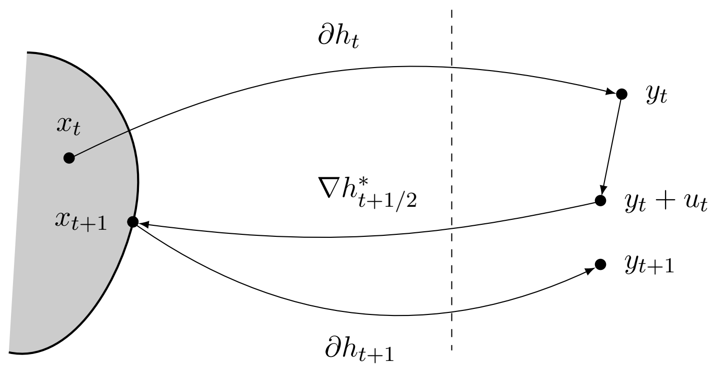


Figure 2.3: UMD iterates with time-dependent regularizers.

**Lemma 2.5.2** (UMD lemma with time-dependent regularizers)**.** *Let $(h_t)_{t \in 1 + \frac{1}{2}\mathbb{N}}$ be a sequence of regularizers, $(u_t)_{t \geqslant 1}$ a sequence in $\mathbb{R}^d$, $((x_t, y_t))_{t \geqslant 1}$ associated UMD iterates, and $x \in \bigcap_{t \in 1 + \frac{1}{2}\mathbb{N}} \operatorname{dom} h_t$. For each $t \geqslant 1$, consider notation*

- $D_t = D_{h_t}(x, x_t;\ y_t)$,
- $D'_t = h_{t+1/2}(x_{t+1}) - h_t(x_t) - \langle y_t, x_{t+1} - x_t \rangle$,
- $D^*_t = h^*_{t+1/2}(y_t + u_t) - h^*_t(y_t) - \langle u_t, x_t \rangle$,
- *For* $x' \in \operatorname{dom} h_t$ *(resp.* $\operatorname{dom} h_{t+1/2}$*),* $\Delta h_t(x') = h_{t+1/2}(x') - h_t(x')$ *(resp.* $\Delta h_{t+1/2}(x') = h_{t+1}(x') - h_{t+1/2}(x')$*),*
- $D^{\Delta}_{t+1/2} = \Delta h_{t+1/2}(x) - \Delta h_{t+1/2}(x_{t+1}) - \langle y_{t+1} - y_t - u_t, x - x_{t+1} \rangle$.

*(i) Then for all* $t \geqslant 1$*,*

$$\langle u_t, x - x_{t+1} \rangle = D_t - D_{t+1} - D'_t + D^{\Delta}_{t+1/2} + \Delta h_t(x),$$

*and*

$$\langle u_t, x - x_t \rangle = D_t - D_{t+1} + D^*_t + D^{\Delta}_{t+1/2} + \Delta h_t(x).$$

*(ii) If* $\Delta h_t = 0$ *for a given* $t \geqslant 1$*, then*

$$\langle u_t, x - x_{t+1} \rangle = D_t - D_{t+1} - D'_t + D^{\Delta}_{t+1/2},$$

*and*

$$\langle u_t, x - x_t \rangle = D_t - D_{t+1} + D^*_t + D^{\Delta}_{t+1/2}.$$

*(iii) If* $\Delta h_{t+1/2} = 0$ *and* $\langle y_{t+1} - y_t - u_t, x - x_{t+1} \rangle \geqslant 0$ *for a given* $t \geqslant 1$*, then*

$$\langle u_t, x - x_{t+1} \rangle \leqslant D_t - D_{t+1} - D'_t + h_{t+1}(x) - h_t(x),$$

*and*

$$\langle u_t, x - x_t \rangle \leqslant D_t - D_{t+1} + D^*_t + h_{t+1}(x) - h_t(x).$$

*(iv) If for a given* $t \geqslant 1$*,* $h_{t+1/2} \geqslant h_t$ *and* $h_t$ *is* $K_t$*-strongly convex with respect to some norm* $\|\cdot\|$ *and some* $K_t > 0$*, then*

$$D^*_t \leqslant \frac{1}{2K_t} \|u_t\|^2_* .$$

*Proof.* The first equality in (i) can be proved by merely simplifying. For the second inequality, we notice that

$$\begin{aligned}
\langle u_t, x_t - x_{t+1} \rangle &= D'_t + h_t(x_t) - h_{t+1/2}(x_{t+1}) - \langle y_t + u_t, x_t - x_{t+1} \rangle \\
&= D'_t + \langle y_t, x_t \rangle - h^*_t(y_t) - \langle y_t + u_t, x_{t+1} \rangle + h^*_{t+1/2}(y_t + u_t) \\
&\qquad - \langle y_t + u_t, x_t - x_{t+1} \rangle \\
&= D'_t + D^*_t,
\end{aligned}$$

where for the second equality, we used Fenchel's identity from Proposition 1.4.7. Adding the above to the first equality give the second equality in (i). Then, (ii) and (iii) are easy consequences.

Let us prove (iv). The assumption $h_{t+1/2} \geqslant h_t$ and implies $h_{t+1/2}^* \leqslant h_t^*$ (see Remark 1.4.4). Therefore,

$$D_t^* = h_{t+1/2}^*(y_t + u_t) - h_t^*(y_t) - \langle u_t, x_t \rangle \leqslant D_{h_t^*}(y_t + u_t, y_t) \leqslant \frac{1}{2K_t} \left\| u_t \right\|_*^2 .$$

$\square$

# Chapter 3

# Online linear optimization

This chapter and the next one first presents the topic of regret minimization in sequential decision problems and then makes the connection with online learning and optimization through the frameworks of online linear optimization and online convex optimization. Several important families of algorithms (dual averaging, mirror descent, follow the regularized leader) are introduced and analyzed using UMD theory from Chapter 2.

Throughout the chapter, $\mathcal{X}$ is a nonempty closed convex set, and $\Pi_{\mathcal{X}}$ denotes the Euclidean projection onto $\mathcal{X}$. For a function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, we denote $\mathcal{D}_f = \operatorname{int} \operatorname{dom} f$.

## 3.1 Introduction to regret minimization

Let us first consider a simple sequential decision problem where the Decision Maker chooses its actions in the finite set $\{1, \dots, d\}$, possibly at random: at step $t \geqslant 1$,

- the Decision Maker chooses $x_t \in \Delta_d$,
- Nature chooses and reveals *payoff vector* $u_t \in [0,1]^d$,
- $i_t$ is a random element in $\{1, \dots, d\}$ drawn according to distribution $x_t$ and revealed to the Decision Maker,
- the Decision Maker obtains payoff $u_{t,i_t}$.

The choice of $x_t$ by the Decision Maker may depend on all past information known to him, meaning $(x_1, u_1, i_1, \dots, x_{t-1}, u_{t-1}, i_{t-1})$. Similarly, the choice of $u_t$ by Nature may depend on all past information including $x_t$: $(x_1, u_1, i_1, \dots, x_{t-1}, u_{t-1}, i_{t-1}, x_t)$.

In a restrictive variant of this problem, called *multi-armed bandit* and which will be considered later, the decision maker only observes the actual payoff $u_{t,i_t}$, and not the whole payoff vector $u_t$.

The Decision Maker wishes to maximize its cumulative payoff $\sum_{t=1}^{T} u_{t,i_t}$. More specifically, we aim at constructing decision rules for the Decision Maker (which we will simply call *algorithms*) that offer some *worst-case guarantee* on the cumulative payoff that holds *for all possible sequence* $(u_t)_{t\geqslant 1}$ chosen by Nature. Therefore, the guarantee must be relative to the sequence of payoff vectors $(u_t)_{t\geqslant 1}$. One possible type of guarantee is an upper bound on the *regret*, defined as

$$\max_{1\leqslant i\leqslant d} \sum_{t=1}^{T} u_{t,i} - \sum_{t=1}^{T} u_{t,i_t},$$

which compares the actual cumulative payoff with the best cumulative payoff that would have been obtained[1] by constantly choosing a given element $i \in \{1,\dots,d\}$, meaning $i_t = i$ for all $t \geqslant 1$.

For a given sequence of payoff vectors $(u_t)_{t\geqslant 1}$, the above regret is a random variable (because for each $t \geqslant 1$, $i_t$ is a random variable), and we are interested in analyzing its expectation. Using the law of total expectation, we can write

$$\begin{aligned}\mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=1}^{T} u_{t,i} - \sum_{t=1}^{T} u_{t,i_t}\right] &= \mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=1}^{T} u_{t,i} - \sum_{t=1}^{T} \mathbb{E}\left[u_{t,i_t} | x_t\right]\right] \\ &= \mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=1}^{T} u_{t,i} - \sum_{t=1}^{T} \langle u_t, x_t\rangle\right].\end{aligned} \tag{3.1}$$

Besides, because the maximum of linear function on a convex compact set is attained at its edges,

$$\max_{1\leqslant i\leqslant d} \sum_{t=1}^{T} u_{t,i} = \max_{1\leqslant i\leqslant d} \left\langle \sum_{t=1}^{T} u_t, e_i \right\rangle = \max_{x\in\Delta_d} \sum_{t=1}^{T} \langle u_t, x_t\rangle .$$

Therefore, upper bounds on the quantity

$$\max_{x\in\Delta_d} \sum_{t=1}^{T} \langle u_t, x\rangle - \sum_{t=1}^{T} \langle u_t, x_t\rangle = \max_{x\in\Delta_d} \sum_{t=1}^{T} \langle u_t, x - x_t\rangle , \tag{3.2}$$

which we will also call the *regret*, will yield the same bound on the expected regret from (3.1).

A common assumption is that the sequence of payoff vectors $(u_t)_{t\geqslant 1}$ is bounded. In that case, we will see that there exists algorithms that

[1] This interpretation holds only if we assume that the choice of the payoff vectors $(u_t)_{t\geqslant 1}$ by Nature does not depend on the choices of the Decision Maker. Nevertheless, all guarantees on the regret will still hold, even in the case where Nature does react to the choices of the Decision Maker; only the interpretation may not stand.

guarantee that above (cumulative) regret grows at most in $\sqrt{T}$, in other words, the *average* regret is bounded from above by a quantity that vanish as $1/\sqrt{T}$. This has the following interpretation called *prediction with expect advice*: there are $d$ experts, and at each step $t \geqslant 1$, the Decision Maker has to choose one of the experts and follow his advice, and obtains the corresponding payoff. Then, as will be proved below, there exists algorithms such that the Decision Maker is guaranteed to perform *as well as the best expert* (asymptotically and in average).

Another important question will be the optimal dependence of the regret bound in $d$, in the case e.g. where the payoff vectors are assumed to be in $[0,1]^d$.

**Example 3.1.1** (Follow the leader)**.** The simplest algorithm we may think of is called *follow the leader*, and picks the decision $i \in \{1,\dots,d\}$ which would have yielded the highest cumulative payoff on the previous steps:

$$i_{t+1} = \operatorname*{arg\,max}_{1\leqslant i\leqslant d} \sum_{s=1}^{t} u_{s,i}, \quad t \geqslant 1,$$

which can be equivalently written

$$x_{t+1} = \operatorname*{arg\,max}_{x\in\Delta_d} \left\langle \sum_{s=1}^{t} u_s, x \right\rangle, \quad t \geqslant 1. \tag{3.3}$$

Unfortunately, this algorithm is too simple and it is easy to find a bounded sequence for which the regret grows linearly, e.g. for $d=2$,

$$u_1 = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_4 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \dots$$

yields for all $T \geqslant 1$,

$$\max_{1\leqslant i\leqslant d} \sum_{t=1}^{T} u_{t,i} - \sum_{t=1}^{T} u_{t,i_t} \geqslant \frac{T-1}{2}.$$

Intuitively, the issue with the follow the leader algorithm is that it follows previous data too closely, so that there exists a payoff vector for the next step for which the decision of the algorithm is the worst. This is a kind of *overfitting*. To address the issue, one possible approach is to *regularize* the quantity that is maximized in (4.3), which will lead to the *dual averaging* and *follow the regularized leader* algorithms below.

We now make the connection with online learning and optimization by presenting two successive extensions of the above framework.

**Online linear optimization** The quantity (3.2) inspires the following natural extension. Let $\mathcal{X} \subset \mathbb{R}^d$ be a nonempty closed convex set, and $\mathcal{U} \subset \mathbb{R}^d$ any nonempty set. At each step $t \geqslant 1$,

- the Decision Maker chooses $x_t \in \mathcal{X}$,
- Nature chooses and reveals $u_t \in \mathcal{U}$,
- the Decision Maker obtains payoff $\langle u_t, x_t \rangle$.

In the case where $\mathcal{X}$ is bounded, the natural definition of the regret is

$$\max_{x \in \mathcal{X}} \sum_{t=1}^{T} \langle u_t, x - x_t \rangle .$$

When $\mathcal{X}$ is unbounded, it will be possible to guarantee upper bounds on

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle ,$$

that depend on the *comparison point* $x \in \mathcal{X}$. In the case where the set $\mathcal{U}$ of payoff vectors is bounded, typical regret bounds also grow as $\sqrt{T}$, as will be established below in Sections 3.2 and 3.3.

**Online convex optimization** We now further generalize by considering convex loss functions, which is motivated by e.g. online learning. Let $\mathcal{X} \subset \mathbb{R}^d$ be a nonempty closed convex set and $\mathcal{L}$ a nonempty set of convex function $\ell : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ such that $\mathcal{X} \subset \operatorname{dom} \ell$. At each step $t \geqslant 1$,

- the Decision Maker chooses $x_t \in \mathcal{X}$,
- Nature chooses and reveals $\ell_t \in \mathcal{L}$,
- the Decision Maker incurs loss $\ell_t(x_t)$.

The corresponding definition for the regret is

$$\max_{x \in \mathcal{X}} \sum_{t=1}^{T} \left( \ell_t(x_t) - \ell_t(x) \right) .$$

In the case where the loss functions have some *curvature*, the latter can be leveraged to achieve regret bounds that grow strictly slower than $\sqrt{T}$.

The case where the loss function is constant boils down to convex optimization.

**Example 3.1.2** (Online linear regression)**.** By slightly modifying the above problem, the Decision Maker can observe some contextual information $z_t$ before choosing $x_t \in \mathcal{X}$. Online versions of supervised learning problem can then be considered with Nature choosing loss function of the form e.g.

$$\ell(x) = (\langle z, x \rangle - y)^2, \qquad z \in \mathbb{R}^d, y \in \mathbb{R}.$$

## 3.2 Dual averaging

We define and analyze the dual averaging family of algorithm which can be seen as a *regularized* version of the follow the leader algorithm from Example (4.3). It is a special case of UMD iterates where for all $t \geqslant 1$, the regularizers satisfy $h_{t+1/2} = h_{t+1}$ (in other words $\Delta h_{t+1/2} = 0$) and where the next dual point $y_{t+1}$ is uniquely defined as $y_{t+1} = y_t + u_t$.

We obtain in Proposition 3.2.5 below a regret bound in the context of online linear optimization, which will be applied and transposed to numerous problems.

**Definition 3.2.1.** Let $(h_t)_{t\geqslant 1}$ be a sequence of regularizers, $(u_t)_{t\geqslant 1}$ a sequence in $\mathbb{R}^d$, and $y_1 \in \mathbb{R}^d$. The associated sequence $(x_t)_{t\geqslant 1}$ of *dual averaging* (DA) iterates is defined for $t \geqslant 1$ as

$$x_t = \nabla h_t^*(y_t) \quad \text{and} \quad y_{t+1} = y_t + u_t.$$

*Remark* 3.2.2. The above definition can be equivalently written

$$\begin{aligned} x_{t+1} &= \operatorname*{arg\,max}_{x\in\mathbb{R}^d} \left\{ \left\langle y_1 + \sum_{s=1}^{t} u_s, x \right\rangle - h_t(x) \right\} \\ &= \operatorname*{arg\,max}_{x\in\mathbb{R}^d} \left\{ \langle u_t, x\rangle - D_{h_t}(x, x_t;\ y_t) \right\}. \end{aligned}$$

In the case $y_1 = 0$, the above second expression gives the following interpretation: $x_{t+1}$ is the maximizer not of the past cumulative payoff function $x \mapsto \left\langle \sum_{s=1}^{t} u_s, x \right\rangle$ as in the *follow the leader* algorithm (see Section 3.1), but of a *regularized* version. For this reason, this algorithm is sometimes called *follow the regularized leader*. We will use this name below to designate a more general algorithm in the context of regret minimization with convex losses.

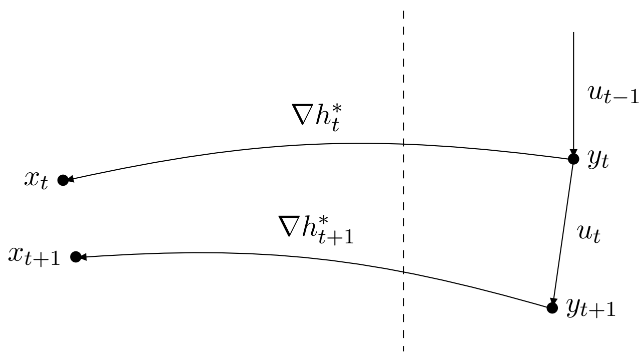


Figure 3.1: Dual averaging

**Example 3.2.3** (Lazy Gradient Descent)**.** Let $y_1 \in \mathbb{R}^d$ and denote $\Pi_{\mathcal{X}}$ the Euclidean projection onto $\mathcal{X}$. For a given sequence $(u_t)_{t \geqslant 1}$, $x_1 = \Pi_{\mathcal{X}}(y_1)$ and

$$y_{t+1} = y_t + u_t \quad \text{and} \quad x_{t+1} = \Pi_{\mathcal{X}}(y_{t+1}), \quad t \geqslant 1$$

corresponds to dual averaging iterates with constant Euclidean regularizer $h = \frac{1}{2}\|\cdot\|_2^2 + I_{\mathcal{X}}$ (see Proposition 2.2.6 and Example 2.3.6).

**Example 3.2.4** (Exponential weights algorithm)**.** For a given sequence $(u_t)_{t \geqslant 1}$, consider

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

These correspond to dual averaging iterates with constant entropic regularizer $h_{\mathrm{ent}}$ (see Proposition 2.2.8)

The following statement gives a regret bound for online linear optimization, which will also be applied and transposed to numerous problems.

**Proposition 3.2.5** (Regret bounds for DA)**.** *Let $(x_t)_{t \geqslant 1}$ and $(y_t)_{t \geqslant 1}$ be defined as in Definition 3.2.1, and $x \in \bigcap_{t \geqslant 1} \operatorname{dom} h_t$. Then,*

*(i) $((x_t, y_t))_{t \geqslant 1}$ is a sequence of UMD iterates associated with regularizers $(h_t)_{t \in 1 + \frac{1}{2}\mathbb{N}}$ and dual increments $(u_t)_{t \geqslant 1}$, where $h_{t+1/2} := h_{t+1}$ for all $t \geqslant 1$;*

*(ii) for all $T \geqslant 1$,*

$$\sum_{t=1}^T \langle u_t, x - x_t \rangle \leqslant D_1 + (h_{T+1}(x) - h_1(x)) + \sum_{t=1}^T D_t^*,$$

*where*

$$\begin{aligned} D_1 &= h_1(x) - h_1(x_1) - \langle y_1, x - x_1 \rangle \\ D_t^* &= h_{t+1}^*(y_t + u_t) - h_t^*(y_t) - \langle u_t, x_t \rangle, \quad t \geqslant 1; \end{aligned}$$

*(iii) if for $t \geqslant 1$, $h_{t+1} \geqslant h_t$ and $h_t$ is $K_t$-strongly convex for some norm $\|\cdot\|$, then $D_t^*$ is bounded as*

$$D_t^* \leqslant \frac{1}{2K_t} \|u_t\|_*^2 .$$

*Proof.* (i) holds because the conditions from Definition 2.3.1 are trivially satisfied. (ii) and (iii) easily follow from Lemma 2.5.2. $\square$

**Dual averaging with nonincreasing parameter** Let $h$ be a regularizer, $(u_t)_{t\geqslant 1}$ a sequence in $\mathbb{R}^d$ , $y_1 = 0$ and $(\eta_t)_{t\geqslant 1}$ a positive and nonincreasing sequence. Consider iterates $(x_t)_{t\geqslant 1}$ defined for $t \geqslant 1$ as

$$x_t = \nabla h^*(\eta_t y_t) \quad \text{and} \quad y_{t+1} = y_t + u_t. \tag{3.4}$$

Then, $((x_t, y_t))_{t\geqslant 1}$ are UMD iterates associated with dual increments $(u_t)_{t\geqslant 1}$ and regularizers $(h_t)_{t\in 1+\frac{1}{2}\mathbb{N}}$ where

$$h_t(x) = \frac{h(x) - \min h}{\eta_t} \quad \text{and} \quad t \in \{1, 2, \dots\}, \tag{3.5}$$

and $h_{t+1/2} = h_{t+1}$ for $t \in \{1, 2, \dots\}$. The definition of UMD iterates is invariant when constants are added to regularizers, so it would be equivalent to simply consider $h_t = h/\eta_t$ but the above regularizers have the advantage of ensuring $h_{t+1} \geqslant h_t$ and therefore makes the analysis simpler.

**Proposition 3.2.6** (Regret bounds for DA with time-dependent parameters)**.** *Consider the iterates defined in* (3.4)*.*

*(i) For all $T \geqslant 1$ and $x \in \operatorname{dom} h$,*

$$\sum_{t=1}^{T} \langle u_t, x - x_t\rangle \leqslant \frac{h(x) - \min h}{\eta_T} + \sum_{t=1}^{T} \frac{D_{h^*}(\eta_t y_{t+1}, \eta_t y_t)}{\eta_t}.$$

*(ii) Moreover, if $h$ is $K$-strongly convex for some norm $\|\cdot\|$, for all $T \geqslant 1$,*

$$\sum_{t=1}^{T} \langle u_t, x - x_t\rangle \leqslant \frac{h(x) - \min h}{\eta_T} + \frac{1}{2K}\sum_{t=1}^{T} \eta_t \|u_t\|_*^2 .$$

*(iii) Moreover, if there exists $L > 0$ such that $\|u_t\|_* \leqslant L$ for all $t \geqslant 1$, then the choice $\eta_t = \eta\sqrt{K}/(L\sqrt{t})$ (for all $t \geqslant 1$) for some $\eta > 0$ yields*

$$\sum_{t=1}^{T} \langle u_t, x - x_t\rangle \leqslant \left(\frac{h(x) - \min h}{\eta} + \eta\right) L\sqrt{\frac{T}{K}}.$$

*(iv) Moreover, if $\sup_{x\in\operatorname{dom} h} h(x) < +\infty$, then $\eta = \sqrt{\delta_h}$ yields*

$$\max_{x\in\operatorname{cl}\operatorname{dom} h} \sum_{t=1}^{T} \langle u_t, x - x_t\rangle \leqslant 2L\sqrt{\frac{\delta_h T}{K}}.$$

*where $\delta_h := \sup_{x\in\operatorname{dom} h} h(x) - \min_{x\in\operatorname{dom} h} h(x)$.*

*Proof.* (i) Applying Proposition 3.2.5, using $y_1 = 0$, and simplifying gives

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{h(x) - \min h}{\eta_{T+1}} + \sum_{t=1}^{T} (h^*_{t+1}(y_{t+1} + u_t) - h^*_t(y_t) - \langle u_t, x_t \rangle).$$

For a given $T \geqslant 1$, iterates $x_1, \dots, x_T$ do not depend on $\eta_{T+1}$, therefore, Proposition 3.2.5 can be applied for $\eta_{T+1} = \eta_T$. Besides, because $h_{t+1} \geqslant h_t$ by definition in (3.5), it follows that $h^*_{t+1}(y_{t+1}) \leqslant h^*_t(y_{t+1})$, and it is easy to verify that for all $y \in \mathbb{R}^d$, $h^*_t(y) = (h^*(\eta_t y) + \min h)/\eta_t$, therefore

$$\begin{aligned} h^*_{t+1/2}(y_{t+1}) - h^*_t(y_t) - \langle u_t, x_t \rangle &\leqslant h^*_t(y_{t+1}) - h^*_t(y_t) - \langle u_t, x_t \rangle \\ &= \eta_t^{-1} \left( h^*(\eta_t y_{t+1}) - h^*(\eta_t y_t) - \langle \eta_t u_t, x_t \rangle \right) \\ &= \eta_t^{-1} D_{h^*}(\eta_t y_{t+1}, \eta_t y_t). \end{aligned}$$

Hence the result.

(ii) follows from Proposition 1.6.4, (iii) and (iv) are immediate consequences. ☐

*Remark* 3.2.7. According to the above statements (iii) and (iv) give that the average regret is minimized at speed $1/\sqrt{T}$, for all $T \geqslant 1$, without prior knowledge of the latter: the algorithm and the guarantee are said to be *horizon-free*. In the case $\sup_{\operatorname{dom} h} h < +\infty$, the bound (iv) does not depend on the comparison point $x$. When $\sup_{\operatorname{dom} h} h = +\infty$, the bound (iii) depend on the comparison point $x$, but still guarantees a horizon-free average regret bound of order $1/\sqrt{T}$, which is an important feature of dual averaging.

## 3.3 Online mirror descent

We define and analyze the the online mirror descent family of algorithms, which is an extension of the online projected gradient descent:

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t + u_t) = \arg\max_{x \in \mathcal{X}} \left\{ \langle u_t, x \rangle - \frac{1}{2} \left\| x - x_t \right\|_2^2 \right\},$$

where the above Euclidean distance is replaced by the Bregman divergence associated with a differentiable function $H$. This yields a special case of UMD iterates where for all $t \geqslant 1$, $h_t = h_{t+1/2}$ and where dual point $y_t$ is uniquely defined as $y_t = \nabla H_t(x_t)$.

Proposition 3.3.13 below gives a regret bound in the context of online linear optimization, and will also be applied and transposed to various problems.

**Definition 3.3.1.** Let $H : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$. Denote $\mathcal{D}_H := \operatorname{int} \operatorname{dom} H$. $H$ is a *mirror map* compatible with $\mathcal{X}$ if

(i) $H$ is lower semi-continuous, strictly convex, and differentiable on $\mathcal{D}_H$,

(ii) the gradient of $H$ takes all possible values, i.e. $\nabla H(\mathcal{D}_H) = \mathbb{R}^d$,

(iii) $\mathcal{X} \subset \operatorname{cl} \mathcal{D}_H$,

(iv) $\mathcal{X} \cap \mathcal{D}_H \neq \varnothing$.

**Example 3.3.2** (Squared $\ell^p$ norms)**.** For $1 < p < +\infty$, $x \mapsto \frac{1}{2} \|x\|_p^2$ is a mirror map compatible with all nonempty closed convex sets.

**Example 3.3.3** (Square Mahalanobis norms)**.** For $A \in \mathbb{R}^{d\times d}$ a symmetric positive definite matrix, $x \mapsto \frac{1}{2} \|x\|_A^2$ is a mirror map compatible w ith all nonempty closed convex sets.

**Example 3.3.4** (Generalized negative entropy)**.** The generalized negative entropy on the closed positive orthant, defined as

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

with convention $0 \cdot \log 0 = 0$, is a mirror map compatible with all nonempty closed convex subsets of $\mathbb{R}_+^d$ (e.g. the simplex).

**Example 3.3.5** (Log barrier)**.** The log barrier on the open positive orthant, defined as

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

is a mirror map compatible with all nonempty closed convex subsets of $\mathbb{R}_+^d$ (e.g. the simplex), although its domain is the open positive orthant $(\mathbb{R}_+^*)^d$.

**Proposition 3.3.6.** *Let $H$ be a mirror map compatible with $\mathcal{X}$, $H^*$ its Legendre–Fenchel transform. Then,*

*(i)* $\operatorname{dom} H^* = \mathbb{R}^d$,

*(ii) $H^*$ is differentiable on $\mathbb{R}^d$,*

*(iii) $\nabla H^*(\mathbb{R}^d) = \mathcal{D}_H$,*

*(iv) For $x \in \mathcal{D}_H$ and $y \in \mathbb{R}^d$, $\nabla H^*(\nabla H(x)) = x$ and $\nabla H(\nabla H^*(y)) = y$.*

*Proof.* Let $y_t \in \mathbb{R}^n$. By property (ii) from Definition 3.3.1, there exists $x_1 \in \mathcal{D}_H$ such that $\nabla H(x_1) = y$. Therefore, function $\varphi_y : x \mapsto \langle y|x\rangle - H(x)$ is differentiable at $x_1$ and $\nabla\varphi_y(x_1) = 0$. Moreover, $\varphi_y$ is strictly concave as a consequence of property (i) from Definition 3.3.1. Therefore, $x_1$ is the unique maximizer of $\varphi_y$ and:

$$H^*(y) = \max_{x\in\mathbb{R}^n} \{\langle y|x\rangle - H(x)\} < +\infty,$$

which proves property (i).

Besides, we have

$$x_1 \in \partial H^*(y) \quad \Longleftrightarrow \quad y = \nabla H(x_1) \quad \Longleftrightarrow \quad x_1 \text{ minimizer of } \phi_y, \tag{3.6}$$

where the first equivalence comes from Proposition 1.4.7. Point $x_1$ being the unique maximizer of $\varphi_y$, we have that $\partial H^*(y)$ is a singleton. In other words, $H^*$ is differentiable in $y$ and

$$\nabla H^*(y) = x_1 \in \mathcal{D}_H. \tag{3.7}$$

First, the above (3.7) proves property (ii). Second, this equality combined with the equality from (3.6) gives the second identity from property (iv). Third, this proves that $\nabla H^*(\mathbb{R}^n) \subset \mathcal{D}_H$.

It remains to prove the reverse inclusion to get property (iii). Let $x \in \mathcal{D}_H$. By property (i) from Definition 3.3.1, $H$ is differentiable in $x$. Consider

$$y := \nabla H(x), \tag{3.8}$$

and all the above holds with this special point $y$. In particular, $x_1 = x$ by uniqueness of $x_1$. Therefore (3.7) gives

$$\nabla H^*(y) = x, \tag{3.9}$$

and this proves $\nabla H^*(\mathbb{R}^n) \supset \mathcal{D}_H$ and thus property (iii). Combining (3.8) and (3.9) gives the first identity from property (iv). □

**Proposition 3.3.7** (OMD iteration)**.** *Let $H$ be a mirror map compatible with $\mathcal{X}$, $x \in \mathcal{D}_H$, $u \in \mathbb{R}^d$ and consider*

$$x' = \operatorname*{arg\,max}_{x'' \in \mathcal{X}} \left\{ \left\langle u, x'' \right\rangle - D_H(x'', x) \right\}.$$

*Then, $x' \in \mathcal{X} \cap \mathcal{D}_H$ and can also be written*

$$\begin{aligned} x' &= \operatorname*{arg\,max}_{x'' \in \mathcal{X}} \left\{ \left\langle \nabla H(x) + u, x'' \right\rangle - H(x'') \right\} \\ &= \operatorname*{arg\,min}_{x'' \in \mathcal{X}} D_H(x'', \nabla H^*(\nabla H(x) + u)). \end{aligned}$$

*Proof.* The proof that $x'$ is well-defined and belongs to $\mathcal{D}_H$ is given in [JKM23]. Besides,

$$\begin{aligned} x' &= \operatorname*{arg\,max}_{x'' \in \mathcal{X}} \left\{ \left\langle u, x'' \right\rangle - D_H(x'', x) \right\} \\ &= \operatorname*{arg\,max}_{x'' \in \mathcal{X}} \left\{ \left\langle u, x'' \right\rangle - H(x'') + \left\langle \nabla H(x), x'' - x \right\rangle \right\} \\ &= \operatorname*{arg\,max}_{x'' \in \mathcal{X}} \left\{ \left\langle \nabla H(x) + u, x'' \right\rangle - H(x'') \right\} \\ &= \operatorname*{arg\,max}_{x'' \in \mathcal{X}} \left\{ -H(x'') + H(\nabla H^*(\nabla H(x) + u)) + \left\langle \nabla H(x) + u, x'' - x \right\rangle \right\} \\ &= \operatorname*{arg\,min}_{x'' \in \mathcal{X}} D_H(x'', \nabla H^*(\nabla H(x) + u)). \end{aligned}$$

□

**Definition 3.3.8.** Let $(H_t)_{t\geqslant 1}$ be a sequence of mirror maps compatible with $\mathcal{X}$ that have the same domain and $(u_t)_{t\geqslant 1}$ a sequence in $\mathbb{R}^d$. A sequence $(x_t)_{t\geqslant 1}$ in $\mathbb{R}^d$ is an sequence of *online mirror descent (OMD)* iterates on $\mathcal{X}$ associated with mirror maps $(H_t)_{t\geqslant 1}$ and *dual increments* $(u_t)_{t\geqslant 1}$ if $x_1 \in \mathcal{X} \cap \operatorname{int} \operatorname{dom} H_1$ and

$$x_{t+1} = \arg\max_{x\in\mathcal{X}} \left\{ \langle u_t, x\rangle - D_{H_t}(x, x_t) \right\}, \quad t \geqslant 1.$$

*Remark* 3.3.9. The above iterates are well-defined. Indeed, we first note that $x \in \mathcal{D}_{H_1}$, and then by induction, as soon as $x_t \in \mathcal{D}_{H_t}$, Proposition 3.3.7 ensures that $x_{t+1}$ is well-defined an unique, and that it belongs to $\mathcal{D}_{H_t} = \mathcal{D}_{H_{t+1}}$.

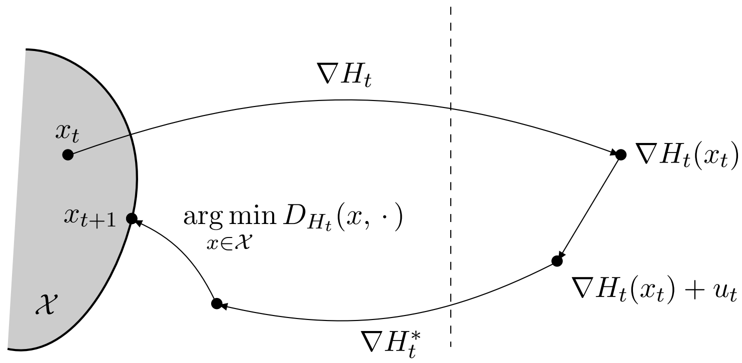


Figure 3.2: Online mirror descent seen with a projection step

**Example 3.3.10** (Online (projected) gradient descent)**.** Denote $\Pi_{\mathcal{X}}$ the Euclidean projection onto $\mathcal{X}$. Let $y_1 \in \mathbb{R}^d$ and a sequence $(u_t)_{t\geqslant 1}$ in $\mathbb{R}^d$. Then, $x_1 = \Pi_{\mathcal{X}}(y_1)$ and

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t + u_t), \quad t \geqslant 1,$$

are OMD iterates on $\mathcal{X}$ associated with constant Euclidean mirror map $H_2 = \frac{1}{2}\left\|\cdot\right\|_2^2$.

**Example 3.3.11** (Exponential weights algorithm)**.** The exponential weights algorithm from Example 3.2.4 gives OMD iterates on the simplex $\Delta_d$ associated with generalized entropic mirror map from Example 3.3.4.

The following proposition indicates the regularizers to consider for making the connection with the definition of UMD iterates.

**Proposition 3.3.12.** *Let $H$ be a mirror map compatible with $\mathcal{X}$. Then $h = H + I_{\mathcal{X}}$ is a regularizer,* $\operatorname{dom} h \subset \mathcal{X}$*, and for all $x \in \mathcal{D}_H$, $\nabla H(x) \in \partial h(x)$.*

*Proof.* See [JKM23]. □

The following statement gives a regret bound for online linear optimization, and will also be applied and transposed to numerous problems.

**Proposition 3.3.13** (Regret bounds for OMD). *Let $(H_t)_{t\geqslant 1}$ be a sequence of mirror maps compatible with $\mathcal{X}$ with a common domain, $(u_t)_{t\geqslant 1}$ a sequence in $\mathbb{R}^d$, $(x_t)_{t\geqslant 1}$ associated OMD iterates in $\mathcal{X}$ and $x \in \mathcal{X} \cap \operatorname{dom} H_1$. Then,*

*(i)* *$((x_t, \nabla H_t(x_t)))_{t\geqslant 1}$ is a sequence of UMD iterates associated with regularizers $(h_t)_{t\in 1+\frac{1}{2}\mathbb{N}}$ defined as*

$$h_t = h_{t+1/2} = H_t + I_{\mathcal{X}}, \quad t \in \{1, 2, \dots\},$$

*and dual increments $(u_t)_{t\geqslant 1}$;*

*(ii)* *for all $T \geqslant 1$,*

$$\sum_{t=1}^{T} \langle u_t, x - x_t\rangle \leqslant D_{H_1}(x, x_1) - D_{H_{T+1}}(x, x_{T+1}) + \sum_{t=1}^{T} (\tilde{D}^{\Delta}_{t+1/2} + \tilde{D}^*_t),$$

*where*

$$\begin{aligned} \tilde{D}^{\Delta}_{t+1/2} &= D_{H_{t+1}-H_t}(x, x_{t+1}), \\ \tilde{D}^*_t &= D_{H_t^*}(\nabla H_t(x_t) + u_t, \nabla H_t(x_t)); \end{aligned}$$

*(iii)* *if for $t \geqslant 1$, $H_t$ is $K_t$-strongly convex for some norm $\|\cdot\|$, then $\tilde{D}^*_t$ is bounded as*

$$\tilde{D}^*_t \leqslant \frac{1}{2K_t} \|u_t\|_*^2 .$$

*Proof.* (i) $(h_t)_{t\in 1+\frac{1}{2}\mathbb{N}}$ are indeed regularizers thanks to Proposition 3.3.12. Let us prove that $((x_t, \nabla H_t(x_t)))_{t\geqslant 1}$ satisfy the definition of UMD iterates. Let $t \geqslant 1$ and $x' \in \mathbb{R}^d$. By definition of $h_t$, $h_t \geqslant H_t$. Moreover, $h_t(x_t) = H_t(x_t)$ because $x_t \in \mathcal{D}_{H_t}$ by Remark 3.3.9. Then,

$$h_t(x') - h_t(x_t) \geqslant H_t(x') - H_t(x_t) \geqslant \langle \nabla F_t(x_t), x - x_t\rangle ,$$

which means $\nabla F_t(x_t) \in \partial h_t(x_t)$, and the first condition from Definition 2.3.1 is satisfied. Besides, using Proposition 3.3.7 and the definition of $h_{t+1/2}$,

$$\begin{aligned} x_{t+1} &= \operatorname*{arg\,max}_{x\in\mathcal{X}} \{\langle \nabla H_t(x_t) + u_t, x\rangle - H_t(x)\} \\ &= \operatorname*{arg\,max}_{x\in\mathbb{R}^d} \left\{\langle \nabla H_t(x_t) + u_t, x\rangle - h_{t+1/2}(x)\right\} = \nabla h^*_{t+1/2}(\nabla H_t(x_t) + u_t), \end{aligned}$$

which proves the second condition, and $((x_t, \nabla H_t(x_t)))_{t\geqslant 1}$ is indeed a sequence of UMD iterates associated with $(H_t)_{t\geqslant 1}$ and $(u_t)_{t\geqslant 1}$.

(ii) For all $t \in 1 + \frac{1}{2}\mathbb{N}$, the definition of $h_t$ imply that $\operatorname{dom} h_t = \mathcal{X} \cap \operatorname{dom} H_t$. Because mirror maps $(H_t)_{t\geqslant 1}$ have a common domain by assumption,

$$x \in \mathcal{X} \cap \operatorname{dom} H_1 = \bigcap_{t\in\{1,2,\dots\}} (\mathcal{X} \cap \operatorname{dom} H_t) = \bigcap_{t\in\{1,2,\dots\}} \operatorname{dom} h_t$$
$$= \bigcap_{t\in 1+\frac{1}{2}\mathbb{N}} \operatorname{dom} h_t.$$

Therefore, Lemma 2.5.2 can be applied with $x$ and with notation therein, we get for all $T \geqslant 1$,

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle = D_1 - D_{T+1} + \sum_{t=1}^{T} (D^{\Delta}_{t+1/2} + D^*_t).$$

First, because $x, x_1 \in \mathcal{X} \cap \operatorname{dom} H_1$,

$$\begin{aligned} D_1 &= h_1(x) - h_1(x_1) - \langle \nabla H_1(x_1), x - x_1 \rangle \\ &= H_1(x) - H_1(x_1) - \langle \nabla H_1(x_1), x - x_1 \rangle \\ &= D_{H_1}(x, x_1). \end{aligned}$$

Similarly, $x$ belongs to $\mathcal{X} \cap \operatorname{dom} H_{T+1}$ and so does $x_{T+1}$ as a consequence of $(x_t)_{t\geqslant 1}$ being OMD iterates (see Remark 3.3.9) and $D_{T+1} = D_{H_{T+1}}(x, x_{T+1})$.

Let us now bound the two remaining Bregman divergences. For $t \geqslant 1$, denote $y_t = \nabla H_t(x_t)$, then Proposition 3.3.7 ensures that

$$x_{t+1} = \operatorname*{arg\,max}_{x\in\mathcal{X}} \left\{ \langle y_t + u, x \rangle - H_t(x) \right\},$$

in other words $x_{t+1}$ is a minimizer of a differentiable convex function over $\mathcal{X}$. Therefore, Proposition 1.2.11 gives the following variational characterization:

$$\left\langle \nabla H_t(x_{t+1}) - y_t - u_t, x' - x_{t+1} \right\rangle \geqslant 0, \quad x' \in \mathcal{X}.$$

We use it for $x' = x$ to bound $D^{\Delta}_{t+1/2}$ as follows.

$$\begin{aligned} D^{\Delta}_{t+1/2} &= \Delta h_{t+1/2}(x) - \Delta h_{t+1/2}(x_{t+1}) \\ &\qquad - \langle y_{t+1} - y_t - u_t, x - x_{t+1} \rangle \\ &\leqslant H_{t+1}(x) - H_t(x) - H_{t+1}(x_{t+1}) + H_t(x_{t+1}) \\ &\qquad - \langle \nabla H_{t+1}(x_{t+1}) - \nabla H_t(x_{t+1}), x - x_{t+1} \rangle \\ &= D_{H_{t+1} - H_t}(x, x_{t+1}) \\ &= \tilde{D}^{\Delta}_{t+1/2}. \end{aligned}$$

We now turn to $D_t^*$. Because $h_t \geqslant H_t$ by definition of $h_t$, the reverse inequality holds for the Legendre–Fenchel transform, and in particular, $h_t^*(y_t+u_t) \leqslant H_t^*(y_t+u_t)$. Besides, it holds that

$$x_t = \nabla H_t^*(y_t) \quad \text{and} \quad x_t = \nabla h_t^*(y_t),$$

where the first equality comes from Proposition 3.3.6, whereas the second is obtained combining the fact that $\nabla H_t(x_t) = y_t \in \partial h_t(x_t)$ (proved above) with Proposition 1.4.7. The latter proposition also gives a characterization of both above gradients as maximizers of concave functions, which yields

$$\begin{aligned} h_t^*(y_t) &= \max_{x\in\mathbb{R}^d} \left\{ \langle y_t, x\rangle - h_t(x) \right\} = \langle y_t, x_t\rangle - h_t(x_t) \\ &= \langle y_t, x_t\rangle - H_t(x_t) = \max_{x\in\mathbb{R}^d} \left\{ \langle y_t, x\rangle - H_t(x) \right\} \\ &= H_t^*(y_t). \end{aligned}$$

Hence

$$\begin{aligned} D_t^* &= h_{t+1/2}^*(y_t+u_t) - h_t^*(y_t) - \langle u_t, x_t\rangle \leqslant H_t^*(y_t+u_t) - H_t^*(y_t) - \langle u_t, x_t\rangle \\ &= D_{H_t^*}(y_t+u_t, y_t). \end{aligned}$$

□

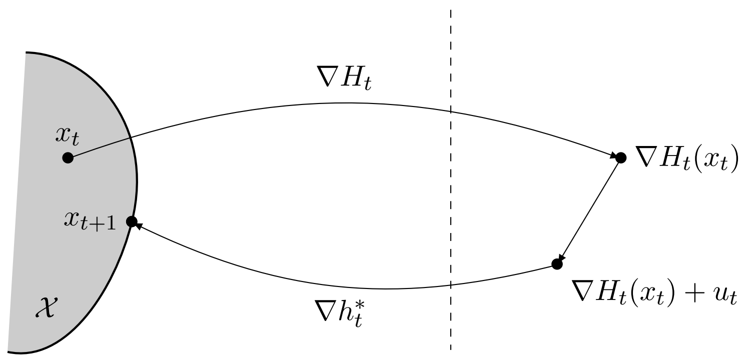


Figure 3.3: Online mirror descent seen with regularizers $h_t = H_t + I_{\mathcal{X}}$

**OMD with nonincreasing step-sizes** An important special case of OMD iterates is the following. Let $H$ be a mirror map compatible with $\mathcal{X}$, $(\gamma_t)_{t\geqslant 1}$ a positive and nonincreasing sequence, $(u_t)_{t\geqslant 1}$ a sequence in $\mathbb{R}^d$, and $x_1 \in \mathcal{X} \cap \mathcal{D}_{H_1}$ an initial point. Then, associated OMD iterates on $\mathcal{X}$ with time-dependent *step-sizes* (or *learning rate*) are defined as

$$\begin{aligned} x_{t+1} &= \operatorname*{arg\,max}_{x\in\mathcal{X}} \left\{ \langle \gamma_t u_t, x\rangle - D_H(x, x_t) \right\} && (3.10) \\ &= \operatorname*{arg\,max}_{x\in\mathcal{X}} \left\{ \langle \nabla H(x_t) + \gamma_t u_t, x\rangle - H(x) \right\}. && (3.11) \end{aligned}$$

This corresponds to OMD iterates from Definition 3.3.8 with time-dependent mirror maps $(H_t)_{t\geqslant 1}$ defined as $H_t = \gamma_t^{-1}H$ for all $t \geqslant 1$, which is indeed a sequence of mirror maps compatible with $\mathcal{X}$ with a common domain.

**Proposition 3.3.14** (Regret bounds for OMD with time-dependent step–sizes)**.** *Consider the OMD iterates with nonincreasing step-sizes defined in* (3.10)*.*

*(i) For all $T \geqslant 1$ and $x \in \mathcal{X} \cap \operatorname{dom} H$,*

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{\max_{1\leqslant t\leqslant T} D_H(x, x_t)}{\gamma_T} + \sum_{t=1}^{T} \tilde{D}_t^*,$$

*where $\tilde{D}_t^* = \gamma_t^{-1} D_{H^*}(\nabla H(x_t) + \gamma_t u_t, \nabla H_t(x_t))$.*

*(ii) Moreover, if $H$ is $K$-strongly convex with respect to some norm $\|\cdot\|$, for all $T \geqslant 1$,*

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{\max_{1\leqslant t\leqslant T} D_H(x, x_t)}{\gamma_T} + \frac{1}{2K} \sum_{t=1}^{T} \gamma_t \|u_t\|_*^2 .$$

*(iii) Moreover, if there exists known constants $R > 0$ and $L > 0$, such that for all $t \geqslant 1$, $\max_{x\in\mathcal{X}} D_H(x, x_t) \leqslant R^2$ and $\|u_t\|_* \leqslant L$, then step-sizes $\gamma_t = R\sqrt{K}/(L\sqrt{t})$ yield for all $T \geqslant 1$,*

$$\max_{x\in\mathcal{X}} \sum_{t=1}^{T} \langle u_t, x - x_t \rangle \leqslant 2RL\sqrt{\frac{T}{K}}.$$

*Proof.* (i) Applying Proposition 3.3.13,

$$\begin{aligned}
\sum_{t=1}^{T} \langle u_t, x - x_t \rangle &\leqslant \frac{D_H(x, x_1)}{\gamma_1} - \frac{D_H(x, x_{T+1})}{\gamma_{T+1}} \\
&\qquad + \sum_{t=1}^{T} \left( \frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \right) D_H(x, x_{t+1}) + \sum_{t=1}^{T} \tilde{D}_t^* \\
&= \frac{D_H(x, x_1)}{\gamma_1} + \sum_{t=1}^{T-1} \left( \frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \right) D_H(x, x_{t+1}) \\
&\qquad - \frac{D_H(x, x_{T+1})}{\gamma_T} + \sum_{t=1}^{T} \tilde{D}_t^* \\
&\leqslant \left( \max_{1\leqslant t\leqslant T} D_H(x, x_t) \right) \left( \frac{1}{\gamma_1} + \sum_{t=1}^{T-1} \left( \frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \right) \right) + \sum_{t=1}^{T} \tilde{D}_t^* \\
&= \frac{\max_{1\leqslant t\leqslant T} D_H(x, x_t)}{\gamma_T} + \sum_{t=1}^{T} \tilde{D}_t^*.
\end{aligned}$$

Besides, for all $t \geqslant 1$ and $y \in \mathbb{R}^d$,

$$H_t^*(y) = \max_{x \in \mathbb{R}^d} \left\{ \langle y, x \rangle - \frac{H(x)}{\gamma_t} \right\} = \frac{1}{\gamma_t} \max_{x \in \mathbb{R}^d} \left\{ \langle \gamma_t y, x \rangle - H(x) \right\},$$

which yields

$$\tilde{D}_t^* = D_{H_t^*}(\nabla H_t(x_t) + u_t, \nabla H_t(x_t)) = \frac{D_{H^*}(\nabla H(x_t) + \gamma_t u_t, \nabla H(x_t))}{\gamma_t}.$$

(ii) and (iii) follow immediately. □

*Remark* 3.3.15. Even for a fixed point $x \in \mathcal{X} \cap \operatorname{dom} H$, unless there is a known bound on $\max_{t \geqslant 1} D_H(x, x_t)$, for instance in the case where $\mathcal{X}$ is bounded, there is no known *horizon-free* sublinear regret bound for OMD. This is contrast with DA which, with the right nonincreasing parameters from Proposition 3.2.5, guarantees a *horizon-free* regret bound that grows as $\sqrt{T}$ with time, as soon as a bound on the norms of vectors $(u_t)_{t \geqslant 1}$ is known. When the horizon $T$ is known however, one can easily establish a $\sqrt{T}$ regret bound using OMD with a constant step-size $\gamma$, chosen as a function of $T$.

**Corollary 3.3.16.** *Consider the special case of* online gradient descent *iterations*

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t + \gamma_t u_t), \quad t \geqslant 1.$$

*Then, for all $T \geqslant 1$ and $x \in \mathcal{X}$,*

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{\max\limits_{1 \leqslant t \leqslant T} \|x - x_t\|_2^2}{2\gamma_T} + \sum_{t=1}^{T} \frac{\gamma_t \|u_t\|_2^2}{2}.$$

*Proof.* Combine Corollaries 3.3.16 and 1.6.7. □

## 3.4 Finite action set

We recall the regret minimization problem where the Decision Maker chooses at each step an element in $\{1, \dots, d\}$ possibly at random, described in Section 3.1. At step $t \geqslant 1$,

- the Decision Maker chooses $x_t \in \Delta_d$,
- Nature chooses and reveals *payoff vector* $u_t \in [0,1]^d$,
- $i_t$ is a random element in $\{1, \dots, d\}$ drawn according to distribution $x_t$ and revealed to the Decision Maker,
- the Decision Maker obtains payoff $u_{t,i_t}$.

As already discussed, the regret minimization in this setting can be reduced to an online linear optimization on the simplex $\Delta_d$, as formalized in the following lemma.

**Lemma 3.4.1.** *For all $T \geqslant 1$,*

$$\mathbb{E}\left[\max_{1\leqslant i\leqslant d}\sum_{t=1}^{T}(u_{t,i}-u_{t,i_t})\right]=\mathbb{E}\left[\max_{x\in\Delta_d}\sum_{t=1}^{T}\langle u_t,x-x_t\rangle\right].$$

Therefore, any upper bound on $\max_{x\in\Delta_d}\sum_{t=1}^{T}\langle u_t,x-x_t\rangle$ will also be an upper bound on $\mathbb{E}\left[\max_{1\leqslant i\leqslant d}\sum_{t=1}^{T}(u_{t,i}-u_{t,i_t})\right]$. A common assumption in this setting is that there is a known bound on the $\ell_\infty$ norm of the payoff vectors. We establish in Proposition 3.4.4 below the important $\sqrt{T\log d}$ regret bound in this case.

**Definition 3.4.2** (Exponential weights algorithm)**.** Let $(u_t)_{t\geqslant 1}$ be a sequence in $\mathbb{R}^d$ and $(\eta_t)_{t\geqslant 1}$ a positive sequence. The associated iterates of *exponential weights algorithm* (EW) iterates are defined in the simplex $\Delta_d$ as

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

We recall the definition of the entropic regularizer on the simplex and gather some of its key properties.

$$h_{\mathrm{ent}}(x)=\begin{cases}\sum_{i=1}^{d}x_i\log x_i & \text{if } x\in\Delta_d, \text{ with convention } 0\cdot\log 0=0\\ +\infty & \text{otherwise.}\end{cases}$$

**Proposition 3.4.3.** *(i)* $h^*_{\mathrm{ent}}(y)=\log\left(\sum_{i=1}^{d}\exp(y_i)\right)$ *for all $y\in\mathbb{R}^d$.*

*(ii)* $\nabla h^*_{\mathrm{ent}}(y)=\left(\frac{\exp(y_i)}{\sum_{j=1}^{d}\exp(y_j)}\right)_{1\leqslant i\leqslant d}$ *for all $y\in\mathbb{R}^d$,*

*(iii)* $\max_{\Delta_d}h_{\mathrm{ent}}-\min h_{\mathrm{ent}}=\log d$*,*

*(iv)* $h_{\mathrm{ent}}$ *is 1-strongly convex for $\|\cdot\|_1$.*

*Proof.* (iii) $h_{\mathrm{ent}}$ being convex, its maximum on $\Delta_d$ is attained at one of the extreme points. At each extreme point, the value of $h_{\mathrm{ent}}$ is zero. Therefore, $\max_{\Delta_d}h_{\mathrm{ent}}=0$. As for the minimum, $h_{\mathrm{ent}}$ being convex and symmetric with respect to the components $x_i$, its minimum is attained at the centroid $(1/d, \dots, 1/d)$ of the simplex $\Delta_d$, where its value is $-\log d$. Therefore, $\min_{\Delta_d} h_{\mathrm{ent}} = -\log d$, hence the result.

(iv) Consider $F : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ defined by

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

Let us prove that $F$ is 1-strongly convex with respect to $\|\cdot\|_1$. By definition, the domain of $F$ is $\mathbb{R}_+^d$. It is differentiable on the interior of the domain $(\mathbb{R}_+^*)^d$ and $\nabla F(x) = (\log x_i)_{1\leqslant i\leqslant d}$ for $x \in (\mathbb{R}_+^*)^d$. Therefore, the norm of $\nabla F(x)$ goes to $+\infty$ when $x$ converges to a boundary point of $\mathbb{R}_+^d$. [Roc70, Theorem 26.1] then assures that the subdifferential $\partial F(x)$ is empty as soon as $x \notin (\mathbb{R}_+^*)^d$. Therefore, condition (iii) from Proposition 1.6.4, which we aim at proving, can be written

$$\left\langle \nabla F(x') - \nabla F(x), x' - x \right\rangle \geqslant \left\| x' - x \right\|_1^2, \quad x, x' \in (\mathbb{R}_+^*)^d. \tag{3.12}$$

Let $x, x' \in (\mathbb{R}_+^*)^d$.

$$\left\langle \nabla F(x') - \nabla F(x), x' - x \right\rangle = \sum_{i=1}^d \log \frac{x_i'}{x_i}(x_i' - x_i).$$

A simple study of function shows that $(z-1)\log z - 2(z-1)^2/(z+1) \geqslant 0$ for $z \geqslant 0$. Applied with $z = x_i'/x_i$, this gives

$$\sum_{i=1}^d \log \frac{x_i'}{x_i}(x_i' - x_i) \geqslant \left\| x' - x \right\|_1^2,$$

and (4.4) is proved. $F$ is therefore 1-strongly convex with respect to $\|\cdot\|_1$ and so is $h_{\mathrm{ent}}$. □

**Proposition 3.4.4** (Regret bound for EW)**.** *Let $(x_t)_{t\geqslant 1}$ be defined as in Definition 3.4.2 and $y_t = \sum_{s=1}^{t-1} u_s$ for all $t \geqslant 1$. Then,*

*(i) $(x_t)_{t\geqslant 1}$ is a sequence of DA iterates associated with regularizer $h_{\mathrm{ent}}$, parameters $(\eta_t)_{t\geqslant 1}$ and dual increments $(u_t)_{t\geqslant 1}$.*

*(ii) Moreover, if $(\eta_t)_{t\geqslant 1}$ is nonincreasing, for all $T \geqslant 1$,*

$$\mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=1}^T (u_{t,i} - u_{t,i_t})\right] \leqslant \frac{\log d}{\eta_T} + \sum_{t=1}^T \frac{D_{h_{\mathrm{ent}}^*}(\eta_t(y_t + u_t), \eta_t y_t)}{\eta_t}.$$

*(iii) If there exists $L > 0$ such that $\|u_t\|_\infty \leqslant L$ for all $t \geqslant 1$, and if $\eta_t = \sqrt{\log d}/(L\sqrt{t})$, then for all $T \geqslant 1$,*

$$\mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=1}^T (u_{t,i} - u_{t,i_t})\right] \leqslant 2L\sqrt{T \log d}.$$

*Proof.* (i) follows from Definition 3.4.2 and the expression of $\nabla h^*_{\mathrm{ent}}$ from Proposition 3.4.3. (ii) and (iii) are then a simple application of Proposition 3.2.6 with the properties from Proposition 3.4.3 together with Lemma 3.4.1. $\square$

The above $\sqrt{T \log d}$ bound is known to be essentially unimprovable without further assumptions.

## 3.5 Multi-armed bandit

The multi-armed bandit problem is a variant of the regret minimization with a finite set of decisions where only the actual payoff is revealed to the Decision Maker. At step $t \geqslant 1$,

- the Decision Maker chooses $x_t \in \Delta_d$,
- Nature chooses $u_t \in \mathbb{R}^d$,
- $i_t$ is drawn in $\{1, \ldots, d\}$ according to $x_t$,
- $u_{t,i_t}$ is revealed to the Decision Maker.

We aim at obtaining guarantees on

$$\mathbb{E}\left[\max_{1 \leqslant i \leqslant d} \sum_{t=1}^{T} (u_{t,i} - u_{t,i_t})\right].$$

Since the vector $u_t$ is unknown to the Decision Maker (only $u_{t,i_t}$ is revealed), one possible approach is to construct, in the case where $x_t$ has positive coeffients, an unbiased estimator of $u_t$ as follows

$$\hat{u}_t = \left(\mathbb{1}_{\{i=i_t\}} \frac{u_{t,i_t}}{x_{t,i_t}}\right)_{1 \leqslant i \leqslant d}, \quad t \geqslant 1, \tag{3.13}$$

which indeed satisfies $\mathbb{E}\left[\hat{u}_t \,|\, x_t\right] = u_t$, and use it as a replacement to $u_t$ in some online linear optimization algorithm on the simplex, e.g. the exponential weights algorithm which indeed output points $x_t$ in the simplex with positive coefficients. The resulting algorithm is called EXP3 (for exponential weights for exploration and exploitation) and is proved below to guarantee a regret bound of order $\sqrt{Td \log d}$, in the case where the payoff vectors are bounded with respect to $\|\cdot\|_\infty$.

**Lemma 3.5.1.** *For all $T \geqslant 1$,*

$$\mathbb{E}\left[\max_{1 \leqslant i \leqslant d} \sum_{t=1}^{T} (u_{t,i} - u_{t,i_t})\right] \leqslant \mathbb{E}\left[\max_{x \in \Delta_d} \sum_{t=1}^{T} \langle \hat{u}_t, x - x_t \rangle\right].$$

*Proof.* Using the fact that $\mathbb{E}\max \geqslant \max \mathbb{E}$,

$$\begin{aligned}
\mathbb{E}\left[\max_{1\leqslant i\leqslant d}\sum_{t=1}^T \hat{u}_{t,i} - \sum_{t=1}^T \langle \hat{u}_t, x_t\rangle\right] &\geqslant \max_{1\leqslant i\leqslant d}\mathbb{E}\left[\sum_{t=1}^T \hat{u}_{t,i}\right] - \mathbb{E}\left[\sum_{t=1}^T \langle \hat{u}_t, x_t\rangle\right] \\
&= \max_{1\leqslant i\leqslant d}\mathbb{E}\left[\sum_{t=1}^T \mathbb{E}\left[\hat{u}_{t,i}\,|\,x_t\right]\right] \\
&\quad - \mathbb{E}\left[\sum_{t=1}^T \mathbb{E}\left[\langle \hat{u}_t, x_t\rangle\,|\,x_t\right]\right] \\
&= \max_{1\leqslant i\leqslant d}\mathbb{E}\left[\sum_{t=1}^T u_{t,i}\right] - \mathbb{E}\left[\sum_{t=1}^T \langle u_t, x_t\rangle\right] \\
&= \max_{1\leqslant i\leqslant d}\mathbb{E}\left[\sum_{t=1}^T u_{t,i}\right] - \mathbb{E}\left[\sum_{t=1}^T \mathbb{E}\left[u_{t,i_t}\,|\,x_t\right]\right] \\
&= \mathbb{E}\left[\max_{1\leqslant i\leqslant d}\sum_{t=1}^T u_{t,i} - \sum_{t=1}^T u_{t,i_t}\right],
\end{aligned}$$

where for the last equality, we used the fact that $u_{t,i}$ is deterministic to swap the maximum and the expectation. □

For the analysis for EXP3, we need the following bound which provides a finer control over the Bregman divergence associated with $h^*_{\text{ent}}$.

**Lemma 3.5.2.** *For $y \in \mathbb{R}^d$, $u \in \mathbb{R}^d_-$, $\eta > 0$ and $x = \nabla h^*_{\text{ent}}(\eta y)$,*

$$D_{h^*_{\text{ent}}}(\eta(y+u), \eta y) \leqslant \frac{\eta^2}{2}\sum_{i=1}^d u_i^2 x_i.$$

*Proof.* Using the explicit expressions from Proposition 3.4.3

$$\begin{aligned}
D_{h^*_{\text{ent}}}(\eta(y+u), \eta u) &= h^*_{\text{ent}}(\eta(y+u)) - h^*_{\text{ent}}(\eta y) - \langle \nabla h^*_{\text{ent}}(\eta y) | \eta u\rangle \\
&= \log\left(\sum_{i=1}^d e^{\eta(y_i+u_i)}\right) - \log\left(\sum_{i=1}^d e^{\eta y_i}\right) - \eta\langle u, x\rangle \\
&= \log\left(\sum_{i=1}^d \frac{e^{\eta u_i}e^{\eta y_i}}{\sum_{j=1}^d e^{\eta y_j}}\right) - \eta\langle u, x\rangle \\
&= \log\left(\sum_{i=1}^d x_i e^{\eta u_i}\right) - \eta\langle u, x\rangle .
\end{aligned}$$

For $z \leqslant 0$, a simple differentiation proves that $e^z \leqslant 1 + z + \frac{z^2}{2}$. Therefore,

$$\begin{aligned} D_{h^*_{\mathrm{ent}}}(\eta(y+u), \eta y) &\leqslant \log\left(\sum_{i=1}^d x_i \left(1 + \eta u_i + \frac{\eta^2 u_i^2}{2}\right)\right) - \eta \langle u, x\rangle \\ &= \log\left(1 + \eta\langle u, x\rangle + \frac{\eta^2}{2}\sum_{i=1}^d x_i u_i^2\right) - \eta\langle u, x\rangle \\ &\leqslant \eta\langle u, x\rangle + \frac{\eta^2}{2}\sum_{i=1}^d x_i u_i^2 - \eta\langle u, x\rangle \\ &= \frac{\eta^2}{2}\sum_{i=1}^d x_i u_i^2, \end{aligned}$$

which gives the result. □

**Proposition 3.5.3** (Regret bound for EXP3)**.** *Let $L > 0$ and $(u_t)_{t \geqslant 1}$ be a sequence in $[-L, 0]^d$. Let*

$$x_t = \left(\frac{\exp\left(\eta_t \sum_{s=1}^{t-1} \hat{u}_{s,i}\right)}{\sum_{j=1}^d \exp\left(\eta_t \sum_{s=1}^{t-1} \hat{u}_{s,j}\right)}\right)_{1 \leqslant i \leqslant d}, \quad t \geqslant 1,$$

*where $\hat{u}_t$ is defined as in (3.13) and $\eta_t = \sqrt{(\log d)/(dL^2 t)}$. Then, for all $T \geqslant 1$,*

$$\mathbb{E}\left[\max_{1 \leqslant i \leqslant d} \sum_{t=1}^T (u_{t,i} - u_{t,i_t})\right] \leqslant 2L\sqrt{Td\log d}.$$

*Proof.* Combine Proposition 3.4.4 with Lemmas 3.5.1 and 3.5.2. □

*Remark* 3.5.4. An important limitation of the above result is that payoff vectors must belong to $[-L, 0]^d$. Although the regret is invariant when an additive constant is added to all components of a payoff vector, the construction of the estimator (3.13) however is not invariant by such a transformation. If payoff vectors are given in e.g. $[0, L]^d$, the same regret bound can be achieved by considering

$$\hat{u}_t = \left(\mathbb{1}_{\{i = i_t\}} \frac{u_{t,i_t} - L}{x_{t,i_t}}\right)_{1 \leqslant i \leqslant d}, \quad t \geqslant 1,$$

which are unbiased estimators of $u_t - L\mathbb{1}$.

# Chapter 4

# Online convex optimization

## 4.1 Introduction

**Definition 4.1.1.** $\ell : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is a *convex loss function on* $\mathcal{X}$ if it is convex, lower semicontinuous, and has nonempty subdifferential on $\mathcal{X}$.

*Remark* 4.1.2. In particular, $\mathcal{X} \subset \operatorname{dom} \ell$ according to Proposition 1.3.6.

We consider the online convex optimization problem. At each step $t \geqslant 1$,

- the Decision Maker chooses $x_t \in \mathcal{X}$,
- Nature chooses and reveals a convex loss function $\ell_t$,
- the Decision Maker incurs loss $\ell_t(x_t)$.

We aim at constructing algorithms which provide guarantees on the corresponding *regret*, defined as

$$\sum_{t=1}^{T} \left(\ell_t(x_t) - \ell_t(x)\right), \quad T \geqslant 1,\ x \in \mathcal{X}.$$

**Example 4.1.3.** *Online linear optimization* is a special case of online convex optimization where for all $t \geqslant 1$, loss functions are of the form $\ell_t(x) = -\langle u_t, x \rangle$ for some $u_t \in \mathbb{R}^d$. Then, the above definition of the regret coincide with the definition from Section 3.

**Example 4.1.4.** In the *online prediction with square loss* problem, for all $t \geqslant 1$, loss functions are of the form $\ell_t(x) = \frac{1}{2} \left\| x - z_t \right\|_2^2$ for some $z_t \in \mathbb{R}^d$.

**Example 4.1.5.** The *online portfolio optimization problem* models the sequential rebalancing of a portfolio between $d$ assets. At each step $t \geqslant 1$, the Decision Maker chooses a distribution $x_t \in \Delta_d$ over the assets and changes the composition of its portfolio accordingly, so that the proportion (in value) corresponding to asset $i \in \{1, \ldots, d\}$ is $x_{t,i}$. At the end of the step, the value of each asset $i \in \{1, \dots, d\}$ is multiplied by a factor $r_{t,i} \in \mathbb{R}_+^*$ chosen by Nature, which corresponds to its performance. The value of the portfolio is then multiplied by $\langle r_t, x_t \rangle$. To fit into the online convex optimization framework, we consider the negative logarithm of the above quantity:

$$\ell_t(x) = -\log \langle r_t, x \rangle \,, \quad x \in \Delta_d,$$

so that the cumulative loss of the Decision Maker corresponds to the logarithm of the total variation ratio of the value of its portfolio.

## 4.2 Loss linearization

Online convex optimization can be reduced to an online linear optimization problem. Indeed, for $t \geqslant 1$, if $g_t \in \partial \ell_t(x_t)$, then by definition of a subgradient,

$$\ell_t(x_t) - \ell_t(x) \leqslant \langle g_t, x_t - x \rangle \,, \quad x \in \mathcal{X}. \tag{4.1}$$

Therefore, an upper bound on the regret corresponding to the online linear optimization problem with payoff vectors $(-g_t)_{t \geqslant 1}$ is also an upper bound on the initial regret corresponding to loss functions $(\ell_t)_{t \geqslant 1}$. The approach of using an online linear optimization algorithm with $(-g_t)_{t \geqslant 1}$ as payoff vectors is called *loss linearization*. The following guarantees are direct adaptations from Sections 3.2 and 3.3.

**Corollary 4.2.1** (Online mirror descent for convex losses)**.** *Let $K > 0$, $\|\cdot\|$ a norm in $\mathbb{R}^d$, $H$ be a mirror map compatible with $\mathcal{X}$ and $K$-strongly convex for $\|\cdot\|$, $(\gamma_t)_{t \geqslant 1}$ a positive and nonincreasing sequence, and $x_1 \in \mathcal{X} \cap \mathcal{D}_H$. Consider associated OMD iterates:*

$$x_{t+1} = \operatorname*{arg\,min}_{x \in \mathcal{X}} \left\{ \langle \gamma_t g_t, x \rangle + D_H(x, x_t) \right\}, \quad t \geqslant 1.$$

*Then for all $T \geqslant 1$ and $x \in \mathcal{X} \cap \operatorname{dom} H$,*

$$\sum_{t=1}^{T} \left( \ell_t(x_t) - \ell_t(x) \right) \leqslant \frac{\max_{1 \leqslant t \leqslant T} D_H(x, x_t)}{\gamma_T} + \frac{1}{2K} \sum_{t=1}^{T} \gamma_t \left\| g_t \right\|_*^2 .$$

*Moreover, if there exists $R, L > 0$ such that for all $t \geqslant 1$, $\max_{x \in \mathcal{X}} D_H(x, x_t) \leqslant R^2$ and $\ell_t$ is $L$-Lipschitz for $\|\cdot\|$, then the choice*

$$\gamma_t = \frac{R\sqrt{K}}{L\sqrt{t}}, \quad t \geqslant 1$$

*guarantee for all $T \geqslant 1$,*

$$\max_{x \in \mathcal{X}} \sum_{t=1}^{T} \left( \ell_t(x_t) - \ell_t(x) \right) \leqslant 2RL\sqrt{\frac{T}{K}}.$$

**Corollary 4.2.2** (Dual averaging for convex losses)**.** *Let $K > 0$, $\|\cdot\|$ a norm in $\mathbb{R}^d$, $h$ a regularizer such that* $\operatorname{dom} h \subset \mathcal{X}$*, which we assume $K$-strongly convex for $\|\cdot\|$, and $(\eta_t)_{t\geqslant 1}$ a positive and nonincreasing sequence. Consider associated DA iterates:*

$$x_t = \operatorname*{arg\,min}_{x\in\mathbb{R}^d} \left\{ \left\langle \eta_t \sum_{s=1}^{t-1} g_s, x \right\rangle + h(x) \right\}, \quad t \geqslant 1.$$

*(i) Then for all $T \geqslant 1$ and $x \in \operatorname{dom} h$,*

$$\sum_{t=1}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{h(x) - \min h}{\eta_T} + \frac{1}{2K} \sum_{t=1}^{T} \eta_t \|g_t\|_*^2 .$$

*(ii) Moreover, if there exists $L > 0$ such that for all $t \geqslant 1$, $\ell_t$ is $L$-Lipschitz continuous for $\|\cdot\|$, then the choice*

$$\eta_t = \frac{\eta\sqrt{K}}{L\sqrt{t}}, \quad t \geqslant 1,$$

*for some $\eta > 0$, guarantee for all $T \geqslant 1$,*

$$\sum_{t=1}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \left( \frac{h(x) - \min h}{\eta} + \eta \right) L\sqrt{\frac{T}{K}}.$$

*(iii) Moreover, if $\sup_{x\in\operatorname{dom} h} h(x) < +\infty$, then $\eta = \sqrt{\delta_h}$ yields,*

$$\max_{x\in\operatorname{cl}\operatorname{dom} h} \sum_{t=1}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant 2L\sqrt{\frac{\delta_h T}{K}},$$

*where $\delta_h := \sup_{x\in\operatorname{dom} h} h(x) - \min_{x\in\operatorname{dom} h} h(x)$.*

## 4.3 Follow the regularized leader

Loss linearization has the computational advantage of using only a subgradient of each loss function as the input to the algorithm, and not the whole function. On the other hand, this approach forgets about the curvature of the loss functions and convexity inequality (3.6) may be far from tight when the loss functions do have curvature.

The follow the regularized leader (FTRL) algorithm does not rely on such a majorization and instead outputs the regularized minimizer of the past cumulative loss based on the actual loss functions.

**Lemma 4.3.1.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous convex function and $\rho$ a regularizer such that* $\operatorname{dom} f \cap \operatorname{dom} \rho \neq \varnothing$*. Then, $f + \rho$ admits a unique minimizer.*

*Proof.* $f$ and $\rho$ are both lower semicontinuous, and so is $f + \rho$. $f + \rho$ is proper because $\operatorname{dom}(f + \rho) = \operatorname{dom} f \cap \operatorname{dom} \rho \neq \varnothing$. And because $f$ is convex and $\rho$ strictly convex, $f + \rho$ is also strictly convex. Let us prove that $\operatorname{dom}(f + \rho)^* = \mathbb{R}^d$, this would prove that $f + \rho$ is a regularizer, and the result would follow from Proposition 2.2.4.

$\operatorname{dom}(f+\rho)$ is a nonempty convex set, and therefore has nonempty relative interior [Roc70, Theorem 6.2]. Let $x_0$ be a point in the relative interior of $\operatorname{dom}(f + \rho)$. Then, $x_0$ also belongs to the relative interior of $\operatorname{dom} f$. By [Roc70, Theorem 23.4], $\partial f(x_0) \neq \varnothing$. Let $y_0 \in \partial f(x_0)$, which by definition satisfies

$$\forall x \in \mathbb{R}^d, \quad f(x) \geqslant f(x_0) + \langle y_0, x - x_0 \rangle .$$

Let $y \in \mathbb{R}^d$.

$$\begin{aligned}
(f + \rho)^*(y) &= \sup_{x \in \mathbb{R}^d} \left\{ \langle y, x \rangle - f(x) - \rho(x) \right\} \\
&\leqslant \sup_{x \in \mathbb{R}^d} \left\{ \langle y, x \rangle - f(x_0) - \langle y_0, x - x_0 \rangle - \rho(x) \right\} \\
&\leqslant \langle y_0, x_0 \rangle - f(x_0) + \sup_{x \in \mathbb{R}^d} \left\{ \langle y - y_0, x \rangle - \rho(x) \right\} \\
&\leqslant \langle y_0, x_0 \rangle - f(x_0) + \rho^*(y - y_0) \\
&< +\infty,
\end{aligned}$$

where $\rho^*$ only has finite values because $\rho$ is a regularizer. Hence the result. □

**Definition 4.3.2** (Follow the regularized leader)**.** Let $(\rho_t)_{t \geqslant 1}$ be a sequence of regularizers with domains contained in $\mathcal{X}$ and $(\ell_t)_{t \geqslant 1}$ a sequence of loss functions on $\mathcal{X}$. Then, the associated *follow the regularized leader* (FTRL) iterates are defined as $x_1 = \arg\min_{x \in \mathbb{R}^d} \rho_1(x)$ and

$$x_{t+1} = \arg\min_{x \in \mathbb{R}^d} \left\{ \sum_{s=1}^{t} \ell_s(x) + \rho_{t+1}(x) \right\}, \quad t \geqslant 1.$$

*Remark* 4.3.3. The above iterates are well-defined thanks to Lemma 4.3.1.

*Remark* 4.3.4. In the case of linear losses, FTRL reduces to DA from Definition 3.2.1.

**Proposition 4.3.5** (Regret bound for FTRL)**.** *Let $(x_t)_{t \geqslant 1}$ be defined as in Definition 4.3.2 and $h_1 = \rho_1 - \min \rho_1$. For all $t \geqslant 1$, let $g_t \in \partial \ell_t(x_t)$, $y_t = -\sum_{s=1}^{t-1} g_s$ and*

$$h_{t+1/2}(x) = h_{t+1}(x) = \sum_{s=1}^{t} D_{\ell_s}(x, x_s;\ g_s) + \rho_{t+1}(x) - \min \rho_{t+1}, \quad x \in \mathbb{R}^d.$$

*Then,*

*(i)* $(h_t)_{t\in 1+\frac{1}{2}\mathbb{N}}$ *is a sequence of regularizers with domains contained in* $\mathcal{X}$*,*

*(ii)* $((x_t,y_t))_{t\geqslant 1}$ *is a sequence of UMD iterates associated with regularizers* $(h_t)_{t\in 1+\frac{1}{2}\mathbb{N}}$ *and dual increments* $(-g_t)_{t\geqslant 1}$*,*

*(iii) for all* $T\geqslant 1$ *and* $x\in\bigcap_{t\geqslant 1}\operatorname{dom}\rho_t$*,*

$$\sum_{t=1}^{T}(\ell_t(x_t)-\ell_t(x))\leqslant \rho_T(x)-\min\rho_T+\sum_{t=1}^{T}D_t^*,$$

*where* $D_t^*=h_{t+1}^*\left(y_{t+1}\right)-h_t^*\left(y_t\right)+\langle g_t,x_t\rangle$*.*

*(iv) Let* $\|\cdot\|$ *be a norm in* $\mathbb{R}^d$*,* $K>0$*,* $\rho$ *a regularizer with domain contained in* $\mathcal{X}$ *which we assume to be* $K$*-strongly convex for* $\|\cdot\|$ *and* $(\eta_t)_{t\geqslant 1}$ *a positive nonincreasing sequence. If* $\rho_t=\eta_t^{-1}\rho$ *for all* $t\geqslant 1$*, then for all* $T\geqslant 1$*,*

$$\sum_{t=1}^{T}(\ell_t(x_t)-\ell_t(x))\leqslant\frac{\rho(x)-\min_{\mathcal{X}}\rho}{\eta_T}+\frac{1}{2K}\sum_{t=1}^{T}\eta_t\left\|g_t\right\|_*^2.$$

*(v) Moreover, if there exists* $L>0$ *such that loss functions* $(\ell_t)_{t\geqslant 1}$ *are* $L$*-Lipschitz continuous for* $\|\cdot\|$*, and if* $\sup_{x\in\operatorname{dom}\rho}\rho(x)<+\infty$*, then the choice*

$$\eta_t=\frac{1}{L}\sqrt{\frac{K\delta_\rho}{t}},\quad t\geqslant 1$$

*guarantees*

$$\max_{x\in\operatorname{cl}\operatorname{dom}\rho}\sum_{t=1}^{T}(\ell_t(x_t)-\ell_t(x))\leqslant 2L\sqrt{\frac{\delta_\rho T}{K}},$$

*where* $\delta_\rho:=\sup_{x\in\operatorname{dom}\rho}\rho(x)-\min\rho$*.*

*Proof.* (i) $h_1=\rho_1-\min\rho_1$ is a regularizer. For all $t\geqslant 1$, $h_{t+1}$ is lower semicontinuous as the sum of lower semi-continuous functions. It is strictly convex, as the sum of a strictly convex function (regularizer $\rho_{t+1}$) and convex functions. Besides,

$$\operatorname{dom}h_{t+1}\subset\operatorname{dom}\rho_{t+1}\subset\mathcal{X}.$$

It remains to prove that $\operatorname{dom}h_{t+1}^*=\mathbb{R}^d$. For $y\in\mathbb{R}^d$,

$$\begin{aligned}h_{t+1}^*(y)&=\sup_{y\in\mathbb{R}^d}\left\{\langle y,x\rangle-h_{t+1}(x)\right\}\\&=\sup_{y\in\mathbb{R}^d}\left\{\langle y,x\rangle-\sum_{s=1}^{t}D_{\ell_s}(x,x_s;\ g_s)-\rho_{t+1}(x)+\min\rho_{t+1}\right\}\end{aligned}$$

which is a finite quantity because Lemma 4.3.1 applies. Hence, $h_{t+1}$ is indeed a regularizer with domain contained in $\mathcal{X}$.

(ii) For all $t \geqslant 1$,

$$\begin{aligned}
x_t &= \operatorname*{arg\,min}_{x\in\mathbb{R}^d} \left\{ \sum_{s=1}^{t-1} \ell_s(x) + \rho_t(x) \right\} \\
&= \operatorname*{arg\,min}_{x\in\mathbb{R}^d} \left\{ \sum_{s=1}^{t-1} (\ell_s(x_s) + \langle g_s, x - x_s\rangle + D_{\ell_s}(x, x_s;\ g_s)) + \rho_t(x) \right\} \\
&= \operatorname*{arg\,min}_{x\in\mathbb{R}^d} \left\{ \sum_{s=1}^{t-1} (\langle g_s, x\rangle + D_{\ell_s}(x, x_s;\ g_s)) + \rho_t(x) \right\} \\
&= \operatorname*{arg\,min}_{x\in\mathbb{R}^d} \left\{ \langle -y_t, x\rangle + h_t(x) \right\} \\
&= \operatorname*{arg\,min}_{x\in\mathbb{R}^d} \left\{ \langle -y_t, x\rangle + h_t(x) \right\} \\
&= \nabla h_t^* \left(y_t\right).
\end{aligned}$$

The above is equivalent to $y_t \in \partial h_t^*(x_t)$ by Proposition 1.4.7. Because $h_{t+1/2} = h_{t+1}$, the above also imply for all $t \geqslant 1$,

$$x_{t+1} = \nabla h_{t+1/2}^*(y_{t+1}) = \nabla h_{t+1/2}^* \left(y_t - g_t\right),$$

hence the result.

(iii) Applying Lemma 2.5.2 with notation therein gives for all $t \geqslant 1$,

$$\langle g_t, x_t - x\rangle \leqslant D_t - D_{t+1} + D_t^* + (\rho_{t+1}(x) - \min \rho_{t+1}) - (\rho_t(x) - \min \rho_t) + D_{\ell_t}(x, x_t;\ g_t).$$

Reorganizing the terms gives

$$\ell_t(x_t) - \ell_t(x) = D_t - D_{t+1} + D_t^* + (\rho_{t+1}(x) - \min \rho_{t+1}) - (\rho_t(x) - \min \rho_t).$$

Summing yields

$$\sum_{t=1}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant D_1 + \sum_{t=1}^{T} D_t^* + (\rho_{T+1}(x) - \min \rho_{T+1}) - (\rho_1(x) - \min \rho_1).$$

Note that for a given $T \geqslant 1$, iterates $x_1, \dots, x_T$ do not depend on $\rho_{T+1}$, and thus, the above analysis is valid with e.g. $\rho_{T+1} = \rho_T$, and we can make the this substitution in the above right-hand side[1]. Besides,

$$D_1 = h_1(x) - h_1(x_1) - \langle 0, x - x_1\rangle = \rho_1(x) - \min \rho_1.$$

Simplifying gives the result.

[1] This argument remains valid for any choice of $\rho_{T+1}$, but we should keep in mind that $D_T^*$ depends on $\rho_{T+1}$.

(iv) Continuing the above analysis (fixing $T \geqslant 1$ and considering $\rho_{T+1} = \rho_T$, which here corresponds to $\eta_{T+1} = \eta_T$), for all $1 \leqslant t \leqslant T$, $h_t$ is $K/\eta_t$-strongly convex for $\|\cdot\|$ as the sum of convex functions and a $K/\eta_t$-strongly convex function ($\rho_t$); besides, $h_{t+1} \geqslant h_t$ because parameters $(\eta_t)_{t\geqslant 1}$ are nonincreasing. Then,

$$h_{t+1} - h_t = D_{\ell_t}(\,\cdot\,, x_t;\; g_t) + \left(\frac{1}{\eta_{t+1}} - \frac{1}{\eta_t}\right)\rho \geqslant 0,$$

which implies $h^*_{t+1} \leqslant h^*_t$. Therefore,

$$\begin{aligned} D^*_t &= h^*_{t+1}(y_t - g_t) - h^*_t(y_t) + \langle g_t, x_t\rangle \\ &\leqslant h^*_t(y_t - g_t) - h^*_t(y_t) + \langle g_t, x_t\rangle \\ &= D_{h^*_t}(y_t - g_t, y_t) \leqslant \frac{\eta_t}{2K}\|g_t\|^2_*, \end{aligned}$$

where we used Proposition 1.6.4, hence the result. (v) is then an easy consequence. □

*Remark* 4.3.6. The above last regret bound (v) provide a guarantee similar to those given in Section 4.2 by OMD and DA with loss linearization, and thus does not make explicit the advantage of FTRL. However, this bound corresponds to the case where the losses are assumed Lipschitz continuous and therefore may very well have no curvature. In the case where losses do have curvature, above regret bound (iii) may be much smaller, and FTRL performs in practice much better and OMD and DA.

## 4.4 Strongly convex losses

We now assume that loss functions have strong convexity and establish below that this can be leveraged to achieve much smaller regret bounds of order $\log T$ instead of $\sqrt{T}$.

**Example 4.4.1** (Square loss)**.** Square loss functions of the form $\ell(x) = \frac{1}{2}\|x - z\|_2^2$ for some $z \in \mathbb{R}^d$ are 1-strongly convex with for $\|\cdot\|_2$ by Corollary 1.6.7.

If the decision maker has access, after having chosen $x_t$, to a subgradient $g_t \in \partial\ell_t(x_t)$ and to the strong convexity parameter of loss function $\ell_t$, it is enough to use online gradient descent with a well-chosen step-size that depends on the strong convexity parameters to obtain the following guarantee.

**Proposition 4.4.2** (OGD with strongly convex losses)**.** *Let $(K_t)_{t\geqslant 1}$ be a positive sequence, $(\ell_t)_{t\geqslant 1}$ loss functions on $\mathcal{X}$ such that for all $t \geqslant 1$, $\ell_t$ is*

*$K_t$-strongly convex for $\|\cdot\|_2$ and $x_1 \in \mathcal{X}$. Consider online gradient descent iterates:*

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t - \gamma_t g_t), \quad t \geqslant 1,$$

*where $g_t \in \partial \ell_t(x_t)$ and $\gamma_t = \left(\sum_{s=1}^t K_s\right)^{-1}$. Then, for all $T \geqslant 1$,*

$$\max_{x \in \mathcal{X}} \sum_{t=1}^T (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{1}{2} \sum_{t=1}^T \frac{\|g_t\|_2^2}{\sum_{s=1}^t K_s}.$$

*In particular, if there exists $K, L > 0$ such that for all $t \geqslant 1$, $K_t = K$ and $\ell_t$ is $L$-Lipschitz continuous for $\|\cdot\|_2$, then for all $T \geqslant 1$,*

$$\max_{x \in \mathcal{X}} \sum_{t=1}^T (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{L^2}{2K}(1 + \log T).$$

*Proof.* Using the characterization of strong convexity from Proposition 1.6.4, we write

$$\sum_{t=1}^T (\ell_t(x_t) - \ell_t(x)) \leqslant \sum_{t=1}^T \left( \langle g_t, x_t - x \rangle - \sum_{t=1}^T \frac{K_t}{2} \|x - x_t\|_2^2 \right). \tag{4.2}$$

$(x_t)_{t \geqslant 1}$ can be seen as OMD iterates on $\mathcal{X}$ with mirror maps $H_t = (2\gamma_t)^{-1} \|\cdot\|_2^2$ for $t \geqslant 1$. Therefore, Applying Proposition 3.3.13 gives

$$\begin{aligned}
\sum_{t=1}^T \langle g_t, x_t - x \rangle \leqslant \frac{\|x - x_1\|_2^2}{2\gamma_1} - \frac{\|x - x_{T+1}\|_2^2}{2\gamma_{T+1}} \\
+ \sum_{t=1}^T \left( \frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \right) \frac{\|x - x_{t+1}\|_2^2}{2} + \sum_{t=1}^T \frac{\gamma_t \|g_t\|_2^2}{2} \\
= \frac{K_1 \|x - x_1\|_2^2}{2} - \frac{\left(\sum_{s=1}^{T+1} K_s\right) \|x - x_{T+1}\|_2^2}{2} \\
+ \sum_{t=1}^T \frac{K_{t+1} \|x - x_{t+1}\|_2^2}{2} + \sum_{t=1}^T \frac{\|g_t\|_2^2}{2 \sum_{s=1}^t K_s} \\
\leqslant \sum_{t=1}^T \frac{K_t \|x - x_t\|_2^2}{2} + \sum_{t=1}^T \frac{\|g_t\|_2^2}{2 \sum_{s=1}^t K_s}.
\end{aligned}$$

Injecting the above into (4.2) gives

$$\sum_{t=1}^T (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{1}{2} \sum_{t=1}^T \frac{\|g_t\|_2^2}{\sum_{s=1}^t K_s}.$$

The right-hand side does not depend on $x$, and the left-hand side is a strongly concave function. By Proposition 1.6.10, it admits a maximum. The result then follows by taking the maximum over $x \in \mathcal{X}$. □

The above result will be used to derive convergence guarantees for SGD the context of empirical risk minimization with convex losses and Ridge regularization.

We now turn to the *follow the leader* (FTL not FTRL) algorithm which gurantees the same regret bound as OMD, but for any norm $\|\cdot\|$, without requiring the knowledge of the norm $\|\cdot\|$ nor of the strong convexity parameters. This is therefore a much stronger guarantee.

**Proposition 4.4.3** (FTL with strongly convex losses)**.** *Let $\|\cdot\|$ be a norm in $\mathbb{R}^d$, $(K_t)_{t\geqslant 1}$ a positive sequence, and $(\ell_t)_{t\geqslant 1}$ loss functions on $\mathcal{X}$ such that for all $t \geqslant 1$, $\ell_t$ is $K_t$-strongly convex for $\|\cdot\|$ and $x_1 \in \mathcal{X}$. Consider*

$$x_{t+1} = \operatorname*{arg\,min}_{x\in\mathcal{X}} \left\{ \sum_{s=1}^{t} \ell_s(x) \right\}, \quad t \geqslant 1. \tag{4.3}$$

*Let $h_1 = \varepsilon \left\| x - x_1 \right\|_2^2 + I_{\mathcal{X}}(x)$ (for some $\varepsilon > 0$) and for all $t \geqslant 1$,*

$$h_{t+1/2}(x) = h_{t+1}(x) = \sum_{s=1}^{t} D_{\ell_s}(x, x_s;\ g_s) + I_{\mathcal{X}}(x),$$

*where for all $s \geqslant 1$, $g_s \in \partial \ell_s(x_s)$. Then,*

*(i) $(h_t)_{t\in 1+\frac{1}{2}\mathbb{N}}$ is a sequence of regularizers with $\mathcal{X}$ as domain,*

*(ii) $((x_t, -\sum_{s=1}^{t-1} g_s))_{t\geqslant 1}$ is a sequence of UMD iterates associated with regularizers $(h_t)_{t\in 1+\frac{1}{2}\mathbb{N}}$ and dual increments $(-g_t)_{t\geqslant 1}$,*

*(iii) for all $T \geqslant 1$,*

$$\max_{x\in\mathcal{X}} \sum_{t=1}^{T} \left(\ell_t(x_t) - \ell_t(x)\right) \leqslant \frac{1}{2} \sum_{t=1}^{T} \frac{\|g_t\|_*^2}{\sum_{s=1}^{t} K_s}.$$

*Proof.* The FTL iterates are well-defined because for $t \geqslant 1$, $\sum_{s=1}^{t} \ell_s + I_{\mathcal{X}}$ is a proper lower semicontinuous strongly convex function, and thus admits a unique minimizer by Proposition 1.6.10.

(i) Let $t \in 1 + \frac{1}{2}\mathbb{N}$. $h_t$ is lower semicontinuous as the sum of lower semicontinuous functions. It is also strictly convex because it is strongly convex. For $t = 1$, $\operatorname{dom} h_1 = \mathcal{X}$ as an immediate consequence of the definition and for $t \geqslant 1$, because the domain of the convex losses contain $\mathcal{X}$,

$$\mathcal{X} \subset \operatorname{dom} \left( \sum_{s=1}^{t-1} D_{\ell_s}(\,\cdot\,, x_s;\ g_s) \right),$$

and therefore

$$\operatorname{dom} h_t = \operatorname{dom} \left( \sum_{s=1}^{t-1} D_{\ell_s}(\,\cdot\,, x_s;\ g_s) + I_{\mathcal{X}} \right) = \mathcal{X}.$$

Besides, strong convexity also ensures that $\operatorname{dom} h_t^* = \mathbb{R}^d$ by Proposition 2.2.3, hence the result.

(ii) Similar to the proof of Proposition 4.3.5.

(iii) For $t \geqslant 1$, applying Lemma 2.5.2 gives, with notation therein,

$$\langle g_t, x_t - x\rangle \leqslant D_t - D_{t+1} + D_t^* + D_{\ell_t}(x, x_t;\ y_t).$$

Rearranging gives

$$\ell_t(x_t) - \ell_t(x) \leqslant D_t - D_{t+1} + D_t^*.$$

Summing gives

$$\sum_{t=1}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \varepsilon \|x - x_1\|_2^2 + \sum_{t=1}^{T} D_t^*.$$

The above being true for all $\varepsilon > 0$, the first term in the above right-hand side can be removed. It remains to bound $D_t^*$ from above. Let $t \geqslant 1$. Note that $x \mapsto D_{\ell_t}(x, x_t;\ g_t)$ is a nonnegative function which attains its minimum, which is 0, for $x = x_t$. Then, with notation $y_t = -\sum_{s=1}^{t-1} g_s$ we can write

$$\begin{aligned}
x_t = \nabla h_t^*(y_t) &= \operatorname*{arg\,min}_{x\in\mathbb{R}^d} \left\{ -\langle y_t, x\rangle + \sum_{s=1}^{t-1} D_{\ell_s}(x, x_s;\ g_s) \right\} \\
&= \operatorname*{arg\,min}_{x\in\mathbb{R}^d} \left\{ -\langle y_t, x\rangle + \sum_{s=1}^{t} D_{\ell_s}(x, x_s;\ g_s) \right\} \\
&= \nabla h_{t+1}^*(y_t) = \nabla h_{t+1/2}^*(y_t),
\end{aligned}$$

and similarly $h_t^*(y_t) = h_{t+1/2}^*(y_t) = h_{t+1}^*(y_t)$. Besides, because $\ell_t$ (and therefore $D_{\ell_t}(\,\cdot\,, x_t;\ g_t)$) is $K_t$-strongly convex for $\|\cdot\|$ by assumption, $h_{t+1} = \sum_{s=1}^{t} D_{\ell_t}(\,\cdot\,, x_t;\ g_t)$ is $\left(\sum_{s=1}^{t} K_s\right)$-strongly convex. Hence,

$$\begin{aligned}
D_t^* &= h_{t+1/2}^*(y_t - g_t) - h_t^*(y_t) + \langle g_t, x_t\rangle \\
&= h_{t+1}^*(y_t - g_t) - h_{t+1}^*(y_t) + \left\langle g_t, \nabla h_{t+1}^*(y_t)\right\rangle \\
&= D_{h_{t+1}^*}(y_t - g_t, y_t) \\
&\leqslant \frac{\|g_t\|_*^2}{2\sum_{s=1}^{t} K_s},
\end{aligned}$$

where the inequality holds by Corollary 1.6.5, hence the result. □

## 4.5 Online linear regression

We consider the following online linear regression problem, which does not exactly fit in the framework of online convex optimization. At each step $t \geqslant 1$,

- Nature chooses and reveals $w_t \in \mathbb{R}^d$,
- the Decision Maker chooses $x_t \in \mathbb{R}^d$,
- Nature chooses and reveals $z_t \in \mathbb{R}$,
- the Decision Maker incurs loss $\frac{1}{2}(\langle w_t, x_t\rangle - z_t)^2$.

The corresponding regret then writes

$$\sum_{t=1}^{T} \left( \frac{1}{2}(\langle w_t, x_t\rangle - z_t)^2 - \frac{1}{2}(\langle w_t, x\rangle - z_t)^2 \right), \quad T \geqslant 1,\ x \in \mathbb{R}^d.$$

The following algorithm is a variant of FTRL which at step $t \geqslant 1$ uses the knowledge of $w_t$ to choose $x_t$. Since the loss function at step $t \geqslant 1$ writes

$$\ell_t(x) = \frac{1}{2} \langle w_t, x\rangle^2 - z_t \langle w_t, x\rangle + \frac{1}{2} z_t^2,$$

the algorithm chooses $x_t$ by minimizing the past cumulatives loss functions *plus* the known part of the next loss function $\ell_t$, meaning the first above term $\frac{1}{2}\langle w_t, x\rangle^2$, *plus* a regularization term $\frac{\lambda}{2}\|x\|_2^2$.

**Definition 4.5.1** (Vovk–Azoury–Warmuth algorithm)**.** Let $\lambda > 0$, and $((w_t, z_t))_{t \geqslant 1}$ a sequence in $\mathbb{R}^d \times \mathbb{R}$. The associated *Vovk–Azoury–Warmuth iterates* are defined as

$$x_t = \operatorname*{arg\,min}_{x \in \mathbb{R}^d} \left\{ \frac{1}{2} \sum_{s=1}^{t-1} (z_s - \langle w_s, x\rangle)^2 + \frac{1}{2} \langle w_t, x\rangle^2 + \frac{\lambda}{2} \|x\|_2^2 \right\}, \quad t \geqslant 1.$$

**Lemma 4.5.2** (Lemma 1.11 and Theorem 11.7 in [CBL06])**.** *Let $T \geqslant 1$, $w_1, \dots, w_T \in \mathbb{R}^d$ and $\lambda > 0$. For all $1 \leqslant t \leqslant T$, denote $S_t = \lambda I + \sum_{s=1}^{t} w_s w_s^\top$. Then,*

$$\sum_{t=1}^{T} w_t^\top S_t^{-1} w_t \leqslant \sum_{i=1}^{d} \log\left(1 + \frac{\lambda_i}{\lambda}\right),$$

*where $\lambda_1, \dots, \lambda_d$ are the eigenvalues of $S_T - \lambda I$.*

**Proposition 4.5.3.** *Consider the iterates $(x_t)_{t \geqslant 1}$ defined as Definition 4.5.1, $h_1(x) = \frac{\lambda}{2}\|x\|_2^2$ and*

$$h_{t+1/2}(x) = h_{t+1}(x) = \frac{1}{2} x^\top \left( \lambda I + \sum_{s=1}^{t+1} w_s w_s^\top \right) x, \quad t \geqslant 1.$$

*Then,*

*(i)* $(h_t)_{t \in \frac{1}{2}\mathbb{N}}$ *is a sequence of regularizers on* $\mathbb{R}^d$*,*

*(ii)* $((x_t, \sum_{s=1}^{t-1} z_s w_s))_{t\geqslant 1}$ *is a sequence of UMD iterates associated with* $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ *and dual increments* $(z_t w_t)_{t\geqslant 1}$*,*

*(iii) for all* $T \geqslant 1$*,*

$$\sum_{t=1}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{\lambda}{2}\|x\|_2^2 + \frac{dZ_T^2}{2}\log\left(1 + \frac{W_T^2 T}{\lambda d}\right).$$

*where*

$$Z_T = \max_{1\leqslant t\leqslant T}|z_t| \quad \textit{and} \quad W_T = \max_{1\leqslant t\leqslant T}\|w_t\|_2\,.$$

*Proof.* (i) For $t \geqslant 1$, $\frac{\lambda}{2}I + \sum_{s=1}^{t} w_s w_s^\top$ is symmetric positive definite as the sum of a positive definite matrix ($\frac{\lambda}{2}I$) and positive semi-definite matrices. $h_t$ is therefore the corresponding squared Mahalanobis norm, which is indeed a regularizer on $\mathbb{R}^d$.

(ii) For $t \geqslant 1$, denote $y_t = \sum_{s=1}^{t-1} z_s w_s$. Then,

$$\begin{aligned}
x_t &= \operatorname*{arg\,min}_{x'\in\mathbb{R}^d}\left\{\frac{1}{2}\sum_{s=1}^{t-1}\left(\langle w_s, x'\rangle - z_s\right)^2 + \frac{1}{2}\langle w_t, x\rangle^2 + \frac{\lambda}{2}\|x\|_2^2\right\}\\
&= \operatorname*{arg\,min}_{x'\in\mathbb{R}^d}\left\{\frac{1}{2}\sum_{s=1}^{t-1}\left(\langle w_s, x'\rangle^2 - 2z_s\langle w_s, x'\rangle + z_s^2\right) + \frac{1}{2}\langle w_t, x'\rangle^2 + \frac{\lambda}{2}\|x\|_2^2\right\}\\
&= \operatorname*{arg\,min}_{x'\in\mathbb{R}^d}\left\{-\sum_{s=1}^{t-1}\langle z_s w_s, x\rangle + \frac{1}{2}\sum_{s=1}^{t}\langle w_s, x\rangle^2 + \frac{\lambda}{2}\|x\|_2^2\right\}\\
&= \nabla h_t^*(y_t),
\end{aligned}$$

which can also be written $y_t \in \partial h_t(x_t)$. Then, it also holds that

$$x_{t+1} = \nabla h_{t+1}^*(y_{t+1}) = \nabla h_{t+1/2}^*(y_{t+1}) = \nabla h_{t+1/2}^*(y_t + z_{t-1}w_{t-1}),$$

hence the result.

(iii) Applying Lemma 2.5.2 gives with notation therein

$$\langle z_t w_t, x - x_t\rangle \leqslant D_t - D_{t+1} + D_t^* + h_{t+1}(x) - h_t(x), \quad t \geqslant 1.$$

Summing gives

$$\sum_{t=1}^{T}\langle z_t w_t, x - x_t\rangle \leqslant D_1 + \sum_{t=1}^{T} D_t^* + h_{T+1}(x) - h_1(x).$$

Because iterates $x_1, \dots, x_T$ do not depend on $h_{T+1}$, the above analysis can be carried by $h_{T+1}$ replaced by $h_T$. Besides, $x_1 = y_1 = 0$ by definition, which implies $D_1 = h_1(x)$. Therefore,

$$\sum_{t=1}^{T}\langle z_t w_t, x - x_t\rangle \leqslant \sum_{t=1}^{T} D_t^* + \frac{\lambda}{2}\|x\|_2^2 + \frac{1}{2}x^\top\left(\sum_{t=1}^{T} w_t w_t^\top\right)x. \tag{4.4}$$

For each $t \geqslant 1$, denote $S_t = \frac{\lambda}{2} I + \sum_{s=1}^{t} w_s w_s^\top$ and using Fenchel's inequality from Remark 1.4.3,

$$\begin{aligned}
D_t^* &= h_{t+1}^*(y_{t+1}) - h_t^*(y_t) - \langle z_t w_t, x_t \rangle \\
&= D_{h_t^*}(y_{t+1}, y_t) + h_{t+1}^*(y_{t+1}) - h_t^*(y_{t+1}) \\
&\leqslant D_{h_t^*}(y_{t+1}, y_t) + \langle y_{t+1}, x_{t+1} \rangle - h_{t+1}(x_{t+1}) - \langle y_{t+1}, x_{t+1} \rangle + h_t(x_{t+1}) \\
&= \frac{1}{2} (z_t w_t)^\top S_t^{-1} (z_t w_t) - \langle w_{t+1}, x_{t+1} \rangle^2 ,
\end{aligned} \tag{4.5}$$

where we expressed the Bregman divergence of a squared Mahalanobis norm. Besides, the regret with respect to the actual loss function can be written as follows:

$$\begin{aligned}
&\frac{1}{2} \Big( (\langle w_t, x_t \rangle - z_t)^2 - (\langle w_t, x \rangle - z_t)^2 \Big) \\
&\qquad = \frac{1}{2} \langle w_t, x_t \rangle^2 - \frac{1}{2} \langle w_t, x \rangle^2 + z_t \langle w_t, x - x_t \rangle .
\end{aligned} \tag{4.6}$$

Combining (4.4), (4.5) and (4.6) together with Lemma 4.5.2 gives the result. □

# Chapter 5

# Blackwell's approachability

Blackwell's approachability is a very general and somewhat abstract framework for sequential decision problems, where the Decision Maker obtains after each step an outcome vector in e.g. $\mathbb{R}^d$, that depends on its decision and the decision of Nature through an *outcome function*. The original question was the following: given a subset of $\mathbb{R}^d$ called the *target*, and assuming that the outcome function is bounded, what are the conditions on the outcome function and target set such that the Decision Maker has an algorithm which guarantees that the average outcome vector converges to the target set? And when such a conditions is satisfied, what algorithm does guarantee such a convergence? This framework contains regret minimization, many variants of regret minimization, and other problems such as asymptotic calibration in statistics.

In our presentation, we restrict to target sets that are closed convex cones, which contain all special cases that are of interest to us, and simplifies the link with the tools from previous chapters. This link allows in Section 5.4 below the conversion of online linear optimization algorithms into approachability algorithms. Then in Section 5.5, the problem of regret minimization on the simplex is revisited from an approachability point of view and new algorithms are derived. Among them are the important regret matching (RM) and regret matching+ (RM+) algorithms which demonstrate excellent performance in practice, in particular in the context of learning in games. In Section 5.6, several other applications are presented.

We start with a few definitions and properties about closed convex cones.

## 5.1 Closed convex cones

The proofs are left as exercises.

**Definition 5.1.1.** A nonempty set $\mathcal{C} \subset \mathbb{R}^d$ is a *closed convex cone* if it is closed and if for all $x, x' \in \mathcal{C}$ and $\lambda \geqslant 0$, $x + x' \in \mathcal{C}$ and $\lambda x \in \mathcal{C}$.

**Proposition 5.1.2.** *A closed convex cone is convex.*

*Remark* 5.1.3. A closed convex cone $\mathcal{C}$ being nonempty, closed and convex, the Euclidean projection onto $\mathcal{C}$ is well-defined and denoted $\Pi_{\mathcal{C}}$.

**Proposition 5.1.4.** *For all $x \in \mathbb{R}^d$ and $\lambda \geqslant 0$, $\Pi_{\mathcal{C}}(\lambda x) = \lambda \Pi_{\mathcal{C}}(x)$.*

**Definition 5.1.5.** Let $\mathcal{A} \subset \mathbb{R}^d$. The *polar cone* of $\mathcal{A}$ is the set

$$\mathcal{A}^\circ = \left\{ y \in \mathbb{R}^d,\ \forall x \in \mathcal{A},\ \langle y, x \rangle \leqslant 0 \right\}.$$

**Proposition 5.1.6.** *A polar cone is a closed convex cone.*

**Proposition 5.1.7.** *If $\mathcal{C}$ is a closed convex cone, then $\mathcal{C}^{\circ\circ} = \mathcal{C}$.*

**Proposition 5.1.8.** *The negative orthant $\mathbb{R}^d_-$ is a closed convex cone and its polar cone is the positive orthant $\mathbb{R}^d_+$.*

**Theorem 5.1.9** (Moreau's decomposition theorem)**.** *Let $\mathcal{C} \subset \mathbb{R}^d$ be a closed convex cone. For all $x \in \mathbb{R}^d$, it holds that:*

$$x = \Pi_{\mathcal{C}}(x) + \Pi_{\mathcal{C}^\circ}(x),$$

*where $\Pi_{\mathcal{C}}$ and $\Pi_{\mathcal{C}^\circ}$ denote the Euclidean projection onto $\mathcal{C}$ and $\mathcal{C}^\circ$ respectively.*

## 5.2 Framework

Let $\mathcal{A}, \mathcal{B}$ be two nonempty sets with no particular structure, and $g : \mathcal{A} \times \mathcal{B} \to \mathbb{R}^d$. The elements of $\mathcal{A}$ (resp. $\mathcal{B}$) are called the *actions* of the Decision Maker (resp. of Nature). $g$ is called the *outcome function.* At step $t \geqslant 1$,

- the Decision Maker chooses action $a_t \in \mathcal{A}$,
- Nature chooses action $b_t \in \mathcal{B}$,
- outcome vector $r_t := g(a_t, b_t) \in \mathbb{R}^d$ is revealed.

We aim at defining algorithms which guarantee some bound on the distance of the average or cumulative outcome vector to a given closed convex cone $\mathcal{C} \subset \mathbb{R}^d$ called the *target.*

With no assumption, it is not possible for the Decision Maker to ensure the convergence of average outcome vectors to the target set, as the outcome function may very well output vectors which are far away from the target. The following assumption corresponds to a favorable case, which in the case where $g$ is bounded, can be proved to be a characterization of the Decision Maker having an algorithm ensuring convergence to $\mathcal{C}$.

**Definition 5.2.1.** A closed convex cone $\mathcal{C}$ satisfies *Blackwell's condition* with respect to outcome function $g$ if there exists $\alpha : \mathcal{C}^\circ \to \mathcal{A}$ such that

$$\forall x \in \mathcal{C}^\circ,\ \forall b \in \mathcal{B}, \quad \langle g(\alpha(x), b), x \rangle \leqslant 0.$$

$\alpha$ is then called an *oracle* associated with $\mathcal{C}$ and $g$.

*Remark* 5.2.2. It follows from the above definition that when Blackwell's condition is satisfied, there exists an oracle satisfying

$$x' = \lambda x \quad \text{for some } \lambda > 0 \quad \Longrightarrow \quad \alpha(x') = \alpha(x). \tag{5.1}$$

*Remark* 5.2.3 (Geometric interpretation of Blackwell's condition). Blackwell's condition means that for any given hyperplane containing the target, the Decision Maker has an action which forces the next outcome vector to belong the same side of the hyperplane as the target, regardless of Nature's action. An equivalent, and more informal interpretation is the following. Starting from a point $y \in \mathbb{R}^d$, the "direction to $\mathcal{C}$" is $\Pi_{\mathcal{C}}(y) - y$, which by Moreau's decomposition theorem belongs to $-\mathcal{C}^\circ$. $\mathcal{C}^\circ$ is in fact the set of all possible opposite "directions to $\mathcal{C}$". Blackwell's condition is then equivalent to saying that: given a "direction to $\mathcal{C}$", the Decision Maker can ensure that the outcome vector will not be in the opposite of that direction (in the sense of a negative dot product).

In some situations, it is easier to establish the following equivalent dual condition, which however does not constructively provide an oracle.

**Proposition 5.2.4** (Blackwell's dual condition)**.** *Assume that $\mathcal{A}$ and $\mathcal{B}$ are convex sets of finite dimensional vector spaces such that $\mathcal{A}$ is compact and outcome function $g$ is bi-affine. Then, a closed convex cone $\mathcal{C} \subset \mathbb{R}^d$ satisfies Blackwell's condition with respect to $g$ if, and only if,*

$$\forall b \in \mathcal{B},\ \exists a \in \mathcal{A}, \quad g(a, b) \in \mathcal{C}.$$

*Proof.* Blackwell's condition can be written

$$\max_{x \in \mathcal{C}^\circ} \min_{a \in \mathcal{A}} \max_{b \in \mathcal{B}} \langle g(a, b), x \rangle \leqslant 0.$$

Since the above dot product is affine in each of the variables $a$, $b$ and $x$, by applying Sion's minimax theorem twice, the above is equivalent to

$$\max_{b \in \mathcal{B}} \min_{a \in \mathcal{A}} \max_{x \in \mathcal{C}^\circ} \langle g(a, b), x \rangle \leqslant 0,$$

which is exactly the dual condition. □

Several examples of problems that fit into this framework are presented in Section 5.6 below.

For the remaining of this chapter, $\mathcal{C}$ will be a closed convex cone satisfying Blackwell's condition with respect to outcome function $g$ and $\alpha$ an associated oracle satisfying condition (5.1).

## 5.3 Blackwell's algorithm

We consider the framework and notation introduced in the previous section. In the case where the given target satisfies Blackwell's condition and an associated oracle is known, Blackwell's original algorithm is defined as follows.

**Definition 5.3.1.** The actions given by *Blackwell's algorithm* are defined as

$$a_t = \alpha\left(\Pi_{\mathcal{C}^\circ}\left(\sum_{s=1}^{t-1} r_s\right)\right), \quad t \geqslant 1,$$

where $\Pi_{\mathcal{C}^\circ}$ denotes the Euclidean projection onto $\mathcal{C}^\circ$.

*Remark* 5.3.2 (Geometric interpretation of Blackwell's algorithm)*.* Given the sum of past outcome vectors $R_{t-1} := \sum_{s=1}^{t-1} r_s$, Blackwell's algorithm chooses action $a_t$ that ensures that the next outcome vector $r_t$ will not be in the opposite direction (in the sense of a negative dot product) of $R_{t-1}$ to $\mathcal{C}$ (which corresponds to $\Pi_{\mathcal{C}}(R_{t-1}) - R_{t-1} = -\Pi_{\mathcal{C}^\circ}(R_{t-1})$ by Moreau's decomposition theorem).

*Remark* 5.3.3*.* An important feature is that Blackwell's algorithm does not involve any parameter or step-size to be chosen. It is said to be *parameter-free*.

The following statement first gives a general guarantee for Blackwell's algorithm with no additional assumption. In the case where the outcome vectors are bounded, the distance of the cumulative outcome to the target is bounded as $\sqrt{T}$; consequently, because of the cone structure of the target, the average outcome vector converges to the target at speed $1/\sqrt{T}$. This result is a special case of the general construction and analysis presented in Section 5.4 below, but we here give an elementary proof.

**Proposition 5.3.4** (Guarantees for Blackwell's algorithm)**.** *If $\mathcal{C} \subset \mathbb{R}^d$ is a closed convex cone satisfying Blackwell's condition, Blackwell's algorithm guarantees for all $T \geqslant 1$,*

$$\min_{r\in\mathcal{C}} \left\| \sum_{t=1}^{T} r_t - r \right\|_2 \leqslant \sqrt{\sum_{t=1}^{T} \|r_t\|_2^2}.$$

*Consequently, if there exists $L > 0$ such that $\|r_t\|_2 \leqslant L$ for all $t \geqslant 1$,*

$$\min_{r\in\mathcal{C}} \left\| \sum_{t=1}^{T} r_t - r \right\|_2 \leqslant L\sqrt{T}.$$

*Proof.* For all $t \geqslant 1$, denote $R_t = \sum_{s=1}^{t} r_t$ (and $R_0 = 0$) and $\Pi_t = \Pi_{\mathcal{C}}(R_t)$.

$$\begin{aligned}\min_{r\in\mathcal{C}} \|R_t - r\|_2^2 &\leqslant \|R_t - \Pi_{t-1}\|_2^2 = \|R_{t-1} + r_t - \Pi_{t-1}\|_2^2 \\ &= \|R_{t-1} - \Pi_{t-1}\|_2^2 + 2\left\langle R_{t-1} - \Pi_{t-1}, r_t\right\rangle + \|r_t\|_2^2 .\end{aligned}$$

We bound the above dot product using Moreau's decomposition theorem as follows:

$$\begin{aligned}\langle R_{t-1} - \Pi_{t-1}, r_t\rangle &= \langle R_{t-1} - \Pi_{\mathcal{C}}(R_{t-1}), r_t\rangle \\ &= \langle \Pi_{\mathcal{C}^\circ}(R_{t-1}), g(\alpha(\Pi_{\mathcal{C}^\circ}(R_{t-1})), b_t)\rangle \leqslant 0,\end{aligned}$$

where the last inequality holds by definition of Blackwell's algorithm and the equality by definition of oracle $\alpha$. Therefore,

$$\min_{r\in\mathcal{C}} \|R_t - r\|^2 \leqslant \min_{r\in\mathcal{C}} \|R_{t-1} - r\|^2 + \|r_t\|_2^2 .$$

The result follows from summing over $t = 1, \dots, T$. □

## 5.4 Regret-based approachability algorithms

We present a scheme which converts an online linear optimization algorithm into an approachability algorithm. In particular, we transpose the guarantees of DA with time-dependent parameters from Section 3.2 and of OMD with time-dependent step-sizes.

The following statement gives an alternative expression of the distance to the target, measured by an arbitrary norm.

**Proposition 5.4.1.** *Let $\|\cdot\|$ be a norm in $\mathbb{R}^d$ and $B$ the corresponding closed unit ball. Then,*

$$\max_{x\in\mathcal{C}^\circ\cap B} \langle y, x\rangle = \min_{y'\in\mathcal{C}} \left\|y' - y\right\|_* , \quad y \in \mathbb{R}^d.$$

*Proof.* Let $y \in \mathbb{R}^d$. Using the definition of the dual norm and Sion's minimax theorem,

$$\inf_{y'\in\mathcal{C}} \left\|y' - y\right\|_* = \inf_{y'\in\mathcal{C}} \sup_{x\in B} \left\langle y - y', x\right\rangle = \sup_{x\in B} \inf_{y'\in\mathcal{C}} \left\{\langle y, x\rangle - \left\langle y', x\right\rangle\right\} .$$

Suppose $x$ does not belong to $\mathcal{C}^\circ$. Then, there exists $y'_0 \in \mathcal{C}$ such that $\langle y'_0, x\rangle > 0$. $\mathcal{C}$ being closed by multiplication by $\mathbb{R}_+$, the quantity $\langle y', x\rangle$ (with $y' \in \mathcal{C}$) can be made arbitrarily large by selecting $y' = \lambda y'_0$ and letting $\lambda \to +\infty$, and thus the above infimum is equal to $-\infty$. Therefore, we can restrict the above supremum to $\mathcal{C}^\circ \cap B$. We thus have

$$\inf_{y'\in\mathcal{C}} \left\|y' - y\right\|_* = \sup_{x\in\mathcal{C}^\circ\cap B} \left\{\langle y, x\rangle - \sup_{y'\in\mathcal{C}} \left\langle y', x\right\rangle\right\} .$$

The above embedded supremum is zero because for $x \in \mathcal{C}^\circ \cap B$ and $y' \in \mathcal{C}$ we obviously have $\langle y', x \rangle \leqslant 0$, and 0 is attained with $y' = 0$. Finally,

$$\inf_{y' \in \mathcal{C}} \left\| y' - y \right\|_* = \sup_{x \in \mathcal{C}^\circ \cap B} \langle y, x \rangle .$$

□

The following lemma relates a general class of quantities that can measure the distance of the sum of outcomes vectors to the target with a quantity that can be interpreted as an regret in an auxiliary online linear optimization problem.

**Lemma 5.4.2.** *Let $\mathcal{X}_0 \subset \mathbb{R}^d$ a nonempty compact set, $(x_t)_{t \geqslant 1}$ a sequence in $\mathcal{C}^\circ$, $(b_t)_{t \geqslant 1}$ a sequence in $\mathcal{B}$, and for all $t \geqslant 1$, $r_t = g(\alpha(x_t), b_t)$. Then for all $T \geqslant 1$,*

$$\max_{x \in \mathcal{X}_0} \left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant \max_{x \in \mathcal{X}_0} \sum_{t=1}^{T} \langle r_t, x - x_t \rangle .$$

*Proof.* For all $t \geqslant 1$, because $\alpha$ is an oracle,

$$\langle r_t, x_t \rangle = \langle g(\alpha(x_t), b_t), x_t \rangle \leqslant 0.$$

The result follows. □

Our conversion scheme can be summarized as follows.

- Choose a nonempty compact set $\mathcal{X}_0 \subset \mathcal{C}^\circ$;
- choose an online linear optimization algorithm on an action set $\mathcal{X}$ such that $\mathcal{X}_0 \subset \mathcal{X} \subset \mathcal{C}^\circ$ and
  - use it with $(r_t)_{t \geqslant 1}$ as payoff vectors,
  - get output $(x_t)_{t \geqslant 1}$ in $\mathcal{X}$,
  - choose actions $a_t = \alpha(x_t)$ $(t \geqslant 1)$ in the initial approachability problem.

Then, according to Lemma 5.4.2, quantity $\max_{x \in \mathcal{X}_0} \left\langle \sum_{t=1}^{T} r_t, x \right\rangle$ is bounded from above by the regret bound that is offered by the online linear optimization algorithm.

In particular, any UMD iterates with regularizer with domain $\mathcal{X}$ can be converted. We here focus on two special cases: DA with time-dependent parameters and OMD with time-dependent step-sizes.

**Proposition 5.4.3** (DA for approachability)**.** *Let $K, L > 0$, $\| \cdot \|$ a norm in $\mathbb{R}^d$, $h$ a regularizer which is $K$-strongly convex for $\| \cdot \|$ and $\mathcal{X}_0$ a nonempty compact set such that $\mathcal{X}_0 \subset \operatorname{dom} h \subset \mathcal{C}^\circ$. Assume that*

- $\sup_{\mathcal{X}_0} h < +\infty$.
- *for all* $t \geqslant 1$,

$$a_t = \alpha\left(\nabla h^*\left(\eta_t \sum_{s=1}^{t-1} r_s\right)\right) \quad \text{where } \eta_t = \sqrt{\frac{K(\sup_{\mathcal{X}_0} h - \min h)}{L^2 t}},$$

- *for all* $t \geqslant 1$, $\|r_t\|_* \leqslant L$,

*Then for all* $T \geqslant 1$,

$$\max_{x \in \mathcal{X}_0} \left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant 2L\sqrt{\frac{(\sup_{\mathcal{X}_0} h - \min h)T}{K}}.$$

*Proof.* Combine Proposition 3.2.6 and Lemma 5.4.2. □

**Proposition 5.4.4** (Blackwell's algorithm is a special case of DA)**.** *Let* $(\eta_t)_{t\geqslant 1}$ *be a positive sequence and* $h_2 = \frac{1}{2}\|\cdot\|_2^2 + I_{\mathcal{C}^\circ}$ *the Euclidean regularizer on* $\mathcal{C}^\circ$*. Then, Blackwell's algorithm coincide with dual averaging associated with regularizer* $h_2$ *and parameters* $(\eta_t)_{t\geqslant 1}$*.*

*Proof.* For all $t \geqslant 1$, using Proposition 2.2.6 on the properties of the Euclidean regularizer, the DA algorithm can be rewritten as

$$\begin{aligned}
a_t &= \alpha\left(\nabla h_2^*\left(\eta_t \sum_{s=1}^{t-1} r_s\right)\right) = \alpha\left(\Pi_{\mathcal{C}^\circ}\left(\eta_t \sum_{s=1}^{t-1} r_t\right)\right) \\
&= \alpha\left(\eta_t \cdot \Pi_{\mathcal{C}^\circ}\left(\sum_{s=1}^{t-1} r_t\right)\right) = \alpha\left(\Pi_{\mathcal{C}^\circ}\left(\sum_{s=1}^{t-1} r_t\right)\right),
\end{aligned}$$

where we used Proposition 5.1.4 for the third equality and where the last inequality holds because oracle $\alpha$ satisfies condition (5.1) by assumption. This indeed coincides with the definition of Blackwell's algorithm from Definition 5.3.1. □

*Remark* 5.4.5. Blackwell's algorithm being parameter-free, the above result shows that in the special case of the Euclidean regularizer on $\mathcal{C}^\circ$, the actions chosen by dual averaging do not depend on the parameters $(\eta_t)_{t\geqslant 1}$.

**Proposition 5.4.6** (OMD for approachability)**.** *Let* $K, L, R > 0$, $\|\cdot\|$ *a norm in* $\mathbb{R}^d$*, closed sets* $\mathcal{X}_0 \subset \mathcal{X} \subset \mathcal{C}^\circ$ *where* $\mathcal{X}$ *is convex and* $\mathcal{X}_0$ *compact,* $H$ *a mirror map compatible with* $\mathcal{X}$ *which is* $K$*-strongly convex for* $\|\cdot\|$*, and* $x_1 \in \mathcal{X} \cap \operatorname{int}\operatorname{dom} H$*. Assume that*

- *for all* $t \geqslant 1$,

$$\gamma_t = \frac{R\sqrt{K}}{L\sqrt{t}}$$
$$x_{t+1} = \arg\max_{x \in \mathcal{X}} \left\{ \langle \nabla H(x_t) + \gamma_t r_t, x \rangle - H(x) \right\}$$
$$a_{t+1} = \alpha(x_{t+1}),$$

- *for all* $t \geqslant 1$, $\|r_t\|_* \leqslant L$,
- *for all* $t \geqslant 1$, $\max_{x \in \mathcal{X}_0} D_H(x, x_t) \leqslant R^2$.

*Then for all* $T \geqslant 1$,

$$\max_{x \in \mathcal{X}_0} \left\langle \sum_{t=1}^{T} r_t, x \right\rangle \leqslant 2RL\sqrt{\frac{T}{K}}.$$

*Proof.* Combine Proposition 3.3.14 and Lemma 5.4.2. □

In the case of the Euclidean mirror map, OMD is called Online Gradient Descent (OGD). We now consider the corresponding approachability algorithm which we call *greedy Blackwell*, and which can be considered as the counterpart in the OMD family of of Blackwell's algorithm. We establish below the same guarantee as for Blackwell's algorithm in Proposition 5.3.4.

**Definition 5.4.7.** The actions of the *greedy Blackwell algorithm* are given by $x_1 = 0$, $a_1 = \alpha(0)$ and for $t \geqslant 1$ by

$$x_{t+1} = \Pi_{\mathcal{C}^\circ}(x_t + r_t) \quad \text{and} \quad a_{t+1} = \alpha(x_{t+1}).$$

**Proposition 5.4.8** (Greedy Blackwell is OGD for approachability)**.** *Let* $\gamma > 0$. *The greedy Blackwell algorithm coincide with OGD on* $\mathcal{C}^\circ$ *with initial point* $x_1 = 0$ *and constant step-size* $\gamma$.

*Proof.* Denote $(x_t)_{t \geqslant 1}$ the sequence defined as in Definition 5.4.7 and $(x'_t)_{t \geqslant 1}$ the sequence from Proposition 5.4.6 associated with Euclidean mirror map $H_2 = \frac{1}{2}\|\cdot\|_2^2$ and $\mathcal{X} = \mathcal{C}^\circ$. Let us prove that $x'_t = \gamma x_t$ for all $t \geqslant 1$. It is true for $t = 1$, as $x'_1 = x_1 = 0$. Then, by induction, for $t \geqslant 1$,

$$x'_{t+1} = \Pi_{\mathcal{C}^\circ}(x'_t + \gamma r_t) = \Pi_{\mathcal{C}^\circ}(\gamma x_t + \gamma r_t) = \gamma \cdot \Pi_{\mathcal{C}^\circ}(x_t + r_t) = \gamma x_{t+1},$$

where we used Proposition 5.1.4 for the penultimate equality. The result follows from property (5.1). □

*Remark* 5.4.9. Like Blackwell's algorithm, the greedy Blackwell algorithm is parameter-free.

**Proposition 5.4.10.** *The greedy Blackwell algorithm guarantees for all $T \geqslant 1$,*

$$\min_{r \in \mathcal{C}} \left\| \sum_{t=1}^{T} r_t - r \right\|_2 \leqslant \sqrt{\sum_{t=1}^{T} \|r_t\|_2^2}.$$

*Proof.* According to Proposition 5.4.8, the greedy Blackwell algorithm corresponds to OGD on $\mathcal{C}^\circ$ with $x_1 = 0$ and any step-size $\gamma > 0$. Then the OMD regret bound from Proposition 3.3.13 (applied with constant mirror map $H = H_2/\gamma$) gives, for all $x \in \mathcal{C}^\circ$ and $T \geqslant 1$, with notation therein,

$$\sum_{t=1}^{T} \langle r_t, x - x_t \rangle \leqslant D_1 + \sum_{t=1}^{T} \tilde{D}_t^*,$$

where

$$D_1 = \frac{1}{2\gamma} \|x\|_2^2 \quad \text{and} \quad \tilde{D}_t^* = D_{H^*}(\nabla H(x_t) + r_t, \nabla H(x_t)) = \frac{\|r_t\|_2^2}{2\gamma},$$

because the Bregman divergence corresponds to the Euclidean distance in the case of the Euclidean mirror map and because $H^* = \frac{1}{2\gamma} \|\cdot\|_2^2$ as seen in Example 1.4.8. The above is true for all $\gamma > 0$. In particular, $\gamma = \left( \sum_{t=1}^{T} \|r_t\|_2^2 \right)^{-1}$ yields

$$\sum_{t=1}^{T} \langle r_t, x - x_t \rangle \leqslant \frac{1}{2} \left( \|x\|_2^2 + 1 \right) \sqrt{\sum_{t=1}^{T} \|r_t\|_2^2}.$$

Using Lemma 5.4.2 and Proposition 5.4.1 (applied with $\mathcal{X}_0 = \mathcal{C}^\circ \cap B_2$ where $B_2$ is the closed Euclidean unit ball) then gives the result. □

## 5.5 Approachability-based regret minimization on the simplex

In this section, we consider the online linear optimization problem on the simplex and rewrite it as an approachability problem. This approach yields new algorithms. In particular, the Blackwell and greedy Blackwell algorithms give in this case the so-called regret matching (RM) and regret matching+ (RM+) algorithms respectively.

In this section, $(a_t)_{t \geqslant 1}$ denote the actions of the Decision Maker in the simplex $\Delta_d$ and should not be confused with $(x_t)_{t \geqslant 1}$ which can be interpreted as the output of auxiliary online linear optimization algorithms, which belong to $\mathbb{R}_+^d$, but not necessarily to $\Delta_d$.

Denote $\mathbb{1} = (1, \dots, 1) \in \mathbb{R}^d$.

**Proposition 5.5.1** (Hart–Mas-Colell reduction)**.** *Consider the framework and notation from Section 5.2 with action sets $\mathcal{A} = \Delta_d$, $\mathcal{B} = \mathbb{R}^d$, and outcome function $g : (a, u) \mapsto u - \langle u, a \rangle \mathbb{1}$.*

*(i) $\mathbb{R}^d_-$ satisfies Blackwell's condition with respect to $g$, with*

$$\alpha(x) = \begin{cases} x / \|x\|_1 & \text{if } x \neq 0 \\ a_1 & \text{if } x = 0, \end{cases} \qquad x \in \mathbb{R}^d_+,$$

*where $a_1 \in \Delta_d$, being an associated oracle.*

*(ii) Let $(x_t)_{t \geqslant 1}$ and $(u_t)_{t \geqslant 1}$ sequences in $\mathbb{R}^d_+$ and $\mathbb{R}^d$ respectively, let $a_t = \alpha(x_t)$ and $r_t = g(a_t, u_t)$ for all $t \geqslant 1$. Then, for all $T \geqslant 1$,*

$$\max_{x \in \Delta_d} \sum_{t=1}^{T} \langle u_t, x - a_t \rangle = \max_{x \in \Delta_d} \left\langle \sum_{t=1}^{T} r_t, x \right\rangle .$$

*Proof.* (i) Let $x \in (\mathbb{R}^d_-)^\circ = \mathbb{R}^d_+$. If $x = 0$, the dot product from Blackwell's condition is zero for any value of $\alpha(x)$. If $x \neq 0$, because $\langle \mathbb{1}, x \rangle = \|x\|_1$, for all $u \in \mathbb{R}^d$,

$$\begin{aligned} \langle g(\alpha(x), u), x \rangle &= \left\langle g\left(\frac{x}{\|x\|_1}, u\right), x \right\rangle = \left\langle u - \left\langle u, \frac{x}{\|x\|_1} \right\rangle \mathbb{1}, x \right\rangle \\ &= \langle u, x \rangle - \left\langle u, \frac{x}{\|x\|_1} \right\rangle \langle \mathbb{1}, x \rangle = \langle u, x \rangle - \langle u, x \rangle = 0. \end{aligned}$$

(ii) For all $x \in \Delta_d$,

$$\begin{aligned} \left\langle \sum_{t=1}^{T} r_t, x \right\rangle &= \left\langle \sum_{t=1}^{T} g(u_t, a_t), x \right\rangle = \sum_{t=1}^{T} \langle u_t - \langle u_t, a_t \rangle \mathbb{1}, x \rangle \\ &= \sum_{t=1}^{T} (\langle u_t, x \rangle - \langle u_t, a_t \rangle \langle \mathbb{1}, x \rangle) = \sum_{t=1}^{T} \langle u_t, x - a_t \rangle . \end{aligned}$$

□

In the approachability problem described in Proposition 5.5.1, we now examine the Blackwell's algorithm and the greedy Blackwell algorithm. They are first defined below with their simplest expressions and then related to Definitions 5.3.1 and 5.4.7.

For $y \in \mathbb{R}^d$, denote $y_+$ the corresponding vector obtained by taking the positive part in each component, in other words,

$$y_+ = \Big(\max(y_i, 0)\Big)_{1 \leqslant i \leqslant d} .$$

**Definition 5.5.2.** Let $(u_t)_{t\geqslant 1}$ be a sequence in $\mathbb{R}^d$ and $a_1 \in \Delta_d$.

- The *regret matching* (RM) algorithm gives for all $t \geqslant 1$,

$$a_t = \begin{cases} \dfrac{y_{t,+}}{\|y_{t,+}\|_1} & \text{if } y_{t,+} := \left( \displaystyle\sum_{s=1}^{t-1} (u_s - \langle u_s, a_s \rangle \mathbb{1}) \right)_+ \neq 0 \\ a_1 & \text{otherwise.} \end{cases}, \quad t \geqslant 1.$$

- The *regret matching+* (RM+) algorithm is defined as $x_1 = 0$ and for all $t \geqslant 1$,

$$a_t = \begin{cases} \dfrac{x_t}{\|x_t\|_1} & \text{if } x_t \neq 0 \\ a_1 & \text{otherwise,} \end{cases}$$
$$x_{t+1} = (x_t + u_t - \langle u_t, a_t \rangle \mathbb{1})_+ .$$

**Proposition 5.5.3.** *(i) RM (resp. RM+) corresponds to Blackwell's algorithm (resp. the greedy Blackwell algorithm) in the approachability problem and oracle described in Proposition 5.5.1.*

*(ii) Against a sequence of payoff vectors $(u_t)_{t\geqslant 1}$ in $\mathbb{R}^d$, with notation from Definition 5.5.2, both RM and RM+ guarantee for all $T \geqslant 1$,*

$$\max_{x\in\Delta_d} \sum_{t=1}^T \langle u_t, x - a_t \rangle \leqslant \sqrt{\sum_{t=1}^T \|u_t - \langle u_t, a_t \rangle \mathbb{1}\|_2^2}.$$

*(iii) Moreover, if there exists $L > 0$ such that $\|u_t\|_\infty \leqslant L$ for all $t \geqslant 1$, both RM and RM+ guarantee for all $T \geqslant 1$,*

$$\max_{x\in\Delta_d} \sum_{t=1}^T \langle u_t, x - a_t \rangle \leqslant 2L\sqrt{2dT}.$$

*Proof.* (i) Because $(\mathbb{R}^d_-)^\circ = \mathbb{R}^d_+$, and considering the oracle $\alpha$ from Proposition 5.5.1, establishing $\Pi_{\mathbb{R}^d_+}(y) = y_+$ for all $y \in \mathbb{R}^d$ will make the definition of Blackwell's algorithm (resp. the greedy Blackwell algorithm) coincide with RM (resp. RM+). Let $y \in \mathbb{R}^d$.

$$\begin{aligned} \Pi_{\mathbb{R}^d_+}(y) &= \operatorname*{arg\,min}_{y'\in\mathbb{R}^d_+} \left\|y' - y\right\|_2^2 = \operatorname*{arg\,min}_{y'_1,\dots,y'_d \geqslant 0} \sum_{i=1}^d (y'_i - y_i)^2 \\ &= \left( \operatorname*{arg\,min}_{y_i \geqslant 0} (y'_i - y_i)^2 \right)_{1\leqslant i\leqslant d} = \Big( \max(y_i, 0) \Big)_{1\leqslant i\leqslant d} \\ &= y_+. \end{aligned}$$

(ii) then follows by noticing that $\Delta_d \subset \mathbb{R}_+^d \cap B_2$ (where $B_2$ is the closed unit Euclidean ball), and combining Proposition 5.4.1 with the guarantees for Blackwell's algorithm and the Greedy Blackwell algorithm from Propositions 5.3.4 and 5.4.8.

(iii) simply follows by writing for $t \geqslant 1$,

$$\begin{aligned}\|u_t - \langle u_t, a_t\rangle \mathbb{1}\|_2 &\leqslant \|u_t\|_2 + \langle u_t, a_t\rangle \|\mathbb{1}\|_2 \leqslant \sqrt{d}\,\|u_t\|_\infty + \|u_t\|_\infty \|a_t\|_1 \|\mathbb{1}\|_2 \\ &\leqslant \sqrt{d}\,\|u_t\|_\infty + \|u_t\|_\infty \|a_t\|_1 \sqrt{d} \leqslant 2L\sqrt{d}.\end{aligned}$$

□

*Remark* 5.5.4. The regret bound of order $\sqrt{dT}$ from (iii) in the case where the payoff vectors $(u_t)_{t\geqslant 1}$ are bounded by $L$ with respect to $\|\cdot\|_\infty$ is suboptimal: recall that the exponential weights algorithm achieves the optimal $\sqrt{T\log d}$ regret bound in Proposition 3.4.4. However, the latter bound needs prior knowledge of $L$ for use in the value of the parameters. In contrast, RM and RM+ are parameter-free and achieve the interestingly adaptive regret bound (ii) with no prior knowledge whatsoever.

The RM and RM+ regret minimization algorithm are of high importance because of their simplicity, ease of implementation, their adaptive character, and their excellent practical performance, in particular in the context of learning in games.

## 5.6 Further applications

We quickly mention a few problems that are variants of regret minimization and that can be solved using regret-based approachability algorithm from Section 5.4 and that either cannot be solved directly using online linear/convex optimization algorithms, or that can but for which the latter do not directly give optimal guarantees. The approach has a systematic character:

- find actions sets $\mathcal{A}, \mathcal{B}$, outcome function $g : \mathcal{A} \times \mathcal{B} \to \mathbb{R}^d$, and closed convex cone $\mathcal{C}$ satisfying Blackwell's condition with respect to $g$;
- find set $\mathcal{X}_0 \subset \mathcal{C}^\circ$ such that for sequences of actions $(a_t)_{t\geqslant 1}$ and $(b_t)_{t\geqslant 1}$ in $\mathcal{A}$ and $\mathcal{B}$ respectively, quantity

$$\max_{x\in\mathcal{X}_0} \left\langle \sum_{t=1}^T g(a_t, b_t), x \right\rangle$$

  is equal or is an upper bound on the quantity of interest in the initial problem;

- choose a closed convex set $\mathcal{X}$ such that $\mathcal{X}_0 \subset \mathcal{X} \subset \mathcal{C}^\circ$ and a sequence of regularizers $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ on $\mathcal{X}$;
- consider UMD iterates associated with regularizers $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ and dual increments $(g(a_t,b_t))_{t\geqslant 1}$;
- transpose corresponding guarantee using Lemma 5.4.2.

We only present below each problem and its reduction to and approachability problem. The choice of an algorithm in the latter and the derivation of the corresponding guarantee are left as exercises.

**Regret minimization with global costs** The problem of regret minimization with global costs is motivated by load balancing and job scheduling, where at each step, the Decision Maker first chooses a distribution (task allocation) over $d$ machines, and then observes the cost of using each machine, which may be different for each machine and each step. The goal of the Decision Maker is to minimize, not the sum of the cumulative costs of using each machine, but a given function of the vector of cumulative costs. A typical example of such *global cost* function is the $\ell_p$ norm, which includes as special cases the sum of the costs (for $p=1$), as well as the makespan i.e. the highest cumulative cost (for $p=\infty$).

Assume $d\geqslant 2$ and let $\|\cdot\|$ be a norm on $\mathbb{R}^d$. At step $t\geqslant 1$,

- the Decision Maker chooses $a_t\in\Delta_d$;
- Nature chooses loss vector $\ell_t\in[0,1]^d$.

The Decision Maker aims at minimizing the following regret:

$$\left\|\sum_{t=1}^{T} a_t\odot\ell_t\right\| - \min_{a\in\Delta_d}\left\|\frac{1}{T}\sum_{t=1}^{T} a\odot\ell_t\right\|,\quad T\geqslant 1,$$

where $\odot$ denotes the component-wise multiplication. At each step $t\geqslant 1$, the $i$-th component of vector $a_t\odot\ell$ is equal to $a_{t,i}\ell_i$ which corresponds to the cost of using machine $i$ for a fraction $a_{t,i}$ of the job. The regret is the difference between the actual global cost incurred by the Decision Maker and the best possible global cost in hindsight for a static distribution $a\in\Delta_d$. Important special cases include the makespan which corresponds to $\|\cdot\|=\|\cdot\|_\infty$: the global cost is then the highest average cost over the machines; and for $\|\cdot\|=\|\cdot\|_1$ the global cost simply corresponds to the sum of the costs of all the machines, and the problem then reduces to basic regret minimization on the simplex.

This problem can be reduced to an approachability problem by considering $\mathcal{A}=\Delta_d$, $\mathcal{B}=[0,1]^d$, outcome function $g:\mathcal{A}\times\mathcal{B}\to(\mathbb{R}^d)^2$ defined as

$$g(a,\ell)=(a\odot\ell,\ell),\quad a\in\Delta_d,\ \ell\in[0,1]^d,$$

and target set

$$\mathcal{C} = \left\{ (y, y') \in (\mathbb{R}_+^d)^2,\ \|y\| \leqslant \min_{a \in \Delta_d} \left\| a \odot y' \right\| \right\},$$

which can be proved to satisfy Blackwell's condition.

**Online combinatorial optimization** Let $d, m \geqslant 1$ be integers. Let $\mathcal{I} = \{1, \dots, d\}$ and $P$ a set which contains subsets of $\mathcal{I}$ of cardinality $m$. Denote $\Delta(P)$ the unit simplex in $\mathbb{R}^P$. At step $t \geqslant 1$,

- the Decision Maker chooses $a_t \in \Delta(P)$;
- Nature chooses and reveals $v_t \in \mathbb{R}^d$;
- the Decision Maker draws $p_t \sim a_t$ and gets payoff $\sum_{i \in p_t} v_{t,i}$.

The quantity to minimize is the following regret:

$$\max_{p \in P} \sum_{t=1}^{T} \sum_{i \in p} v_{t,i} - \sum_{t=1}^{T} \sum_{i \in p_t} v_{t,i}.$$

This problem can be seen as a basic regret minimization problem on finite set $P$, and payoff vectors $(\sum_{i \in p} v_i)_{p \in P}$ which belong to $[-m, m]^P$ as soon as we assume $v \in [-1, 1]^d$. The exponential weights algorithm would then guarantee a regret bound of order $m\sqrt{T \log |P|}$ by Proposition 3.4.4. However, it is possible to take advantage of the structure of the problem and to construct an algorithm which guarantees a significantly better bound, of order $m\sqrt{T \log(d/m)}$, which is known to be optimal. This is possible by reducing the problem problem to a well-chosen approachability problem.

Let $A$ be the $d \times |P|$ matrix defined by $A = (\mathbb{1}_{\{i \in p\}})_{\substack{i \in \mathcal{I} \\ p \in P}}$, and for each $p \in P$, let $e_p = (\mathbb{1}_{\{i \in p\}})_{i \in \mathcal{I}} \in \mathbb{R}^d$. We consider action sets $\mathcal{A} = \Delta(P)$ and $\mathcal{B} = [-1, 1]^d$, outcome function

$$g(a, v) = v - \frac{\langle v, a \rangle}{m} \mathbf{1} \in \mathbb{R}^d, \quad a \in \Delta(P),\ v \in [-1, 1]^d,$$

and target set $\mathcal{C} = A(\Delta(P))^{\circ}$ where $A(\Delta(P))$ denotes the image of the set $\Delta(P)$ via $A$ seen as a linear map from $\mathbb{R}^P$ to $\mathbb{R}^d$. $\mathcal{C}$ can then be proved to satisfy Blackwell's condition.

**Internal and swap regret** We here consider the same framework as in Section 3.4, only the quantity to be minimized is different. At step $t \geqslant 1$,

- the Decision Maker chooses $a_t \in \Delta_d$,
- Nature chooses and reveals $v_t \in \mathbb{R}^d$,

- the Decision Maker draws $i_t \sim a_t$.

Let $\Phi$ be a nonempty subset of $\mathcal{I}^{\mathcal{I}}$. The quantity to minimize is the $\Phi$-regret defined as

$$\max_{\varphi \in \Phi} \sum_{t=1}^{T} v_{t,\varphi(i_t)} - \sum_{t=1}^{T} v_{t,i_t},$$

and can be interpreted as follows. For a given map $\varphi \in \Phi$, $\sum_{t=1}^{T} v_{t,\varphi(i_t)}$ is the cumulative payoff that the Decision Maker would have obtained if he had chosen $\varphi(i)$ each time he actually chose $i$ (for all $i \in \mathcal{I}$). The $\Phi$-regret therefore compares the actual cumulative payoff of the Decision Maker with the best such quantity (for $\varphi \in \Phi$) in hindsight. Is is possible to construct an algorithm which guarantees on the $\Phi$-regret a bound of order $\sqrt{T \log |\Phi|}$ by reducing this problem to a well-chosen approachability problem.

Consider action sets $\mathcal{A} = \Delta_d$ and $\mathcal{B} = [-1,1]^d$, outcome function

$$g(a,v) = \left( \sum_{i \in \mathcal{I}} a_i (v_{\varphi(i)} - v_i) \right)_{\varphi \in \Phi}, \quad a \in \Delta_d,\ v \in \mathbb{R}^d.$$

and target set $\mathbb{R}_-^{\Phi}$, which can then be proved to satisfy Blackwell's condition.

An important special case is when $\Phi$ is the set of all transpositions of $\mathcal{I}$, in other words, the set of maps $\varphi : \mathcal{I} \to \mathcal{I}$ such that there exists $i \neq j$ in $\mathcal{I}$ such that

$$\varphi(i) = j, \quad \varphi(j) = i, \quad \text{and} \quad \varphi(k) = k \text{ for all } k \notin \{i,j\}.$$

The $\Phi$-regret is then called the *internal regret* and can be written

$$\max_{i,j \in \mathcal{I}} \sum_{t=1}^{T} \mathbb{1}_{\{i_t = i\}} (v_{t,j} - v_{t,i}).$$

# Chapter 6

# Gradient methods in optimization

Let $\mathcal{X} \subset \mathbb{R}^d$ be a nonempty closed convex set.

## 6.1 Lipschitz convex optimization

Let $L > 0$, $\|\cdot\|$ a norm in $\mathbb{R}^d$, $f : \mathbb{R}^d \to \mathbb{R}$ a convex function that is $L$-Lipschitz continuous for $\|\cdot\|$ and that admits a minimizer on $\mathcal{X}$, meaning there exists $x_* \in \mathcal{X}$ such that

$$f(x_*) = \min_{x \in \mathcal{X}} f(x).$$

The following statement, when used with $x = x_*$, relates the minimization of $f$ on $\mathcal{X}$ using subgradients with a quantity that can be interpreted as a regret in an online linear optimization problem. All algorithms and results below are then transpositions of familiar online linear optimization algorithms together with their guarantees.

**Lemma 6.1.1.** *Let $(x_t)_{t \geqslant 1}$ and $(g_t)_{t \geqslant 1}$ sequences in $\mathbb{R}^d$ such that $g_t \in \partial f(x_t)$. Then for all positive sequence $(\gamma_t)_{t \geqslant 1}$, $x \in \mathbb{R}^d$ and $T \geqslant 1$,*

$$\min_{1 \leqslant t \leqslant T} f(x_t) - f(x) \leqslant \left( \sum_{t=1}^{T} \gamma_t \right)^{-1} \sum_{t=1}^{T} \langle \gamma_t g_t, x_t - x \rangle .$$

*Proof.* For all $t \geqslant 1$, the definition a subgradient gives

$$f(x_t) - f(x) \leqslant \langle g_t, x_t - x \rangle .$$

Multiplying by $\gamma_t$ and summing over $t = 1, \dots, T$ and dividing by $\sum_{t=1}^{T} \gamma_t$ gives

$$\frac{\sum_{t=1}^{T} \gamma_t f(x_t)}{\sum_{t=1}^{T} \gamma_t} - f(x) \leqslant \left( \sum_{t=1}^{T} \gamma_t \right)^{-1} \sum_{t=1}^{T} \langle \gamma_t g_t, x_t - x \rangle .$$

The result follows. □

We first consider the following very general extension of projected (sub)gradient descent, which writes $x_{t+1} = \Pi_{\mathcal{X}}(x_t - \gamma_t g_t)$ $(t \geqslant 1)$, and then derive several corollaries.

**Proposition 6.1.2** (Strict UMD iterates for Lipschitz convex optimization)**.** *Let $K, R > 0$, $h$ a regularizer such that $x_* \in \operatorname{dom} h$ and which is $K$-strongly convex for $\|\cdot\|$, and $((x_t, y_t))_{t\geqslant 1}$ a sequence of strict UMD iterates associated with regularizer $h$ and dual increments $(-\gamma_t g_t)_{t\geqslant 1}$ where $g_t \in \partial f(x_t)$ for all $t \geqslant 1$.*

*(i) Let $T \geqslant 1$. Then,*

$$\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*) \leqslant \left(\sum_{t=1}^{T} \gamma_t\right)^{-1} \left(D_h(x_*, x_1;\ y_1) + \frac{L^2}{2K}\sum_{t=1}^{T}\gamma_t^2\right).$$

*(ii) Let $\gamma > 0$. If $\gamma_t = \gamma\sqrt{2K}/(L\sqrt{t})$ for all $t \geqslant 1$, then*

$$\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*) \leqslant \frac{L}{\sqrt{2KT}}\left(\frac{D_h(x_*, x_1;\ y_1)}{\gamma} + \gamma(1 + \log T)\right).$$

*(iii) If $D_h(x_*, x_1;\ y_1) \leqslant R^2$, then $\gamma = R$ yields,*

$$\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*) \leqslant \frac{RL}{\sqrt{2KT}}\left(2 + \log T\right).$$

**Corollary 6.1.3** (DA for Lipschitz convex optimization)**.** *Consider the assumptions from Proposition 6.1.2. In particular, the same guarantees hold if:*

$$x_t = \nabla h^*\left(y_1 - \sum_{s=1}^{t-1}\gamma_s g_s\right), \quad t \geqslant 1,$$

In the context of optimizing a unique objective function, online mirror descent is called mirror descent (MD). Proposition 6.1.2 reduces to the following.

**Corollary 6.1.4** (MD for Lipschitz convex optimization)**.** *Let $K, R, \gamma > 0$, $\|\cdot\|$ a norm on $\mathbb{R}^d$ , $H$ a mirror map compatible with $\mathcal{X}$ and $K$-strongly convex for $\|\cdot\|$, $x_1 \in \mathcal{X} \cap \operatorname{int}\operatorname{dom} H$ and for all $t \geqslant 1$,*

$$x_{t+1} = \operatorname*{arg\,min}_{x\in\mathcal{X}} \left\{\langle \nabla H(x_t) - \gamma_t g_t, x\rangle - H(x)\right\}$$

*where $g_t \in \partial f(x_t)$ and $\gamma_t = \gamma\sqrt{2K}/(L\sqrt{t})$.*

*(i) Let $T \geqslant 1$. Then,*

$$\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*) \leqslant \frac{L}{\sqrt{2KT}} \left( \frac{D_H(x_*, x_1)}{\gamma} + \gamma(1+\log T) \right).$$

*(ii) If $D_H(x_*, x_1;\ y_1) \leqslant R^2$, then $\gamma = R$ yields,*

$$\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*) \leqslant \frac{RL}{\sqrt{2KT}} (2+\log T).$$

**Corollary 6.1.5** (Projected GD for Lipschitz convex optimization)**.** *Assume that $f$ is $L$-Lipschitz continuous for $\|\cdot\|_2$. Let $R, \gamma > 0$, $x_1 \in \mathcal{X}$ and for all $t \geqslant 1$,*

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t - \gamma_t g_t),$$

*where $g_t \in \partial f(x_t)$ and $\gamma_t = \gamma\sqrt{2}/(L\sqrt{t})$.*

*(i) Let $T \geqslant 1$. Then,*

$$\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*) \leqslant \frac{L}{\sqrt{2T}} \left( \frac{\|x_* - x_1\|_2^2}{2\gamma} + \gamma(1+\log T) \right).$$

*(ii) If $\frac{1}{2}\|x_* - x_1\|_2^2 \leqslant R^2$, then $\gamma = R$ yields,*

$$\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*) \leqslant \frac{RL}{\sqrt{2T}} (2+\log T).$$

The convergence speed obtained in the above statement in $(\log T)/\sqrt{T}$. Although not an issue in practice, it is possible to shave off the $\log T$ factor using DA with time-dependent parameters instead of time-dependent stepsizes.

**Proposition 6.1.6** (DA with time-dependent parameters for Lipschitz convex optimization)**.** *Let $K, R, \eta > 0$, $h$ a regularizer such that $x_* \in \operatorname{dom} h$ and which is $K$-strongly convex for $\|\cdot\|$, $y_1 \in \mathbb{R}^d$ and*

$$x_t = \nabla h^* \left( \eta_t \left( y_1 - \sum_{t=1}^{t-1} g_t \right) \right), \quad t \geqslant 1,$$

*where $g_t \in \partial f(x_t)$ and $\eta_t = \eta\sqrt{K}/(L\sqrt{t})$ for all $t \geqslant 1$.*

*(i) Let $T \geqslant 1$. Then,*

$$\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*) \leqslant \left( \frac{h(x_*) - \min h}{\eta} + \eta \right) \frac{L}{\sqrt{2KT}}.$$

*(ii) If* $h(x_*) - \min h \leqslant R^2$*, then* $\eta = R$ *yields for all* $T \geqslant 1$*,*

$$\min_{1 \leqslant t \leqslant T} f(x_t) - f(x_*) \leqslant 2LR\sqrt{\frac{1}{KT}}.$$

*Proof.* Combine the regret bound for DA with time-dependent parameters from Proposition 3.2.6 with Lemma 6.1.1 applied with constant $\gamma_t = 1$ $(t \geqslant 1)$. □

All above guarantees need prior knowledge of Lipschitz coefficient $L$ because some parameters of step-sizes must be chosen accordingly to obtain the best bound.

## 6.2 Smooth convex optimization

Let $L > 0$, $\|\cdot\|$ be a norm in $\mathbb{R}^d$ and $f : \mathbb{R}^d \to \mathbb{R}$ be a convex that is $L$-smooth for $\|\cdot\|$, and that admits a minimizer on $\mathcal{X}$, meaning there exists $x_* \in \mathcal{X}$ such that

$$f(x_*) = \min_{x \in \mathcal{X}} f(x).$$

We first give the following very general extension of gradient descent with constant step-size and then derive corollaries.

**Proposition 6.2.1.** *Let* $K > 0$*,* $h$ *a regularizer such that* $x_* \in \operatorname{dom} h$ *and which is* $K$*-strongly convex for* $\|\cdot\|$ *and* $((x_t, y_t))_{t \geqslant 1}$ *a sequence of strict UMD iterates associated with regularizer* $h$ *and dual increments* $(-\frac{K}{L}\nabla f(x_t))_{t \geqslant 1}$*. Then for all* $T \geqslant 1$*,*

$$f(x_{T+1}) - f(x_*) \leqslant \frac{L}{K}\frac{D_h(x_*, x_1;\ y_1)}{T}.$$

*Proof.* Let $t \geqslant 1$. The definition of smoothness gives

$$D_f(x_{t+1}, x_t) \leqslant \frac{L}{2}\left\|x_{t+1} - x_t\right\|^2.$$

Moreover, using the strong convexity of $h$, we can write

$$D_f(x_{t+1}, x_t) \leqslant \frac{L}{K} D'_t, \tag{6.1}$$

where $D'_t = D_h(x_{t+1}, x_t;\ y_t)$. Applying Lemma 2.4.1 for any $x \in \operatorname{dom} h$ gives, with notation therein:

$$D_{t+1} \leqslant D_t + \frac{K}{L}\left\langle \nabla f(x_t), x - x_{t+1} \right\rangle - D'_t.$$

Using (6.1) and the convexity of $f$, we obtain

$$
\begin{aligned}
D_{t+1} &\leqslant D_t + \frac{K}{L}\left\langle \nabla f(x_t), x - x_t \right\rangle + \frac{K}{L}\left\langle \nabla f(x_t), x_t - x_{t+1} \right\rangle - \frac{K}{L} D_f(x_{t+1}, x_t) \\
&\leqslant D_t + \frac{K}{L}\left( f(x) - f(x_t) - \left\langle \nabla f(x_t), x_t - x_{t+1} \right\rangle - D_f(x_{t+1}, x_t) \right) \\
&= D_t + \frac{K}{L}\left( f(x) - f(x_{t+1}) \right).
\end{aligned}
$$

Applying the above for $x = x_t$ (which is possible because $x_t = \nabla h^*(y_t)$ does belong to $\operatorname{dom} h$) gives

$$
\begin{aligned}
f(x_{t+1}) - f(x_t) &\leqslant \frac{L}{K}\left( D_h(x_t, x_t;\ y_t) - D_h(x_t, x_{t+1};\ y_{t+1}) \right) \\
&= -\frac{L}{K} \cdot D_h(x_t, x_{t+1};\ y_{t+1}) \leqslant 0.
\end{aligned}
$$

Then, considering $x = x_*$ (which belongs to $\operatorname{dom} h$ by assumption) and summing gives

$$
T\left( f(x_{T+1}) - f(x_*) \right) \leqslant \sum_{t=1}^{T} \left( f(x_{t+1}) - f(x_*) \right) \leqslant \frac{L}{K} D_h(x_*, x_1;\ y_1),
$$

hence the result. □

**Corollary 6.2.2** (DA for smooth convex optimization)**.** *Let $h$ satisfying the assumptions from Proposition 6.2.1, $y_1 \in \mathbb{R}^d$ and*

$$
x_t = \nabla h^*\left( y_1 - \frac{L}{K} \sum_{s=1}^{t-1} \nabla f(x_s) \right), \qquad t \geqslant 1.
$$

*Then for all $T \geqslant 1$,*

$$
f(x_{T+1}) - f(x_*) \leqslant \frac{L}{K} \frac{D_h(x_*, x_1;\ y_1)}{T}.
$$

**Corollary 6.2.3** (MD for smooth convex optimization)**.** *Let $H$ be a mirror map compatible with $\mathcal{X}$ and $K$-strongly convex for $\|\cdot\|$, $x_1 \in \mathcal{X} \cap \operatorname{int} \operatorname{dom} H$ and*

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

*Then for all $T \geqslant 1$,*

$$
f(x_{T+1}) - f(x_*) \leqslant \frac{L}{K} \frac{D_H(x_*, x_1)}{T}.
$$

**Corollary 6.2.4** (Projected GD for smooth convex optimization)**.** *Assume that $f$ is $L$-smooth for $\|\cdot\|_2$. Let $x_1 \in \mathcal{X}$ and*

$$x_{t+1} = \Pi_{\mathcal{X}}\left(x_t - \frac{1}{L}\nabla f(x_t)\right), \quad t \geqslant 1.$$

*Then for all $T \geqslant 1$,*

$$f(x_{T+1}) - f(x_*) \leqslant \frac{L\,\|x_1 - x_*\|_2^2}{2T}, \quad T \geqslant 1.$$

All above guarantees need prior knowledge of $L$.

## 6.3 Nesterov's acceleration

We consider the same smooth convex optimization problem as in Section 6.2, where smoothness is with respect to a given norm $\|\cdot\|$. We now define and analyze an extension of Nesterov's accelerated gradient. This algorithm improves the convergence rate from $1/T$ to $1/T^2$.

**Proposition 6.3.1** (Accelerated UMD for smooth convex optimization)**.** *Let $K > 0$, $h$ a regularizer such that $x_* \in \operatorname{dom} h$ and which is $K$-strongly convex for $\|\cdot\|$, $(\gamma_t)_{t\geqslant 1}$ and $(\lambda_t)_{t\geqslant 1}$ positive sequences. Let $((v_t, w_t, x_t, w_t))_{t\geqslant 1}$ a sequence in $(\mathbb{R}^d)^4$ such that*

- $v_1 = w_1 = x_1 = \nabla h^*(y_1)$,
- *$((x_t, y_t))_{t\geqslant 1}$ is a sequence of strict UMD iterates associated with regularizer $h$ and dual increments $(-\gamma_t \nabla f(w_t))_{t\geqslant 1}$,*
- *For all $t \geqslant 1$,*

$$w_t = (1-\lambda_t)v_t + \lambda_t x_t, \tag{6.2}$$
$$v_{t+1} = (1-\lambda_t)v_t + \lambda_t x_{t+1}. \tag{6.3}$$

*where $\gamma_1 = 1/L$ and for all $t \geqslant 1$,*

$$\gamma_{t+1} = \frac{1 + \sqrt{1 + (2L\gamma_t)^2}}{2L}, \quad \textit{and} \quad \lambda_t = \frac{1}{L\gamma_t}.$$

*Then, for all $T \geqslant 1$,*

$$f(v_{T+1}) - f(x_*) \leqslant \frac{4LD_h(x_*, x_1;\ y_1)}{KT^2}.$$

With first gather a few properties that are immediate from the above definition.

**Lemma 6.3.2.** *With assumptions from Proposition 6.3.1, for $t \geqslant 1$,*

*(i)* $\lambda_t \in (0,1)$,

*(ii)* $w_t - x_t = (\lambda_t^{-1} - 1)(v_t - w_t)$,

*(iii)* $\gamma_t \lambda_t^{-1} - \gamma_{t+1}(\lambda_{t+1}^{-1} - 1) = 0$,

*(iv)* *For* $T \geqslant 1$, $\gamma_T \lambda_T^{-1} = \sum_{t=1}^{T} \gamma_t \geqslant \frac{KT^2}{4L}$.

Sequence $(x_t, y_t)_{t \geqslant 1}$ being strict UMD iterates, the following statement follows from Lemma 2.4.1.

**Lemma 6.3.3.** *With assumptions from Proposition 6.3.1, for $T \geqslant 1$,*

$$\sum_{t=1}^{T} \gamma_t \langle \nabla f(w_t), x_{t+1} - x_* \rangle \leqslant D_h(x_*, x_1;\ y_1) - \sum_{t=1}^{T} D_h(x_{t+1}, x_t;\ y_t).$$

*Proof of Proposition 6.3.1.* Let $t \geqslant 1$. Using the definition of smoothness of $f$ between points $w_t$ and $v_{t+1}$:

$$\begin{aligned} f(v_{t+1}) - f(w_t) &\leqslant \langle \nabla f(w_t), v_{t+1} - w_t \rangle + \frac{L}{2} \|v_{t+1} - w_t\|^2 \\ &= \lambda_t \langle \nabla f(w_t), x_{t+1} - x_t \rangle + \frac{L\lambda_t^2}{2} \|x_{t+1} - x_t\|^2 \\ &\leqslant \lambda_t \langle \nabla f(w_t), x_{t+1} - x_t \rangle + \frac{L\lambda_t^2}{K} D_h(x_{t+1}, x_t;\ y_t), \end{aligned}$$

where we used relation (6.3) from the definition of the algorithm to get the second line, and the $K$-strong convexity of $h$ (Proposition 1.6.4) to get the third line. Multiplying by $L\gamma_t^2/K$ and simplifying gives:

$$\gamma_t \lambda_t^{-1}(f(v_{t+1}) - f(w_t)) \leqslant \gamma_t \langle \nabla f(w_t), x_{t+1} - x_t \rangle + D_h(x_{t+1}, x_t;\ y_t). \tag{6.4}$$

Besides, we can write

$$\begin{aligned} x_{t+1} - x_t &= (x_{t+1} - x_*) + (x_* - w_t) + (w_t - x_t) \\ &= (x_{t+1} - x_*) + (x_* - w_t) + (\lambda_t^{-1} - 1)(v_t - w_t), \end{aligned}$$

where the second line uses relation (ii) from Lemma 6.3.2. Injecting the above into $\langle \nabla f(w_t), x_{t+1} - x_t \rangle$ gives:

$$\begin{aligned} \langle \nabla f(w_t), x_{t+1} - x_t \rangle &= \langle \nabla f(w_t), x_{t+1} - x_* \rangle + \langle \nabla f(w_t), x_* - w_t \rangle \\ &\quad + (\lambda_t^{-1} - 1) \langle \nabla f(w_t), v_t - w_t \rangle \\ &\leqslant \langle \nabla f(w_t), x_{t+1} - x_* \rangle + f(x_*) - f(w_t) \\ &\quad + (\lambda_t^{-1} - 1)\left(f(v_t) - f(w_t)\right). \end{aligned} \tag{6.5}$$

Combining inequalities (7.4) and (7.3), and summing over $t = 1, \ldots, T$ gives:

$$\sum_{t=1}^{T} \gamma_t \lambda_t^{-1} \left( f(v_{t+1}) - f(w_t) \right) \leqslant D_h(x_*, x_1;\ y_1) - \sum_{t=1}^{T} D_h(x_{t+1}, x_t;\ y_t)$$
$$+ \sum_{t=1}^{T} \gamma_t \left( f(x_*) - f(w_t) \right) + \sum_{t=1}^{T} \gamma_t (\lambda_t^{-1} - 1) \left( f(v_t) - f(w_t) \right)$$
$$+ \sum_{t=1}^{T} D_h(x_{t+1}, x_t;\ y_t),$$

where we used Lemma 6.3.3 to get the first two terms of the right-hand side. Then, simplifying and moving all values of $f$ (except for $f(x_*)$) to the left-hand side, we get:

$$\sum_{t=1}^{T} \left( \gamma_t + \gamma_t (\lambda_t^{-1} - 1) - \gamma_t \lambda_t^{-1} \right) f(w_t) + \sum_{t=1}^{T} \left( \gamma_{t-1} \lambda_{t-1}^{-1} - \gamma_t (\lambda_t^{-1} - 1) \right) f(v_t)$$
$$+ \gamma_1 (\lambda_1^{-1} - 1) f(v_1) + \gamma_T \lambda_T^{-1} f(v_{T+1}) \leqslant D_h(x_*, x_1;\ y_1) + \left( \sum_{t=1}^{T} \gamma_t \right) f(x_*).$$

The factor in front of $f(w_t)$ is clearly zero, as well as $\gamma_1(\lambda_1^{-1} - 1)$. The result then follows by applying properties (iii) and (iv) from Lemma 6.3.2. □

**Corollary 6.3.4** (Accelerated dual averaging for smooth convex optimization)**.** *Let $y_1 \in \mathbb{R}^d$, $x_1 = \nabla h^*(y_1)$ and for all $t \geqslant 1$,*

$$\begin{aligned} w_t &= (1 - \lambda_t) v_t + \lambda_t x_t, \\ x_{t+1} &= \nabla h^* \left( y_1 - \sum_{s=1}^{t} \gamma_s \nabla f(w_s) \right), \\ v_{t+1} &= (1 - \lambda_t) v_t + \lambda_t x_{t+1}, \end{aligned}$$

*where $h$, $(\gamma_t)_{t \geqslant 1}$ and $(\lambda_t)_{t \geqslant 1}$ satisfy the assumptions from Proposition 6.3.1. Then for all $T \geqslant 1$,*

$$f(v_{T+1}) - f(x_*) \leqslant \frac{4L \cdot D_h(x_*, x_1;\ y_1)}{K T^2}.$$

*Proof.* Corresponds to Proposition 6.1.2 where $y_{t+1} = y_t - \gamma_t \nabla f(w_t)$ for all $t \geqslant 1$. □

**Corollary 6.3.5** (Accelerated mirror descent smooth convex optimization)**.** *Let $H$ be a mirror map compatible with $\mathcal{X}$, $K$-strongly convex for $\| \cdot \|$, and* *such that* $x_* \in \mathcal{X} \cap \operatorname{dom} H$. *Let* $x_1 \in \mathcal{X} \cap \operatorname{int} \operatorname{dom} H$ *and for all* $t \geqslant 1$,

$$\begin{aligned} w_t &= (1-\lambda_t)v_t + \lambda_t x_t, \\ x_{t+1} &= \operatorname*{arg\,min}_{x \in \mathcal{X}} \left\{ \langle \nabla H(x_t) - \gamma_t \nabla f(w_t), x \rangle + H(x) \right\}, \\ v_{t+1} &= (1-\lambda_t)v_t + \lambda_t x_{t+1}, \end{aligned}$$

*where* $(\gamma_t)_{t \geqslant 1}$ *and* $(\lambda_t)_{t \geqslant 1}$ *satisfy the assumptions from Proposition 6.3.1. Then for all* $T \geqslant 1$,

$$f(v_{T+1}) - f(x_*) \leqslant \frac{4L \cdot D_H(x_*, x_1)}{KT^2}.$$

*Proof.* Corresponds to Proposition 6.1.2 where $y_t = \nabla H(x_t)$ for all $t \geqslant 1$. $\square$

**Corollary 6.3.6** (Accelerated gradient descent smooth convex optimization)**.** *Assume that* $f$ *is* $L$*-smooth for* $\|\cdot\|_2$ *and that* $x_* \in \mathbb{R}^d$ *is a global minimizer. Let* $x_1 = v_1 = w_1 \in \mathbb{R}^d$ *and for all* $t \geqslant 1$,

$$\begin{aligned} w_t &= (1-\lambda_t)v_t + \lambda_t x_t, \\ x_{t+1} &= x_t - \gamma_t \nabla f(w_t), \\ v_{t+1} &= (1-\lambda_t)v_t + \lambda_t x_{t+1}, \end{aligned}$$

*where* $(\gamma_t)_{t \geqslant 1}$ *and* $(\lambda_t)_{t \geqslant 1}$ *satisfy the assumptions from Proposition 6.3.1. Then, the above can be rewritten as*

$$\begin{aligned} v_{t+1} &= w_t - \lambda_t \gamma_t \nabla f(w_t) \\ w_{t+1} &= v_{t+1} + \frac{\lambda_{t+1}(1-\lambda_t)}{\lambda_t}(v_{t+1} - v_t), \end{aligned}$$

*and for all* $T \geqslant 1$,

$$f(v_{T+1}) - f(x_*) \leqslant \frac{2L \left\| x_1 - x_* \right\|_2^2}{T^2}.$$

*Proof.* The definition and the guarantee correspond to Corollary 6.1.4 with $H$ chosen as the Euclidean mirror map and $K = 1$. Besides, for $t \geqslant 1$,

$$v_{t+1} = w_t + \lambda_t(x_{t+1} - x_t) = w_t - \lambda_t \gamma_t \nabla f(w_t),$$

and

$$\begin{aligned} w_{t+1} &= (1-\lambda_{t+1})v_{t+1} + \lambda_{t+1}x_{t+1} = v_{t+1} + \lambda_{t+1}(x_{t+1} - v_{t+1}) \\ &= v_{t+1} + \lambda_{t+1}(1-\lambda_t)(x_{t+1} - v_t) \\ &= v_{t+1} + \frac{\lambda_{t+1}(1-\lambda_t)}{\lambda_t}(v_{t+1} - v_t). \end{aligned}$$

$\square$

## 6.4 Stochastic nonsmooth convex optimization

Let $f : \mathbb{R}^d \to \mathbb{R}$ a convex function that admits a minimizer on $\mathcal{X}$, meaning there exists $x_* \in \mathcal{X}$ such that

$$f(x_*) = \min_{x \in \mathcal{X}} f(x).$$

We adapt the approach from Section 6.1 to the case where the algorithm only accesses unbiaised stochastic estimators of the (sub)gradients. We present a UMD-based generalization of the celerabrated stochastic gradient descent (SGD).

**Example 6.4.1** (Finite-sum optimization)**.** Consider an objective function $f$ that is given as

$$f(x) = \frac{1}{n} \sum_{i=1}^{n} f_i(x),$$

where each function $f_i : \mathbb{R}^d \to \mathbb{R}$ ($1 \leqslant i \leqslant n$) is convex. Important such problems include empirical risk minimization, log-likelihood maximization, etc. A possible way of reducing the per-iteration computational cost is to replace, at each step $t \geqslant 1$, the computation of an exact (sub)gradient of $f$ at $x_t$ by the computation of a unbiaised stochastic estimator by drawing an index $i_t$ uniformly, and considering a subgradient $g_t \in \partial f_{i_t}(x_t)$. Then, it can be verified that indeed $\mathbb{E}\left[g_t | x_t\right] \in \partial f(x_t)$.

**Lemma 6.4.2.** *Let $(x_t)_{t \geqslant 1}$ and $(g_t)_{t \geqslant 1}$ be random sequences in $\mathbb{R}^d$ such that for all $t \geqslant 1$, $\mathbb{E}\left[g_t \,|\, x_t\right] \in \partial f(x_t)$, and $(\gamma_t)_{t \geqslant 1}$ a positive sequence, then for all $x \in \mathbb{R}^d$ and $T \geqslant 1$,*

$$\mathbb{E}\left[\min_{1 \leqslant t \leqslant T} f(x_t) - f(x)\right] \leqslant \left(\sum_{t=1}^{T} \gamma_t\right)^{-1} \mathbb{E}\left[\sum_{t=1}^{T} \langle \gamma_t g_t, x_t - x \rangle\right].$$

*Proof.* For $t \geqslant 1$,

$$\begin{aligned}
\mathbb{E}\left[\langle \gamma_t g_t, x_t - x \rangle\right] &= \mathbb{E}\left[\mathbb{E}\left[\langle \gamma_t g_t, x_t - x \rangle \,|\, x_t\right]\right] \\
&= \mathbb{E}\left[\gamma_t \left\langle \mathbb{E}\left[g_t \,|\, x_t\right], x_t - x \right\rangle\right] \\
&\geqslant \mathbb{E}\left[\gamma_t \left(f(x_t) - f(x)\right)\right] \\
&\geqslant \gamma_t \cdot \mathbb{E}\left[\min_{1 \leqslant t \leqslant T} f(x_t) - f(x)\right],
\end{aligned}$$

where the penultimate inequality uses the definition of a subgradient. The result follows. □

**Proposition 6.4.3** (Stochastic UMD for stochastic nonsmooth convex optimization)**.** *Let $K, R, \sigma > 0$, $\|\cdot\|$ a norm in $\mathbb{R}^d$, $h$ a regularizer such that $x_* \in \operatorname{dom} h$ and which is $K$-strongly convex for $\|\cdot\|$, and random sequences $(x_t)_{t \geqslant 1}$, $(y_t)_{t \geqslant 1}$ and $(g_t)_{t \geqslant 1}$ in $\mathbb{R}^d$ and $(\gamma_t)_{t \geqslant 1}$ a positive sequence such that*

- *$(x_1, y_1)$ is deterministic,*
- *$((x_t, y_t))_{t\geqslant 1}$ is a.s. a sequence of strict UMD iterates associated with regularizer $h$ and dual increments $(-\gamma_t g_t)_{t\geqslant 1}$,*
- *$\mathbb{E}\left[g_t \,|\, x_t\right] \in \partial f(x_t)$ a.s. for all $t \geqslant 1$,*
- *$\mathbb{E}\left[\|g_t\|_*^2 \,\middle|\, x_t\right] \leqslant \sigma^2$ a.s. for all $t \geqslant 1$.*

*(i) Let $T \geqslant 1$. Then,*

$$\mathbb{E}\left[\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*)\right] \leqslant \left(\sum_{t=1}^{T}\gamma_t\right)^{-1}\left(D_h(x_*, x_1;\; y_1) + \frac{\sigma^2}{2K}\sum_{t=1}^{T}\gamma_t^2\right).$$

*(ii) Let $\gamma > 0$. If $\gamma_t = \gamma\sqrt{2K}/(\sigma\sqrt{t})$ for all $t \geqslant 1$, then*

$$\mathbb{E}\left[\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*)\right] \leqslant \frac{\sigma}{\sqrt{2KT}}\left(\frac{D_h(x_*, x_1;\; y_1)}{\gamma} + \gamma(1+\log T)\right).$$

*(iii) If $D_h(x_*, x_1;\; y_1) \leqslant R^2$, then $\gamma = R$ yields,*

$$\mathbb{E}\left[\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*)\right] \leqslant \frac{R\sigma}{\sqrt{2KT}}\left(2+\log T\right).$$

*Proof.* Lemma 6.4.2 gives

$$\mathbb{E}\left[\min_{1\leqslant t\leqslant T} f(x_t) - f(x_*)\right] \leqslant \left(\sum_{t=1}^{T}\gamma_t\right)^{-1}\mathbb{E}\left[\sum_{t=1}^{T}\langle\gamma_t g_t, x_t - x_*\rangle\right].$$

Because $x_* \in \operatorname{dom} h$, UMD Lemma (Lemma 2.4.1) bounds the expectation from the above right-hand side as

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^{T}\langle\gamma_t g_t, x_t - x_*\rangle\right] &\leqslant \mathbb{E}\left[D_h(x_*, x_1;\; y_1) + \frac{1}{2K}\sum_{t=1}^{T}\|\gamma_t g_t\|_*^2\right]\\ &= D_h(x_*, x_1;\; y_1) + \frac{1}{2K}\mathbb{E}\left[\sum_{t=1}^{T}\gamma_t^2\mathbb{E}\left[\|g_t\|_*^2 \,\middle|\, x_t\right]\right]\\ &\leqslant D_h(x_*, x_1;\; y_1) + \frac{\sigma^2}{2K}\sum_{t=1}^{T}\gamma_t^2,\end{aligned}$$

which proves (i). (ii) and (iii) follow. □

# Chapter 7

# AdaGrad

We present in this chapter two instances of the AdaGrad family of algorithms. It is one of the most important innovations in the topic of regret minimization because of the adaptive property of its regret bounds and because it has led to great success in practice in optimization and deep learning through its many variants, such as RMSprop, Adam, etc.

Let $\mathcal{X}$ be a nonempty closed convex subset of $\mathbb{R}^d$.

## 7.1 Definitions

Let $(u_t)_{t\geqslant 1}$ be a sequence in $\mathbb{R}^d$, $x_1 \in \mathcal{X}$ and $\gamma > 0$.

**Definition 7.1.1.** The associated sequence of *AdaGrad-Norm* iterates on $\mathcal{X}$ is given by

$$x_{t+1} = \Pi_{\mathcal{X}}\left(x_t + \frac{\gamma}{\sqrt{\sum_{s=1}^t \|u_s\|_2^2}}u_t\right), \quad t \geqslant 1,$$

with convention $0/0 = 0$.

**Definition 7.1.2.** Let $\varepsilon > 0$. The associated sequence $(x_t)_{t\geqslant 1}$ of *AdaGrad-Diagonal* iterates on $\mathcal{X}$ is defined for each $t \geqslant 1$ as

$$\begin{aligned} x'_{t+1} &= \left(x_{t,i} + \frac{\gamma}{\varepsilon + \sqrt{\sum_{s=1}^t u_{s,i}^2}}u_{t,i}\right)_{1\leqslant i\leqslant d}, \\ x_{t+1} &= \operatorname*{arg\,min}_{x\in\mathcal{X}} \left\|x - x'_{t+1}\right\|_{A_t}, \end{aligned}$$

where

$$A_t = \operatorname{diag}\left(\varepsilon + \sqrt{\sum_{s=1}^t u_{s,i}^2}\right)_{1\leqslant i\leqslant d}.$$

## 7.2 Regret bounds

**Lemma 7.2.1.** *Let $(a_t)_{t\geqslant 1}$ be a nonnegative sequence. Then for all $T \geqslant 1$,*

$$\sum_{t=1}^{T} \frac{a_t}{\sqrt{\sum_{s=1}^{t} a_s}} \leqslant 2\sqrt{\sum_{t=1}^{T} a_t}.$$

*with convention $0/0 = 0$.*

*Proof.* For all $t \geqslant 1$, denote $b_t = \sum_{s=1}^{t} a_s$. If $b_t = 0$ for some $t \geqslant 1$, then $b_s = 0$ for all $s \leqslant t$ and the corresponding terms in the sum are zero because $(a_t)_{t\geqslant 1}$ is nonnegative by assumption. Without loss of generality, we assume $a_1 = b_1 > 0$. Then,

$$\begin{aligned}\sum_{t=1}^{T} \frac{b_t - b_{t-1}}{\sqrt{b_t}} &\leqslant \sum_{t=1}^{T} \frac{b_t}{\sqrt{b_t}} - \sum_{t=1}^{T} \frac{b_{t-1}}{\sqrt{b_t}} = \sum_{t=1}^{T} \frac{b_t}{\sqrt{b_t}} - \sum_{t=1}^{T-1} \frac{b_t}{\sqrt{b_{t+1}}} \\ &= \sqrt{b_T} + \sum_{t=1}^{T-1} b_t \left( \frac{1}{\sqrt{b_t}} - \frac{1}{\sqrt{b_{t+1}}} \right) \\ &= \sqrt{b_T} + \sum_{t=1}^{T-1} \frac{b_t(\sqrt{b_{t+1}} - \sqrt{b_t})}{\sqrt{b_t}\sqrt{b_{t+1}}} \\ &\leqslant \sqrt{b_T} + \sum_{t=1}^{T-1} (\sqrt{b_{t+1}} - \sqrt{b_t}) \leqslant 2\sqrt{b_T},\end{aligned}$$

hence the result. □

**Lemma 7.2.2.** *Let $w_1, \ldots, w_d \geqslant 0$. Then,*

$$\sum_{i=1}^{d} \sqrt{w_i} = \inf_{\substack{v \in (\mathbb{R}_+^*)^d \\ \|v\|_1 \leqslant 1}} \sqrt{\sum_{i=1}^{d} \frac{w_i}{v_i}}.$$

*Proof.* Let $v_1, \ldots, v_d > 0$ such that $\sum_{i=1}^{d} v_i \leqslant 1$. Then using Jensen's inequality for $z \mapsto z^2$,

$$\left( \sum_{i=1}^{d} \left( \frac{v_i}{\sum_{j=1}^{d} v_j} \right) \left( \frac{\sqrt{w_i}}{v_i} \right) \right)^2 \leqslant \sum_{i=1}^{d} \frac{v_i}{\sum_{j=1}^{d} v_j} \frac{w_i}{v_i^2} \leqslant \sum_{i=1}^{d} \frac{w_i}{v_i}.$$

Taking the square root gives

$$\sum_{i=1}^{d} \sqrt{w_i} \leqslant \sqrt{\sum_{i=1}^{d} \frac{w_i}{v_i}}.$$

The above inequality is an equality if $w_1 = \cdots = w_d = 0$. Otherwise, let $\varepsilon > 0$ and consider

$$v_i = \frac{\sqrt{w_i + \varepsilon}}{\sum_{j=1}^d \sqrt{w_j + \varepsilon}}, \quad 1 \leqslant i \leqslant d,$$

which are positive and satisfy $v_1 + \cdots + v_d = 1$. Then,

$$\sqrt{\sum_{i=1}^d \frac{w_i}{v_i}} = \sqrt{\sum_{i=1}^d \frac{w_i}{\sqrt{\varepsilon + w_i}} \sum_{j=1}^d \sqrt{\varepsilon + w_j}},$$

which converges to $\sum_{i=1}^d \sqrt{w_i}$ as $\varepsilon \to 0^+$, hence the result. $\square$

**Proposition 7.2.3** (Regret bound for AdaGrad-Norm)**.** *Let $(u_t)_{t \geqslant 1}$ be a sequence in $\mathbb{R}^d$, $x_1 \in \mathcal{X}$, $\gamma > 0$ and $(x_t)_{t \geqslant 1}$ the associated sequence of AdaGrad-Norm iterates on $\mathcal{X}$.*

*(i) Let $x \in \mathcal{X}$ and $T \geqslant 1$.*

$$\sum_{t=1}^T \langle u_t, x - x_t \rangle \leqslant \left( \frac{\max_{1 \leqslant t \leqslant T} \|x_t - x\|_2^2}{2\gamma} + \gamma \right) \sqrt{\sum_{t=1}^T \|u_t\|_2^2}.$$

*(ii) If $R \geqslant \max_{1 \leqslant t \leqslant T} \|x_t - x\|_2$ then $\gamma = R/\sqrt{2}$ yields*

$$\sum_{t=1}^T \langle u_t, x - x_t \rangle \leqslant R \sqrt{2 \sum_{t=1}^T \|u_t\|_2^2}.$$

*Proof.* If $u_t = 0$ for all $t \geqslant 1$, the result holds. Otherwise, consider

$$\tau = \min \{t \geqslant 1,\ u_t \neq 0\}.$$

Let $(\gamma_t)_{t \geqslant 1}$ be a positive and nonincreasing sequence defined as

$$\gamma_t = \begin{cases} \dfrac{\gamma}{\|u_\tau\|_2} & \text{if } t \leqslant \tau \\ \dfrac{\gamma}{\sqrt{\sum_{s=1}^t \|u_s\|_2^2}} & \text{if } t \geqslant \tau. \end{cases}$$

Then, $(x_t)_{t \geqslant 1}$ is a sequence of online gradient descent iterates on $\mathcal{X}$ with step-sizes $(\gamma_t)_{t \geqslant 1}$, because for $1 \leqslant t \leqslant \tau - 1$, $u_t = 0$ and thus,

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t) = \Pi_{\mathcal{X}}(x_t + \gamma_t u_t),$$

and for $t \geqslant \tau$,

$$x_{t+1} = \Pi_{\mathcal{X}} \left( x_t + \frac{\gamma}{\sqrt{\sum_{s=1}^t \|u_s\|_2^2}} u_t \right) = \Pi_{\mathcal{X}} \left( x_t + \gamma_t u_t \right).$$

If $T < \tau$, the result holds because both quantities are zero. If $T \geqslant \tau$, the regret bound for OGD with time-dependent step-sizes from Corollary 3.3.16 gives

$$\begin{aligned}
\sum_{t=1}^T \langle u_t, x - x_t \rangle &\leqslant \frac{\max\limits_{1 \leqslant t \leqslant T} \|x - x_t\|_2^2}{2\gamma_T} + \sum_{t=1}^T \frac{\gamma_t \|u_t\|_2^2}{2} \\
&= \frac{\max\limits_{1 \leqslant t \leqslant T} \|x_t - x\|_2^2}{2\gamma} \sqrt{\sum_{t=1}^T \|u_t\|_2^2} + \frac{\gamma}{2} \sum_{t=1}^T \frac{\|u_t\|_2^2}{\sqrt{\sum_{s=1}^t \|u_s\|_2^2}} \\
&\leqslant \left( \frac{\max_{1 \leqslant t \leqslant T} \|x_t - x\|_2^2}{2\gamma} + \gamma \right) \sqrt{\sum_{t=1}^T \|u_t\|_2^2}.
\end{aligned}$$

using Lemma 7.2.1. Hence the result. $\square$

**Proposition 7.2.4** (Regret bound for AdaGrad-Diagonal). *Let $(u_t)_{t \geqslant 1}$ be a sequence in $\mathbb{R}^d$, $x_1 \in \mathcal{X}$, $\varepsilon, \gamma > 0$ and $(x_t)_{t \geqslant 1}$ the associated sequence of AdaGrad-Diagonal iterates on $\mathcal{X}$.*

*(i) Then for all $x \in \mathcal{X}$ and $T \geqslant 1$,*

$$\sum_{t=1}^T \langle u_t, x - x_t \rangle \leqslant \frac{\varepsilon}{2\gamma} \|x_1 - x\|_2^2 + \left( \frac{\max_{1 \leqslant t \leqslant T} \|x_t - x\|_\infty^2}{2\gamma} + \gamma \right) \sum_{i=1}^d \sqrt{\sum_{t=1}^T u_{t,i}^2}.$$

*(ii) In particular, if $R \geqslant \max_{1 \leqslant t \leqslant T} \|x_t - x\|_\infty$, then $\gamma = R/\sqrt{2}$ yields,*

$$\sum_{t=1}^T \langle u_t, x - x_t \rangle \leqslant R \left( \frac{\varepsilon d}{\sqrt{2}} + \sum_{i=1}^d \sqrt{2 \sum_{t=1}^T u_{t,i}^2} \right).$$

*(iii) Moreover,*

$$\sum_{i=1}^d \sqrt{\sum_{t=1}^T u_{t,i}^2} = \inf_{\substack{v \in (\mathbb{R}_+^*)^d \\ \|v\|_1 \leqslant d}} \sqrt{d \sum_{t=1}^T \|u_t\|_{\operatorname{diag}(v)^{-1}}^2} \leqslant \sqrt{d \sum_{t=1}^T \|u_t\|_2^2}.$$

*Proof.* Let $(A_t)_{t\geqslant 1}$ be the sequence of symmetric positive definite matrices defined as

$$A_t = \operatorname{diag}\left(\varepsilon + \sqrt{\sum_{t=1}^{T} u_{t,i}^2}\right)_{1\leqslant i\leqslant d},$$

and $(H_t)_{t\geqslant 1}$ the sequence of mirror maps defined as

$$H_t(x) = \frac{1}{2\gamma}\left\|x\right\|_{A_t}^2 = \frac{1}{2\gamma}\left\langle x, A_t x\right\rangle, \quad x\in\mathbb{R}^d,\ t\geqslant 1.$$

Then, $(x_t)_{t\geqslant 1}$ corresponds to a sequence of online mirror descent iterates associated with mirror maps $(H_t)_{t\geqslant 1}$ and dual increments $(u_t)_{t\geqslant 1}$. Then Proposition 3.3.13 gives, with notation therein,

$$\sum_{t=1}^{T}\left\langle u_t, x - x_t\right\rangle \leqslant D_{H_1}(x,x_1) - D_{H_{T+1}}(x,x_{T+1}) + \sum_{t=1}^{T}\tilde{D}_{t+1/2}^{\Delta} + \sum_{t=1}^{T}\tilde{D}_t^*, \quad (7.1)$$

where for $t\geqslant 1$,

$$\tilde{D}_{t+1/2}^{\Delta} = D_{H_{t+1}-H_t}(x,x_{t+1}) = D_{H_{t+1}}(x,x_{t+1}) - D_{H_t}(x,x_{t+1}), \qquad (7.2)$$

and since $H_t$ is $1/\gamma$-strongly convex for $\|\cdot\|_{A_t}$ by Corollary 1.6.8,

$$\tilde{D}_t^* = D_{H_t^*}(\nabla H_t(x_t)+u_t, \nabla H_t(x_t)) \leqslant \frac{\gamma}{2}\left\|u_t\right\|_{A_t^{-1}}^2.$$

Summing $\tilde{D}_{t+1/2}^{\Delta}$ and using using the expression of the associated Bregman divergence from Example 1.5.7 gives

$$\begin{aligned}\sum_{t=1}^{T}\tilde{D}_{t+1/2}^{\Delta} &= \sum_{t=1}^{T}\left(D_{H_{t+1}}(x,x_{t+1}) - D_{H_t}(x,x_{t+1})\right)\\ &= D_{H_{T+1}}(x,x_{T+1}) - D_{H_T}(x,x_{T+1}) + \sum_{t=1}^{T} D_{H_t-H_{t-1}}(x,x_t)\end{aligned} \qquad (7.3)$$

and the above last sum is bounded as

$$\begin{aligned}
\sum_{t=1}^{T} D_{H_t - H_{t-1}}(x, x_t) &= \frac{1}{2\gamma} \sum_{t=1}^{T} \|x - x_t\|^2_{A_t - A_{t-1}} \\
&= \frac{1}{2\gamma} \sum_{t=1}^{T} \sum_{i=1}^{d} (x_i - x_{t,i})^2 (A_{t,ii} - A_{t-1,ii}) \\
&\leqslant \frac{\max\limits_{1 \leqslant t \leqslant T} \|x - x_t\|^2_\infty}{2\gamma} \sum_{i=1}^{d} \sum_{t=1}^{T} (A_{t,ii} - A_{t-1,ii}) \\
&= \frac{\max\limits_{1 \leqslant t \leqslant T} \|x - x_t\|^2_\infty}{2\gamma} \sum_{i=1}^{d} (A_{T,ii} - A_{1,ii}) \\
&= \frac{\max\limits_{1 \leqslant t \leqslant T} \|x - x_t\|^2_\infty}{2\gamma} \sum_{i=1}^{d} \left( \varepsilon + \sqrt{\sum_{t=1}^{T} u_{t,i}^2} - \varepsilon - |u_{1,i}| \right) \\
&= \frac{\max\limits_{1 \leqslant t \leqslant T} \|x - x_t\|^2_\infty}{2\gamma} \left( \sum_{i=1}^{d} \sqrt{\sum_{t=1}^{T} u_{t,i}^2} \right) \\
&\quad - \frac{\max\limits_{1 \leqslant t \leqslant T} \|x - x_t\|^2_\infty}{2\gamma} \sum_{i=1}^{d} |u_{1,i}| \,.
\end{aligned} \tag{7.4}$$

Besides,

$$\begin{aligned}
D_H(x, x_1) &= \frac{1}{2\gamma} \|x - x_1\|^2_{A_1} = \frac{1}{2\gamma} \sum_{i=1}^{d} (x_i - x_{1,i})^2 (\varepsilon + |u_{1,i}|) \\
&\leqslant \frac{\varepsilon}{2\gamma} \|x - x_1\|_2^2 + \frac{\max\limits_{1 \leqslant t \leqslant T} \|x - x_t\|^2_\infty}{2\gamma} \sum_{i=1}^{d} |u_{1,i}|
\end{aligned} \tag{7.5}$$

Finally, summing $\tilde{D}_t^*$ gives

$$\begin{aligned}
\sum_{t=1}^{T} \tilde{D}_t^* &\leqslant \frac{\gamma}{2} \sum_{t=1}^{T} \|u_t\|^2_{A_t^{-1}} = \frac{\gamma}{2} \sum_{t=1}^{T} \sum_{i=1}^{d} \frac{u_{t,i}^2}{\varepsilon + \sqrt{\sum_{s=1}^{t} u_{s,i}^2}} \\
&\leqslant \frac{\gamma}{2} \sum_{t=1}^{T} \sum_{i=1}^{d} \frac{u_{t,i}^2}{\sqrt{\sum_{s=1}^{t} u_{s,i}^2}} \leqslant \gamma \sum_{i=1}^{d} \sqrt{\sum_{t=1}^{T} u_{t,i}^2},
\end{aligned} \tag{7.6}$$

where we used Lemma 7.2.1 for the last inequality.

Then combining (7.1), (7.2), (7.3), (7.4), (7.5) and (7.6) gives (i), and (ii) follows. Using Lemma 7.2.2 gives (iii). □

## 7.3 Application to nonsmooth convex optimization

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function that admits a minimizer $\mathcal{X}$, in other words, there exists $x_* \in \mathcal{X}$ such that

$$f(x_*) = \min_{x \in \mathcal{X}} f(x).$$

The following result demonstrates that AdaGrad-Norm is *adaptive* to the Lipschitz continuity of the objective function, meaning that it guarantees a convergence bound of order $L/\sqrt{T}$ without prior knowledge of the Lipschitz coefficient $L$.

**Proposition 7.3.1** (AdaGrad-Norm for Lipschitz convex optimization)**.** *Assume that $f$ is $L$-Lipschitz for $\|\cdot\|_2$. Let $R, \gamma > 0$, $x_1 \in \mathcal{X}$ and*

$$x_{t+1} = \Pi_{\mathcal{X}}\left(x_t - \frac{\gamma}{\sqrt{\sum_{s=1}^t \|g_s\|_2^2}} g_t\right), \quad t \geqslant 1,$$

*where $g_t \in \partial f(x_t)$ for all $t \geqslant 1$.*

*(i) Then for all $T \geqslant 1$,*

$$\min_{1 \leqslant t \leqslant T} f(x_t) - f(x_*) \leqslant \left(\frac{\max_{1 \leqslant t \leqslant T} \|x_t - x_*\|_2^2}{2\gamma} + \gamma\right) \frac{L}{\sqrt{T}}.$$

*(ii) If $R \geqslant \max_{1 \leqslant t \leqslant T} \|x_t - x_*\|_2$, then $\gamma = R/\sqrt{2}$ yields*

$$\min_{1 \leqslant t \leqslant T} f(x_t) - f(x_*) \leqslant RL\sqrt{\frac{2}{T}}.$$

*Proof.* Combining the general regret bound for AdaGrad-Norm (Proposition 7.2.3) and Lemma 6.1.1 gives

$$\min_{1 \leqslant t \leqslant T} f(x_t) - f(x_*) \leqslant \frac{1}{T}\left(\frac{\max_{1 \leqslant t \leqslant T} \|x_t - x_*\|_2^2}{2\gamma} + \gamma\right)\sqrt{\sum_{t=1}^T \|g_t\|_2^2}.$$

The Lipschitz continuity of $f$ gives $\|g_t\|_2 \leqslant L$ for all $t \geqslant 1$. The result follows. □

The following result demonstrates that, in the context of stochastic convex optimization, AdaGrad-Norm is *adaptive* to the second-order moment of the gradient estimators: if the latter are bounded by $\sigma^2$, it achieves a convergence bound of order $\sigma/\sqrt{T}$ without prior knowledge of $\sigma$.

**Proposition 7.3.2** (AdaGrad-Norm for stochastic nonsmooth convex optimization)**.** *Let $R, \sigma > 0$, $x_1 \in \mathcal{X}$ and*

$$x_{t+1} = \Pi_{\mathcal{X}} \left( x_t - \frac{R}{\sqrt{2 \sum_{s=1}^{t} \|g_s\|_2^2}} g_t \right), \quad t \geqslant 1,$$

*where for each $t \geqslant 1$, almost-surely,*

$$\mathbb{E}\left[g_t \,|\, x_t\right] \in \partial f(x_t), \quad \mathbb{E}\left[ \|g_t\|_2^2 \,\middle|\, x_t \right] \leqslant \sigma^2, \quad \text{and} \quad \|x_t - x_*\|_2 \leqslant R.$$

*Then, for all $T \geqslant 1$,*

$$\mathbb{E}\left[ \min_{1 \leqslant t \leqslant T} f(x_t) - f(x_*) \right] \leqslant R\sigma \sqrt{\frac{2}{T}}.$$

*Proof.* Combining the general regret bound for AdaGrad-Norm (Proposition 7.2.3) and Lemma 6.4.2 gives

$$\mathbb{E}\left[ \min_{1 \leqslant t \leqslant T} f(x_t) - f(x_*) \right] \leqslant \frac{R}{T} \mathbb{E}\left[ \sqrt{2 \sum_{t=1}^{T} \|g_t\|_2^2} \right].$$

Then using Jensen's inequality with the concavity of the square root,

$$\begin{aligned} \mathbb{E}\left[ \sqrt{2 \sum_{t=1}^{T} \|g_t\|_2^2} \right] &\leqslant \sqrt{2 \cdot \mathbb{E}\left[ \sum_{t=1}^{T} \|g_t\|_2^2 \right]} \\ &\leqslant \frac{R}{T} \mathbb{E}\left[ \sqrt{2 \sum_{t=1}^{T} \mathbb{E}\left[ \|g_t\|_2^2 \,\middle|\, x_t \right]} \right] \\ &\leqslant R\sigma \sqrt{\frac{2}{T}}, \end{aligned}$$

hence the result. ☐

## 7.4 Application to smooth convex optimization

Let $L > 0$ and $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function that is $L$-smooth for $\|\cdot\|_2$ and that admits a global minimizer $x_*$:

$$f(x_*) = \min_{x \in \mathbb{R}^d} f(x).$$

**Lemma 7.4.1.** *For all $x \in \mathbb{R}^d$,*

$$\|\nabla f(x)\|_2^2 \leqslant 2L \left( f(x) - f(x_*) \right).$$

*Proof.* Let $x' \in \mathbb{R}^d$. The definition of smoothness gives

$$f(x') - f(x) - \left\langle \nabla f(x), x' - x \right\rangle \leqslant \frac{L}{2} \left\| x' - x \right\|_2^2 .$$

For $x' = x - \frac{1}{L}\nabla f(x)$, the above simplifies into

$$f(x') - f(x) + \frac{1}{2L} \left\| \nabla f(x) \right\|_2^2 \leqslant 0.$$

The result follows because $f(x_*) \leqslant f(x')$. □

The following statement proves that AdaGrad-Norm is *adaptive* to the smoothness of the objective function: it achieves a $L/T$ convergence bound without prior knowledge of the smoothness coefficient $L$.

**Proposition 7.4.2** (AdaGrad-Norm for smooth convex optimization)**.** *Assume $x_* \in \mathcal{X}$. Let $\gamma, R > 0$, $x_1 \in \mathcal{X}$ and*

$$x_{t+1} = \Pi_{\mathcal{X}} \left( x_t - \frac{\gamma}{\sqrt{\sum_{s=1}^{t} \left\| \nabla f(x_s) \right\|_2^2}} \nabla f(x_t) \right), \quad t \geqslant 1.$$

*(i) Let $T \geqslant 1$. Then,*

$$\min_{1 \leqslant t \leqslant T} f(x_t) - f(x_*) \leqslant \left( \frac{\max_{1 \leqslant t \leqslant T} \left\| x_t - x \right\|_2^2}{2\gamma} + \gamma \right)^2 \frac{2L}{T}.$$

*(ii) If $R \geqslant \max_{1 \leqslant t \leqslant T} \left\| x_t - x_* \right\|_2$, then $\gamma = R/\sqrt{2}$ yields,*

$$\min_{1 \leqslant t \leqslant T} f(x_t) - f(x_*) \leqslant \frac{4R^2 L}{T}.$$

*Proof.* Using general regret bound for AdaGrad-Norm (Proposition 7.2.3) and the convexity of $f$,

$$\begin{aligned}
\sum_{t=1}^{T} \left( f(x_t) - f(x_*) \right) &\leqslant \sum_{t=1}^{T} \left\langle \nabla f(x_t), x_t - x_* \right\rangle \\
&\leqslant \left( \frac{\max_{1 \leqslant t \leqslant T} \left\| x_t - x \right\|_2^2}{2\gamma} + \gamma \right) \sqrt{\sum_{t=1}^{T} \left\| \nabla f(x_t) \right\|_2^2} \\
&\leqslant \left( \frac{\max_{1 \leqslant t \leqslant T} \left\| x_t - x \right\|_2^2}{2\gamma} + \gamma \right) \sqrt{2L \sum_{t=1}^{T} \left( f(x_t) - f(x_*) \right)},
\end{aligned}$$

where we used Lemma 7.4.1 for the last inequality. Dividing by $\sqrt{\sum_{t=1}^{T}(f(x_t) - f(x_*))}$ and taking the square gives

$$\sum_{t=1}^{T}(f(x_t) - f(x_*)) \leqslant \left( \frac{\max\limits_{1 \leqslant t \leqslant T} \|x_t - x\|_2^2}{2\gamma} + \gamma \right)^2 2L.$$

The above left-hand side is bounded from below as

$$\sum_{t=1}^{T}(f(x_t) - f(x_*)) \geqslant \sum_{t=1}^{T}\left( \min_{1 \leqslant t \leqslant T} f(x_t) - f(x_*) \right) \geqslant T \left( \min_{1 \leqslant t \leqslant T} f(x_t) - f(x_*) \right),$$

hence the result. $\square$

# Chapter 8

# Monotone operators and fixed point iterations

This chapter gives a quick preview of monotone operators, fixed point iterations, and their relation to regret minimization. Monotone operators are a generalization of the gradient of a convex function, and allow to deal with various problems such as convex-concave saddle-points, convex games, finite two-player zero-sum games, etc.

Monotone operators can be defined with set-valued mappings, to generalize subdifferentials, but we restrict to single-valued mappings for simplicity.

Let $\mathcal{X} \subset \mathbb{R}^d$ be nonempty closed and convex. $I$ denotes the identity map on a set which will be clear from the context.

## 8.1 Monotone operators

Recall the minimization of a differentiable convex function $f : \mathbb{R}^d \to \mathbb{R}$. A global solution corresponds to a *zero* of the gradient: the corresponding notion below is the zero of the monotone operator. Regarding the constrained problem on $\mathcal{X}$, according to Proposition 1.2.11, the minimizer $x_*$ of $f$ on $\mathcal{X}$ is characterized by the following variational inequality:

$$\forall x \in \mathcal{X}, \quad \langle \nabla f(x_*), x - x_* \rangle \geqslant 0,$$

which below is extended to the notion of *strong solution* on $\mathcal{X}$ of a monotone operator.

**Definition 8.1.1.** Let $G : \mathcal{X} \to \mathbb{R}^d$ and $x_* \in \mathcal{X}$.

(i) $G$ is a *monotone operator* if:

$$\forall x, x' \in \mathcal{X}, \quad \big\langle G(x') - G(x), x' - x \big\rangle \geqslant 0.$$

(ii) $x_*$ is a zero of $G$ if $G(x_*) = 0$.

(iii) $x_*$ is a *strong solution* of $G$ on $\mathcal{X}$ if

$$\forall x \in \mathcal{X}, \quad \langle G(x_*), x_* - x \rangle \leqslant 0.$$

(iv) $x_*$ is a *weak solution* of $G$ on $\mathcal{X}$ if

$$\forall x \in \mathcal{X}, \quad \langle G(x), x_* - x \rangle \leqslant 0.$$

*Remark* 8.1.2. An advantage of the concept of weak solution is that such a solution always exist as soon as $\mathcal{X}$ is compact.

*Remark* 8.1.3. If $x_*$ is either a weak or a strong solution and belongs to the interior of $\mathcal{X}$, then it is a zero of the operator.

The following proposition proves that in the case of a continuous monotone operator, both above solution concepts are equivalent.

**Proposition 8.1.4.** *Let $G : \mathcal{X} \to \mathbb{R}^d$ a monotone operator and $x_* \in \mathcal{X}$.*

*(i) If $x_*$ is a strong solution of $G$ on $\mathcal{X}$, then it is a weak solution of $G$ on $\mathcal{X}$.*

*(ii) If $x_*$ is a weak solution of $G$ on $\mathcal{X}$ and $G$ is continuous, then it is a strong solution of $G$ on $\mathcal{X}$.*

*Proof.* (i) If $x_*$ is a strong solution, for all $x \in \mathcal{X}$,

$$\langle G(x), x_* - x \rangle \leqslant \langle G(x_*), x_* - x \rangle \leqslant 0,$$

where the first inequality holds by monotonicity of $G$.

(ii) Now assume that $x_*$ is a weak solution and that $G$ is continuous. Let us prove that $x_*$ is a strong solution. Let $x \in \mathcal{X}$ and for all $\lambda \in (0, 1)$ consider $x_\lambda = (1 - \lambda)x_* + \lambda x$, which belongs to $\mathcal{X}$ by convexity of the latter. Then because $x_*$ is a weak solution,

$$\langle G(x_\lambda), x_* - x_\lambda \rangle \leqslant 0,$$

where $x_* - x_\lambda$ rewrites as $\lambda(x_* - x)$. Then, dividing by $\lambda$ gives

$$\langle G(x_\lambda), x_* - x \rangle \leqslant 0.$$

Taking the limit as $\lambda \to 0^+$ gives, by continuity of $G$,

$$\langle G(x_*), x_* - x \rangle \leqslant 0.$$

$x_*$ is thus a strong solution. □

**Proposition 8.1.5.** *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable convex function. Then, $\nabla f : \mathbb{R}^d \to \mathbb{R}^d$ is a monotone operator.*

*Proof.* Let $x, x' \in \mathbb{R}^d$. By convexity,

$$f(x') \geqslant f(x) + \left\langle \nabla f(x), x' - x \right\rangle$$
$$f(x) \geqslant f(x') + \left\langle \nabla f(x'), x - x' \right\rangle .$$

Summing the above two inequalities gives

$$\left\langle \nabla f(x') - \nabla f(x), x' - x \right\rangle \geqslant 0,$$

hence the result. □

**Definition 8.1.6.** Let $m, n \geqslant 1$ be integers, $g : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$, $\mathcal{A} \subset \mathbb{R}^m$ and $\mathcal{B} \subset \mathbb{R}^n$ nonempty sets. A couple $(a_*, b_*) \in \mathcal{A} \times \mathcal{B}$ is a *saddle-point* of $g$ on $\mathcal{A} \times \mathcal{B}$ if

$$a_* \in \operatorname*{Arg\,min}_{a \in \mathcal{A}} g(a, b_*) \quad \text{and} \quad b_* \in \operatorname*{Arg\,max}_{b \in \mathcal{B}} g(a_*, b).$$

If moreover $\mathcal{A} = \mathbb{R}^m$ and $\mathcal{B} = \mathbb{R}^n$, the saddle-point is said to be *global*.

**Proposition 8.1.7.** *Let $m, n \geqslant 1$ be integer, $g : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ such that*

- *for all $b \in \mathbb{R}^n$, $g(\,\cdot\,, b)$ is convex and differentiable (denote $\nabla_a g(\,\cdot\,, b)$ its gradient);*
- *for all $a \in \mathbb{R}^m$, $g(a, \,\cdot\,)$ is concave and differentiable (denote $\nabla_b g(a, \,\cdot\,)$ its gradient).*

*Let $G : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^m \times \mathbb{R}^n$ be defined as*

$$G(a, b) = (\nabla_a g(a, b), -\nabla_b g(a, b)), \quad a \in \mathbb{R}^m, \; b \in \mathbb{R}^n,$$

*and $\mathcal{A} \subset \mathbb{R}^m$ and $\mathcal{B} \subset \mathbb{R}^n$ be nonempty closed convex sets.*

*(i) $G$ is a monotone operator.*

*(ii) $(a, b) \in \mathcal{A} \times \mathcal{B}$ is a saddle-point of $g$ on $\mathcal{A} \times \mathcal{B}$ if, and only if it is a strong solution of $G$ on $\mathcal{A} \times \mathcal{B}$.*

*(iii) $(a, b) \in \mathbb{R}^m \times \mathbb{R}^n$ is a global saddle-point of $g$ if, and only if, it is a zero of $G$.*

**Example 8.1.8** (KKT operator)**.** Let $p \geqslant 1$ be an integer, $f : \mathbb{R}^d \to \mathbb{R}$ a differentiable convex function, $A \in \mathbb{R}^{p \times d}$ and $b \in \mathbb{R}^p$. Consider the constrained optimization problem

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & Ax = b. \end{aligned}$$

Solving the above is equivalent to finding a saddle-point of the Lagrangian $L(x, w) = f(x) + w^\top (Ax - b)$, which in turn is equivalent to finding a zero of the associated KKT operator:

$$F(x, w) = \begin{pmatrix} \nabla f(x) + A^\top w \\ b - Ax \end{pmatrix}.$$

**Proposition 8.1.9.** *The Euclidean projection on a nonempty closed convex subset of $\mathbb{R}^d$ is a monotone operator.*

## 8.2 Bounded monotone operators

We consider bounded monotone operators, which is the natural extension of Lipschitz convex optimization. The simplest algorithm for this problem is the obvious extension of projected gradient descent, which is called the (projected) *forward step* algorithm:

$$x_{t+1} = \Pi_{\mathcal{X}} \left( x_t - \gamma_t G(x_t) \right), \quad t \geqslant 0.$$

For a class of UMD-based iterates that contains the above, Proposition 8.2.3 guarantees that an approximate weak solution can be obtained at speed $1/\sqrt{T}$. Then, in the special case of AdaGrad-Norm step-sizes, Proposition 7.2.4 provides a guarantee that is *adaptive* to the bound on the operator.

Let $G : \mathcal{X} \to \mathbb{R}^d$ be a monotone operator.

**Example 8.2.1** (Finite two-player zero-sum games)**.** Let $m, n \geqslant 1$ be integers, and $A \in \mathbb{R}^{m \times n}$. The solutions of the corresponding two-player zero-sum games are the solutions of the following constrained saddle-point problem:

$$\max_{a \in \Delta_m} \min_{b \in \Delta_n} \langle a, Ab \rangle .$$

The corresponding monotone operator $G(a, b) = (-Ab, A^\top a)$ is bounded on $\Delta_m \times \Delta_n$.

The following gives a connection between the search for an approximate weak solution and regret minimization.

**Lemma 8.2.2.** *Let $(x_t)_{t \geqslant 0}$ be a sequence in $\mathcal{X}$, $(\gamma_t)_{t \geqslant 0}$ a positive sequence, $T \geqslant 0$, and*

$$\bar{x}_T^{(\gamma)} = \frac{\sum_{t=0}^T \gamma_t x_t}{\sum_{t=0}^T \gamma_t}.$$

*Then for all $x \in \mathcal{X}$,*

$$\left\langle G(x), \bar{x}_T^{(\gamma)} - x \right\rangle \leqslant \left( \sum_{t=0}^T \gamma_t \right)^{-1} \sum_{t=0}^T \left\langle \gamma_t G(x_t), x_t - x \right\rangle .$$

*Proof.* For all $t \geqslant 0$, the monotonicity of $G$ gives

$$\gamma_t \left\langle G(x), x_t - x \right\rangle \leqslant \gamma_t \left\langle G(x_t), x_t - x \right\rangle .$$

Summing and dividing by $\sum_{t=0}^T \gamma_t$ gives the result. □

**Proposition 8.2.3** (Strict UMD iterates with time-dependent step-sizes and non-uniform averaging for variational inequalities with bounded monotone operators)**.** *Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$, $K, R > 0$, and $h$ a regularizer on $\mathcal{X}$. We assume that*

- $(x_t, y_t)_{t\geqslant 0}$ *is a sequence of strict UMD iterates associated with regularizer $h$ and dual increments $(-\gamma_t G(x_t))_{t\geqslant 0}$,*
- *$h$ is $K$-strongly convex for $\|\cdot\|$,*
- *for all $x \in \mathcal{X}$, $\|G(x)\|_* \leqslant L$.*

*(i) Let $T \geqslant 0$, $x \in \operatorname{dom} h$ and $\bar{x}_T^{(\gamma)} = \left(\sum_{t=0}^T \gamma_t\right)^{-1} \sum_{t=0}^T \gamma_t x_t$. Then,*

$$\forall x \in \operatorname{dom} h, \quad \left\langle G(x), \bar{x}_T^{(\gamma)} - x \right\rangle \leqslant \frac{D_h(x, x_0;\ y_0) + \frac{L^2}{2K} \sum_{t=0}^T \gamma_t^2}{\sum_{t=0}^T \gamma_t}.$$

*(ii) Let $\gamma > 0$. If $\gamma_t = \gamma\sqrt{2K}/(L\sqrt{t+1})$ for all $t \geqslant 0$, then*

$$\left\langle G(x), \bar{x}_T^{(\gamma)} - x \right\rangle \leqslant \frac{L}{\sqrt{2KT}} \left( \frac{D_h(x, x_0;\ y_0)}{\gamma} + \gamma(1 + \log(T+1)) \right).$$

*(iii) If $D_h(x, x_0;\ y_0) \leqslant R^2$, then $\gamma = R$ yields,*

$$\left\langle G(x), \bar{x}_T^{(\gamma)} - x \right\rangle \leqslant \frac{RL}{\sqrt{2KT}} (2 + \log(T+1)).$$

*Proof.* Use above Lemma 8.2.2 and perform a similar analysis as for Lipschitz convex optimization (Proposition 6.1.2). □

From the above general statement, we can of course derive corollaries for OMD and DA. Analogously to Proposition 6.1.6, DA with time-dependent parameters can be used instead to shave off the $\log T$ factor.

**Proposition 8.2.4** (AdaGrad-Norm for bounded monotone operators)**.** *Assume that $G$ is bounded by $L$ with respect to $\|\cdot\|_2$. Let $x_0 \in \mathcal{X}$, $\gamma > 0$,*

$$x_{t+1} = \Pi_{\mathcal{X}} \left( x_t - \frac{\gamma}{\sqrt{\sum_{s=0}^t \|G(x_s)\|_2^2}} G(x_t) \right), \quad t \geqslant 0.$$

*(i) Let $T \geqslant 0$ and $x \in \mathcal{X}$. Then,*

$$\langle G(x), \bar{x}_T - x \rangle \leqslant \left( \frac{\max_{0\leqslant t\leqslant T} \|x - x_t\|_2^2}{2\gamma} + \gamma \right) \sqrt{\frac{L}{T+1}},$$

*where $\bar{x}_T = \frac{1}{T+1} \sum_{t=0}^T x_t$.*

*(ii) Let* $R \geqslant \max_{0\leqslant t\leqslant T} \|x_t - x\|_2$*, then* $\gamma = R/\sqrt{2}$ *yields*

$$\langle G(x), \bar{x}_T - x\rangle \leqslant RL\sqrt{\frac{2}{T+1}}.$$

*Proof.* Combine above Lemma 8.2.2 with the regret bound for AdaGrad-Norm (Proposition 7.2.3). □

## 8.3 Lipschitz continuous monotone operators

We now consider Lipschitz continuous monotone operators, for which approximate weak solutions can be obtained at rate $1/T$ instead of $1/\sqrt{T}$. An important special case discussed below is two-player zero-sum games. The simplest method with such a guarantee is the *extragradient* algorithm which writes

$$\begin{aligned} w_t &= \Pi_{\mathcal{X}}(x_t - \gamma G(x_t)), \\ x_{t+1} &= \Pi_{\mathcal{X}}(x_t - \gamma G(w_t)), \quad t \geqslant 0, \end{aligned}$$

where $\gamma > 0$ is a step-size. We present and analyze below a very general a class of UMD-based iterates that extends extragradient.

**Example 8.3.1** (Finite two-player zero-sum games)**.** Let $m, n \geqslant 1$ be integers, and $A \in \mathbb{R}^{m\times n}$. Corresponding monotone operator $G(a, b) = (-Ab, A^\top a)$ is Lipschitz continuous on $\Delta_m \times \Delta_n$.

Let $G : \mathcal{X} \to \mathbb{R}^d$ be a monotone operator.

**Definition 8.3.2.** Let $h$ a regularizer on $\mathcal{X}$ and $\gamma > 0$. A sequence $((x_t, w_t, y_t, z_t))_{t\geqslant 0}$ in $(\mathbb{R}^d)^4$ is a sequence of *UMP iterates* associated with regularizer $h$, operator $G$ and step-size $\gamma$ if $((x_t, y_t))_{t\geqslant 0}$ is a sequence of strict UMD iterates associated with regularizer $h$ and dual iterates $(-\gamma G(w_t))_{t\geqslant 0}$ and for $t \geqslant 0$,

(i) $z_t \in \partial h(x_t)$,

(ii) $\forall x \in \mathcal{X},\ \langle z_t - y_t, x - x_t\rangle \geqslant 0$,

(iii) $w_t = \nabla h^*(z_t - \gamma G(x_t))$,

**Proposition 8.3.3** (UMP iterates for variational inequalities with Lipschitz continuous monotone operator)**.** *Let* $K, L > 0$*,* $\|\cdot\|$ *be a norm on* $\mathbb{R}^d$*,* $h$ *a regularizer on* $\mathcal{X}$*,* $((x_t, y_t, w_t, z_t))_{t\geqslant 0}$ *a sequence of UMP iterates associated with regularizer* $h$*, operator* $G$ *and step-size* $K/L$*. We assume that*

- $h$ *is* $K$*-strongly convex for* $\|\cdot\|$*,*

- *$G : \mathcal{X} \to \mathbb{R}^d$ is $L$-Lipschitz continuous for the following norms:*

$$\left\|G(x') - G(x)\right\|_* \leqslant L \left\|x' - x\right\|, \quad x, x' \in \mathcal{X}.$$

*Then for all $T \geqslant 0$,*

$$\forall x \in \operatorname{dom} h, \quad \langle G(x), \bar{w}_T - x\rangle \leqslant \frac{L \cdot D_h(x, x_0;\ y_0)}{K(T+1)},$$

*where $\bar{w}_T = \frac{1}{T+1}\sum_{t=0}^{T} w_t$.*

*Proof.* Let $x \in \operatorname{dom} h$ and $T \geqslant 0$. Lemma 8.2.2 gives:

$$\begin{aligned}
\frac{K(T+1)}{L}\langle G(x)|\bar{w}_T - x\rangle &\leqslant \sum_{t=0}^{T} \frac{K}{L}\langle G(w_t)|w_t - x\rangle \\
&= \sum_{t=0}^{T}\left(\frac{K}{L}\langle G(w_t)|x_{t+1} - x\rangle + \frac{K}{L}\langle G(w_t)|w_t - x_{t+1}\rangle\right) \\
&\leqslant D_h(x, x_0;\ y_0) + \sum_{t=0}^{T}\left(-D_h(x_{t+1}, x_t;\ y_t) + \frac{K}{L}\langle G(w_t)|w_t - x_{t+1}\rangle\right),
\end{aligned}$$

where the second inequality comes from applying Lemma 2.4.1 (because $(x_t, y_t)_{t\geqslant 1}$ is a sequence of strict UMD iterates). Let $t \geqslant 0$ and denote $\delta_t$ the content of the above last sum and we bound it as follows.

$$\begin{aligned}
\delta_t &= -D_h(x_{t+1}, x_t;\ y_t) - \frac{K}{L}\langle G(w_t)|x_{t+1} - w_t\rangle \\
&= -h(x_{t+1}) + h(x_t) + \langle y_t|x_{t+1} - x_t\rangle + \frac{K}{L}\langle (G(x_t) - G(w_t))|x_{t+1} - w_t\rangle \\
&\quad - \frac{K}{L}\langle G(x_t)|x_{t+1} - w_t\rangle .
\end{aligned} \tag{8.1}$$

Condition (ii) from the Definition 8.3.2:

$$\langle y_t|x_{t+1} - x_t\rangle \leqslant \langle z_t|x_{t+1} - x_t\rangle . \tag{8.2}$$

Besides, using basic inequality $\langle y|x\rangle \leqslant \frac{1}{2}\left\|y\right\|_*^2 + \frac{1}{2}\left\|x\right\|^2$, we can write:

$$\begin{aligned}
\frac{K}{L}\langle (G(x_t) - G(w_t))|x_{t+1} - w_t\rangle &\leqslant \frac{(K/L)^2}{2K}\left\|G(x_t) - G(w_t)\right\|_*^2 + \frac{K}{2}\left\|x_{t+1} - w_t\right\|^2 \\
&\leqslant \frac{(K/L)^2 L^2}{2K}\left\|x_t - w_t\right\|^2 + \frac{K}{2}\left\|x_{t+1} - w_t\right\|^2,
\end{aligned} \tag{8.3}$$

where we used the Lipschitz continuity of operator $G$. Injecting (8.2) and (8.3) into (8.1) and simplifying gives

$$\delta_t \leqslant -h(x_{t+1}) + h(x_t) + \langle z_t | w_t - x_t \rangle + \left\langle z_t - \frac{K}{L} G(x_t) \middle| x_{t+1} - w_t \right\rangle \\ + \frac{K}{2} \left\| w_t - x_t \right\|_*^2 + \frac{K}{2} \left\| x_{t+1} - w_t \right\|^2 . \quad (8.4)$$

First, we have $z_t \in \partial h(x_t)$ by condition (i) and $D_h(w_t, x_t;\ z_t)$ is well-defined. Besides, thanks to Proposition 1.4.7, condition (iii) is equivalent to $z_t - \frac{K}{L} G(x_t) \in \partial h(w_t)$, and $D_h(x_{t+1}, w_t;\ z_t - \frac{K}{L} G(x_t))$ is thus well-defined. We can make those two generalized Bregman divergences appear in the above right-hand side, which is consequently equal to:

$$\delta_t \leqslant \frac{K}{2} \left\| w_t - x_t \right\|^2 - D_h(w_t, x_t;\ z_t) + \frac{K}{2} \left\| x_{t+1} - w_t \right\|^2 \\ - D_h\left( x_{t+1}, w_t;\ z_t - \frac{K}{L} G(x_t) \right) .$$

Using the $K$-strong convexity of $h$ (Proposition 1.6.4), the above simplifies to $\delta_t \leqslant 0$. The result follows. □

The above general statement contains as special cases the following well-known algorithms.

**Corollary 8.3.4** (Mirror-Prox)**.** *Let $K, L > 0$, $\| \cdot \|$ be a norm on $\mathbb{R}^d$, $H$ a mirror map compatible with $\mathcal{X}$, $x_0 \in \mathcal{X} \cap \operatorname{int} \operatorname{dom} H$ and for $t \geqslant 0$,*

$$w_t = \operatorname*{arg\,max}_{x \in \mathcal{X}} \left\{ \left\langle \nabla H(x_t) - \frac{K}{L} G(x_t), x \right\rangle - H(x) \right\} \\ x_{t+1} = \operatorname*{arg\,max}_{x \in \mathcal{X}} \left\{ \left\langle \nabla H(x_t) - \frac{K}{L} G(w_t), x \right\rangle - H(x) \right\} .$$

*Assume that*

- *$H$ is $K$-strongly convex for $\| \cdot \|$,*
- *$G : \mathcal{X} \to \mathbb{R}^d$ is $L$-Lipschitz continuous for the following norms:*

$$\left\| G(x') - G(x) \right\|_* \leqslant L \left\| x' - x \right\| , \quad x, x' \in \mathcal{X}.$$

*Then for all $T \geqslant 0$,*

$$\forall x \in \mathcal{X} \cap \operatorname{dom} H, \quad \langle G(x), \bar{w}_T - x \rangle \leqslant \frac{L \cdot D_H(x, x_0)}{K(T+1)},$$

*where $\bar{w}_T = \frac{1}{T+1} \sum_{t=0}^{T} w_t$.*

**Corollary 8.3.5** (Extragradient). *Let $L > 0$, $x_0 \in \mathcal{X}$ and for $t \geqslant 0$,*

$$w_t = \Pi_{\mathcal{X}} \left( x_t - \frac{1}{L} G(x_t) \right)$$
$$x_{t+1} = \Pi_{\mathcal{X}} \left( x_t - \frac{1}{L} G(w_t) \right).$$

*Assume that $G$ is $L$-Lipschitz continuous for $\|\cdot\|_2$. Then, for all $T \geqslant 0$,*

$$\forall x \in \mathcal{X}, \quad \langle G(x), \bar{w}_T - x \rangle \leqslant \frac{L \|x - x_0\|_2^2}{2(T+1)},$$

*where $\bar{w}_T = \frac{1}{T+1} \sum_{t=0}^{T} w_t$.*

*Proof.* Apply Corollary 8.3.4 with the Euclidean mirror map. □

**Corollary 8.3.6** (Dual extrapolation). *Under the assumptions of Corollary 8.3.4, let $h = H + I_{\mathcal{X}}$, and for $t \geqslant 0$,*

$$x_t = \nabla h^* \left( y_0 - \frac{K}{L} \sum_{s=0}^{t-1} G(w_s) \right),$$
$$w_t = \operatorname*{arg\,max}_{x \in \mathcal{X}} \left\{ \left\langle \nabla H(x_t) - \frac{K}{L} G(x_t), x \right\rangle - H(x) \right\}.$$

*Then for all $T \geqslant 0$,*

$$\forall x \in \operatorname{dom} h, \quad \langle G(x), \bar{w}_T - x \rangle \leqslant \frac{L \cdot D_h(x, x_0;\ y_0)}{K(T+1)},$$

*where $\bar{w}_T = \frac{1}{T+1} \sum_{t=0}^{T} w_t$.*

## 8.4 Co-coercive operators and fixed point iterations

We now consider co-coercivity, which is a stronger notion than monotonicity. We establish the equivalence between being a zero of a co-coercive operator and being a fixed point of a corresponding *nonexpansive* map (meaning 1-Lipschitz continuous for the Euclidean norm). This subtopic is particularly important as it covers the construction and analysis of many optimization algorithms, such as the proximal gradient descent, ADMM, Douglas-Rachford splitting, Chambolle-Pock, and many more.

We present in Lemma 8.4.8 a connection between the search for an approximate solution of a co-coercive operator and regret minimization. We use this connection to recover in Proposition 8.4.9 the classical guarantee for the Krasnoselskii-Mann method, which is the workhorse fixed point iteration for nonexpansive maps. We further utilize this connection to define an AdaGrad-Norm-based method for fixed points which comes in Proposition 8.4.10 with an *adaptive* guarantee.

**Definition 8.4.1.** Let $L > 0$. A map $G : \mathcal{X} \to \mathbb{R}^d$ is a $L$-*co-coercive* operator if for all $x, x' \in \mathcal{X}$,

$$\langle G(x') - G(x), x' - x \rangle \geqslant \frac{1}{L} \left\| G(x') - G(x) \right\|_2^2 .$$

**Definition 8.4.2.** A map $F : \mathcal{X} \to \mathcal{X}$ is *nonexpansive* if it is 1-Lipschitz continuous for $\|\cdot\|_2$, in other words

$$\forall x, x' \in \mathcal{X}, \quad \left\| F(x') - F(x) \right\|_2 \leqslant \left\| x' - x \right\|_2 .$$

**Definition 8.4.3.** Let $F : \mathcal{X} \to \mathcal{X}$. $x_*$ is a fixed point of $F$ if $F(x_*) = x_*$.

**Proposition 8.4.4.** *Let $F : \mathcal{X} \to \mathcal{X}$, $x_* \in \mathcal{X}$ and $L > 0$. The following statements are equivalent.*

*(i) $x_*$ is a fixed point of $F$,*

*(ii) $x_*$ is a fixed point of $\frac{1}{L}F + (1 - \frac{1}{L})I$,*

*(iii) $x_*$ is a zero of $I - F$.*

*Proof.* Immediate. □

**Proposition 8.4.5.** *Let $L > 0$. A map $G : \mathcal{X} \to \mathbb{R}^d$ is a $L$-co-coercive operator if, and only if, $I - \frac{2}{L}G$ is nonexpansive.*

*Proof.* For $G : \mathcal{X} \to \mathbb{R}^d$, being $L$-co-coercive can be written, for all $x, x' \in \mathcal{X}$,

$$\frac{1}{L} \left\| G(x') - G(x) \right\|_2^2 - \langle G(x') - G(x), x' - x \rangle \leqslant 0.$$

Multiplying by $4/L$ and adding $\|x - x'\|_2^2$, the above equivalently rewrites: for all $x, x' \in \mathcal{X}$,

$$\left\| \frac{2}{L}G(x') - \frac{2}{L}G(x) \right\|_2^2 - 2 \left\langle \frac{2}{L}G(x') - \frac{2}{L}G(x), x' - x \right\rangle + \left\| x' - x \right\|_2^2 \leqslant \left\| x' - x \right\|_2^2 ,$$

which simplifies into

$$\left\| \left( \frac{2}{L}G - I \right)(x') - \left( \frac{2}{L}G - I \right)(x) \right\|_2^2 \leqslant \left\| x' - x \right\|_2^2 ,$$

in other words, $I - \frac{2}{L}G$ is nonexpansive. □

**Corollary 8.4.6.** *Let $L > 0$ and $F : \mathcal{X} \to \mathcal{X}$. Then, $\frac{1}{L}F + (1 - \frac{1}{L})I$ is nonexpansive if, and only if $\frac{1}{2}(I - F)$ is $L$-co-coercive.*

*Proof.* Follows from Proposition 8.4.5. □

**Proposition 8.4.7.** *Let $L > 0$ and $f : \mathbb{R}^d \to \mathbb{R}$ a differentiable convex function. Then, $f$ is $L$-smooth if, and only if, $\nabla f$ is a $L$-co-coercive operator.*

*Proof.* This characterization is part of Proposition 1.6.4. □

The easiest fixed points problems are those with a contraction map $F : \mathcal{X} \to \mathcal{X}$ (meaning $L$-Lipschitz continuous for $\|\cdot\|_2$ with $0 \leqslant L < 1$). Banach's fixed point theorem states that, from any initial point $x_0 \in \mathcal{X}$, iteration $x_{t+1} = F(x_t)$ ensures *geometric convergence* to the fixed point, which is necessarily unique.

When the map is only nonexpansive, the above iteration does not ensure convergence anymore. This motivates the Krasnoselskii-Mann iteration (see Proposition 8.4.9 below) which is a damped version: the next iterate $x_{t+1}$ is obtained by the convex combination of the image $F(x_t)$ with the current iterate $x_t$. Convergence to a fixed point is then ensured for e.g. constant coefficients (for the convex combination), but there no more guarantees on the distance to the fixed point. Instead, a vanishing bound on $\|F(x_t) - x_t\|$ is obtained, which is a weaker guarantee but somehow still measures how far $x_t$ is from being a fixed point.

**Lemma 8.4.8.** *Let $L > 0$, $F : \mathcal{X} \to \mathcal{X}$ such that $\frac{1}{L}F + (1 - \frac{1}{L})I$ is nonexpansive, $x_*$ a fixed point of $F$, $(x_t)_{t\geqslant 0}$ be a sequence in $\mathcal{X}$, $(\gamma_t)_{t\geqslant 0}$ a positive sequence, and $T \geqslant 0$. Then,*

$$\sum_{t=0}^{T} \gamma_t \|F(x_t) - x_t\|_2^2 \leqslant 2L \sum_{t=0}^{T} \langle \gamma_t (F(x_t) - x_t), x_* - x_t \rangle .$$

*Proof.* For $t \geqslant 0$, because $\frac{1}{2}(I - F)$ is $L$-co-coercive by Corollary 8.4.6,

$$\begin{aligned} \gamma_t \langle F(x_t) - x_t, x_* - x_t \rangle &= 2\gamma_t \left\langle \frac{F(x_t) - x_t}{2} - \frac{F(x_*) - x_*}{2}, x_* - x_t \right\rangle \\ &\geqslant \frac{2\gamma_t}{L} \left\| \frac{F(x_t) - x_t}{2} - \frac{F(x_*) - x_*}{2} \right\|_2^2 \\ &= \frac{\gamma_t}{2L} \|F(x_t) - x_t\|_2^2 . \end{aligned}$$

The result follows by summing the weighted average. □

**Proposition 8.4.9** (Krasnosleskii-Mann iterations)**.** *Let $F : \mathcal{X} \to \mathcal{X}$ be a nonexpansive map, $x_*$ a fixed point of $F$, $(\gamma_t)_{t\geqslant 0}$ a sequence in $(0, 1)$, $x_0 \in \mathcal{X}$ and for $t \geqslant 0$,*

$$x_{t+1} = (1 - \gamma_t)x_t + \gamma_t F(x_t).$$

*(i)* $(x_t)_{t\geqslant 0}$ *is a sequence of OGD iterates on* $\mathbb{R}^d$ *associated with dual increments* $(\gamma_t(F(x_t)-x_t))_{t\geqslant 0}$.

*(ii)* *Let* $T \geqslant 0$*. Then,*

$$\|F(x_T)-x_T\|_2 \leqslant \frac{\|x_0-x_*\|_2}{\sqrt{\sum_{t=0}^{T}\gamma_t(1-\gamma_t)}}.$$

*(iii)* *In particular, if* $\gamma_t = 1/2$ *for all* $t \geqslant 0$*,*

$$\|F(x_T)-x_T\|_2 \leqslant \frac{2\,\|x_0-x_*\|_2}{\sqrt{T+1}}.$$

*Proof.* For $t \geqslant 0$, the definition of the iterates rewrites

$$x_{t+1} = x_t + \gamma_t(F(x_t)-x_t),$$

which gives (i).

Combining the regret bound for OGD (Corollary 3.3.16) with Lemma 8.4.8 gives

$$\sum_{t=0}^{T}\gamma_t\,\|F(x_t)-x_t\|_2^2 \leqslant \|x_0-x_*\|_2^2 - \|x_{T+1}-x_*\|_2^2 + \sum_{t=0}^{T}\gamma_t^2\,\|F(x_t)-x_t\|_2^2\,.$$

Getting rid of term $\|x_{T+1}-x_*\|_2^2$ and rearranging gives

$$\sum_{t=0}^{T}\gamma_t(1-\gamma_t)\,\|F(x_t)-x_t\|_2^2 \leqslant \|x_0-x_*\|_2^2\,.$$

The result follows as soon as $(\|F(x_t)-x_t\|_2^2)_{t\geqslant 0}$ is nonincreasing, which we now prove. Let $t \geqslant 0$. Note that by definition of the iterates, $F(x_t)-x_{t+1} = (1-\gamma_t)(F(x_t)-x_t)$. Then,

$$\begin{aligned}
\|F(x_{t+1})-x_{t+1}\|_2 &= \|F(x_{t+1})-F(x_t)+F(x_t)-x_{t+1}\|_2 \\
&\leqslant \|F(x_{t+1})-F(x_t)\|_2 + \|F(x_t)-x_{t+1}\|_2 \\
&\leqslant \|x_{t+1}-x_t\|_2 + (1-\gamma_t)\,\|F(x_t)-x_t\|_2 \\
&= \gamma_t\,\|F(x_t)-x_t\|_2 + (1-\gamma_t)\,\|F(x_t)-x_t\|_2\,,
\end{aligned}$$

where we used the nonexpansiveness of $F$ for the second inequality. Hence (ii) and (iii) follows. □

We now transpose the AdaGrad-Norm algorithm to the problem of finding a fixed point of a map $F$, which may not be nonexpansive, but such that $\frac{1}{L}F + (1-\frac{1}{L})I$ is nonexpansive, for some $L > 0$. The adaptive character of AdaGrad allows to be adaptive to such a coefficient $L$, without prior knowledge.

**Proposition 8.4.10** (AdaGrad-Norm for fixed points)**.** *Let $L, R > 0$, $F : \mathcal{X} \to \mathcal{X}$ such that $\frac{1}{L}F + (1 - \frac{1}{L})I$ is nonexpansive, $x_*$ a fixed point of $F$, $x_0 \in \mathcal{X}$ such that $F(x_0) \neq x_0$, $\gamma > 0$ and for $t \geqslant 0$,*

$$x_{t+1} = x_t + \frac{\gamma}{\sqrt{\sum_{t=0}^{T} \|F(x_t) - x_t\|_2^2}}(F(x_t) - x_t).$$

*Let $T \geqslant 0$.*

*(i) Then,*

$$\min_{0 \leqslant t \leqslant T} \|F(x_t) - x_t\|_2 \leqslant \left( \frac{\max_{0 \leqslant t \leqslant T} \|x_t - x_*\|_2^2}{2\gamma} + \gamma \right) \frac{2L}{\sqrt{T}}.$$

*(ii) If $R \geqslant \max_{0 \leqslant t \leqslant T} \|x_t - x_*\|_2$, then $\gamma = R/\sqrt{2}$ yields*

$$\min_{0 \leqslant t \leqslant T} \|F(x_t) - x_t\|_2 \leqslant \frac{2\sqrt{2}RL}{\sqrt{T}}.$$

*Proof.* Combining Lemma 8.4.8 with the regret bound for AdaGrad-Norm (Proposition 7.2.3) gives

$$\sum_{t=0}^{T} \|F(x_t) - x_t\|_2^2 \leqslant 2L \left( \frac{\max_{0 \leqslant t \leqslant T} \|x_t - x_*\|_2^2}{2\gamma} + \gamma \right) \sqrt{\sum_{t=0}^{T} \|F(x_t) - x_t\|_2^2}.$$

Dividing by $\sqrt{\sum_{t=0}^{T} \|F(x_t) - x_t\|_2^2}$ gives

$$\sqrt{\sum_{t=0}^{T} \|F(x_t) - x_t\|_2^2} \leqslant 2L \left( \frac{\max_{0 \leqslant t \leqslant T} \|x_t - x_*\|_2^2}{2\gamma} + \gamma \right).$$

The result follows. ∎

# Chapter 9

# Regret learning in games

For $d \geqslant 1$, recall that the unit simplex in $\mathbb{R}^d$ is defined as

$$\Delta_d = \left\{ x \in \mathbb{R}_+^d,\ \sum_{i=1}^d x_i = 1 \right\},$$

which represents the set of probability distributions over the set $\{1, \dots, d\}$. Similarly, if $\mathcal{S}$ is a finite set, we denote

$$\Delta(\mathcal{S}) = \left\{ x \in \mathbb{R}_+^{\mathcal{S}},\ \sum_{s \in \mathcal{S}} x_s = 1 \right\}$$

the unit simplex in $\mathbb{R}^{\mathcal{S}}$.

## 9.1 Normal-form games

We quickly recall basic definitions about finite normal-form games.

**Definition 9.1.1.** A *finite (normal-form) game* is given by

- a finite number of players $N \geqslant 1$,
- and for each player $1 \leqslant k \leqslant N$,
  - a set of $\mathcal{S}^{(k)}$ of *pure strategies*,
  - a payoff function $g^{(k)} : \prod_{k'=1}^N \mathcal{S}^{(k')} \to \mathbb{R}$

The interpretation is the following. Each player $1 \leqslant k \leqslant N$ chooses a strategy $s^{(k)} \in \mathcal{S}^{(k)}$, independently of the other players, and gets payoff $g^{(k)}(s^{(1)}, \dots, s^{(N)})$. Each player aims at maximizing his payoff.

Let such a game be given.

**Definition 9.1.2.** A $N$-tuple $(s^{(1)}, \ldots, s^{(N)}) \in \mathcal{S}^{(1)} \times \cdots \times \mathcal{S}^{(N)}$ is a *Nash equilibrium in pure strategies* if for all $1 \leqslant k \leqslant N$,

$$s^{(k)} \in \operatorname*{Arg\,max}_{\tilde{s}^{(k)} \in \mathcal{S}^{(k)}} g^{(k)}(s^{(1)}, \ldots, s^{(k-1)}, \tilde{s}^{(k)}, s^{(k+1)}, \ldots, s^{(N)}).$$

**Example 9.1.3** (Penalty game)**.** A goalkeeper (Player 1) plays against the shooter (Player 2). Each player chooses left or right. The goalkeeper (resp. the shooter) wins (resp. looses) if their choices match. $N = 2$, $\mathcal{S}^{(1)} = \mathcal{S}^{(2)} = \{\mathrm{L}, \mathrm{R}\}$, and

$$g^{(1)}(s^{(1)}, s^{(2)}) = -g^{(2)}(s^{(1)}, s^{(2)}) = \begin{cases} 1 & \text{if } s^{(1)} = s^{(2)} \\ -1 & \text{otherwise.} \end{cases}$$

The game has no Nash equilibrium in pure strategies.

**Definition 9.1.4.** Let $1 \leqslant k \leqslant N$. A *mixed strategy* for Player $k$, is a probability distribution $a^{(k)} \in \Delta(\mathcal{S}^{(k)})$ over $\mathcal{S}^{(k)}$, where $\Delta(\mathcal{S}^{(k)})$ is the unit simplex in $\mathbb{R}^{\mathcal{S}^{(k)}}$.

**Definition 9.1.5.** The *mixed extension* of the game from Definition 9.1.1 is the game with sets of strategies $\Delta(\mathcal{S}^{(k)})$ $(1 \leqslant k \leqslant N)$ and payoff functions given for $1 \leqslant k \leqslant N$ and $a^{(1)} \in \Delta(\mathcal{S}^{(1)}), \ldots, a^{(N)} \in \Delta(\mathcal{S}^{(N)})$ by

$$g^{(k)}(a^{(1)}, \ldots, a^{(N)}) = \sum_{\substack{s^{(1)} \in \mathcal{S}^{(1)} \\ \vdots \\ s^{(N)} \in \mathcal{S}^{(N)}}} \left( \prod_{k'=1}^{N} a^{(k')}_{s^{(k')}} \right) g^{(k)}(s^{(1)}, \ldots, s^{(N)}).$$

Interpretation of the mixed extension is as follows. Each player $1 \leqslant k \leqslant N$ independently chooses a mixed strategy $a^{(k)} \in \Delta(\mathcal{S}^{(k)})$ and draws a pure strategy $s^{(k)} \sim a^{(k)}$. We then consider the expectation of each payoff $g^{(k)}(s^{(1)}, \ldots, s^{(N)})$ $(1 \leqslant k \leqslant N)$.

**Definition 9.1.6.** A $N$-tuple $(a^{(1)}, \ldots, a^{(N)}) \in \Delta(\mathcal{S}^{(1)}) \times \cdots \times \Delta(\mathcal{S}^{(N)})$ is a *Nash equilibrium in mixed strategies* if for all $1 \leqslant k \leqslant N$,

$$a^{(k)} \in \operatorname*{Arg\,max}_{\tilde{a}^{(k)} \in \Delta(\mathcal{S}^{(k)})} g^{(k)}(a^{(1)}, \ldots, a^{(k-1)}, \tilde{a}^{(k)}, a^{(k+1)}, \ldots, a^{(N)}).$$

**Theorem 9.1.7** (Nash, 1951)**.** *A finite game admits a Nash equilibrium in mixed strategies.*

**Example 9.1.8** (Penalty game)**.** In the penalty game, each player choosing left or right with probability 1/2 is a Nash equilibrium.

## 9.2 Two-player zero-sum games

Let $m, n \geqslant 1$ be integers. We focus on zero-sum two-player games which can be represented by a matrix $A \in \mathbb{R}^{m\times n}$. Denote $\mathcal{S}^{(1)} = \{1, \dots, m\}$ (resp. $\mathcal{S}^{(2)} = \{1, \dots, n\}$) the set of *pure strategies* of Player 1 (resp. Player 2). When Player 1 and Player 2 choose strategies $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$ respectively, Player 1 (resp. Player 2) obtains payoff $A_{ij}$ (resp. $-A_{ji}$). For mixed strategies $a \in \Delta_m$ and $b \in \Delta_n$, if $i \sim a$ and $j \sim b$ and sampled independently,

$$\mathbb{E}\left[A_{ij}\right] = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j A_{ij} = \langle a, Ab\rangle .$$

*Remark* 9.2.1. $(a_*, b_*) \in \Delta_m \times \Delta_n$ is a Nash equilibrium in mixed strategies if, and only if,

$$a_* \in \operatorname*{Arg\,max}_{a\in\Delta_m} \langle a, Ab_*\rangle \quad \text{and} \quad b_* \in \operatorname*{Arg\,min}_{b\in\Delta_n} \langle a_*, Ab\rangle .$$

**Definition 9.2.2.** The *duality gap* $\delta_A : \Delta_m \times \Delta_n \to \mathbb{R}_+$ is defined as

$$\delta_A(a, b) = \max_{a'\in\Delta_m} \left\langle a', Ab\right\rangle - \min_{b'\in\Delta_n} \left\langle a, Ab'\right\rangle, \quad a \in \Delta_m,\ b \in \Delta_n.$$

**Proposition 9.2.3.** $(a_*, b_*) \in \Delta_m \times \Delta_n$ *is a Nash equilibrium of the two-player zero-sum game associated with A if, and only if,* $\delta_A(a_*, b_*) = 0$.

*Proof.* It always holds that

$$\max_{a\in\Delta_m} \langle a, Ab_*\rangle \geqslant \langle a_*, Ab_*\rangle \geqslant \min_{b\in\Delta_n} \langle a_*, Ab\rangle .$$

Therefore, $\delta_A(a_*, b_*) = 0$ if, and only, if the above inequalities are equalities. The above first inequality being an equality is equivalent to

$$a_* \in \operatorname*{Arg\,max}_{a\in\Delta_m} \langle a, Ab_*\rangle ,$$

and similarly the second inequality being an equality is equivalent to

$$b_* \in \operatorname*{Arg\,min}_{b\in\Delta_n} \langle a_*, Ab\rangle .$$

Hence the result. □

*Remark* 9.2.4. The duality gap is easy to compute, as it rewrites

$$\delta_A(a, b) = \max_{1\leqslant i\leqslant m} (Ab)_i - \min_{1\leqslant j\leqslant n} (A^\top a)_j, \quad a \in \Delta_m,\ b \in \Delta_n.$$

It is therefore the quantity of choice to measure how far a couple $(a, b)$ is from being a Nash equilibrium.

**Theorem 9.2.5** (Von Neumann's minimax theorem)**.** *There exists a Nash equilibrium* $(a_*, b_*) \in \Delta_m \times \Delta_n$ *of the two-player zero-sum game associated with* $A$ *and*

$$\langle a_*, Ab_* \rangle = \max_{a\in\Delta_m} \min_{b\in\Delta_n} \langle a, Ab \rangle = \min_{b\in\Delta_n} \max_{a\in\Delta_m} \langle a, Ab \rangle .$$

There are several proofs of von Neumann's theorem. In the next section, we give one based on regret minimization.

## 9.3 Regret learning

The following statement demonstrates that the duality gap can be minimized if the game is played repeatedly, with each player minimizing its regret.

**Lemma 9.3.1.** *Let* $T \geqslant 1$*,* $a_1, \dots, a_T \in \Delta_m$*,* $b_1, \dots, b_T \in \Delta_n$*,* $\bar{a}_T = \frac{1}{T}\sum_{t=1}^T a_t$ *and* $\bar{b}_T = \frac{1}{T}\sum_{t=1}^T b_t$*. Then,*

$$\delta_T\left(\bar{a}_T, \bar{b}_T\right) = \frac{1}{T}\left(\max_{a\in\Delta_m}\sum_{t=1}^T \langle Ab_t, a - a_t\rangle + \max_{b\in\Delta_n}\sum_{t=1}^T \left\langle -A^\top a_t, b - b_t\right\rangle\right).$$

*Proof.*

$$\begin{aligned}
\delta_T\left(\bar{a}_T, \bar{b}_T\right) &= \max_{a\in\Delta_m}\left\langle a, A\left(\frac{1}{T}\sum_{t=1}^T b_t\right)\right\rangle - \min_{b\in\Delta_n}\left\langle \frac{1}{T}\sum_{t=1}^T a_t, Ab\right\rangle \\
&= \frac{1}{T}\left(\max_{a\in\Delta_m}\sum_{t=1}^T \langle a, Ab_t\rangle - \max_{b\in\Delta_n}\sum_{t=1}^T\left\langle -A^\top a_t, b\right\rangle\right) \\
&= \frac{1}{T}\left(\max_{a\in\Delta_m}\sum_{t=1}^T \langle a, Ab_t\rangle - \sum_{t=1}^T \langle a_t, Ab_t\rangle + \sum_{t=1}^T\left\langle A^\top a_t, b_t\right\rangle\right. \\
&\qquad \left. + \max_{b\in\Delta_n}\sum_{t=1}^T\left\langle -A^\top a_t, b\right\rangle\right),
\end{aligned}$$

hence the result. □

The following proposition uses the exponential weights algorithm to minimize each of the two regrets appearing the the above lemma. The guarantee on the duality gap is then an immediate adaptation from the regret bound for the exponential weights algorithm (Proposition 3.4.4).

**Proposition 9.3.2** (Exponential weights for two-player zero-sum games)**.** *For* $t \geqslant 1$*, let*

$$y_t = \eta_t \sum_{s=1}^{t-1} Ab_s, \qquad z_t = -\eta'_t \sum_{s=1}^{t-1} A^\top a_s,$$

*where*

$$\eta_t = \frac{\sqrt{\log m}}{\|A\|_\infty \sqrt{t}}, \qquad \eta'_t = \frac{\sqrt{\log n}}{\|A\|_\infty \sqrt{t}},$$

*and*

$$a_t = \left(\frac{\exp\left(y_{t,i}\right)}{\sum_{i'=1}^m \exp\left(y_{t,i'}\right)}\right)_{1\leqslant i\leqslant m}, \qquad b_t = \left(\frac{\exp\left(z_{t,j}\right)}{\sum_{j'=1}^n \exp\left(z_{t,j'}\right)}\right)_{1\leqslant j\leqslant n},$$

*For* $T \geqslant 1$*, let* $\bar{a}_T = \frac{1}{T}\sum_{t=1}^T a_t$ *and* $\bar{b}_T = \frac{1}{T}\sum_{t=1}^T b_t$*. Then,*

$$\delta_A\left(\bar{a}_T, \bar{b}_T\right) \leqslant \frac{2\|A\|_\infty\left(\sqrt{\log m}+\sqrt{\log n}\right)}{\sqrt{T}}.$$

*Proof.* For all $t \geqslant 1$, $\|Ab_t\|_\infty \leqslant \|A\|_\infty \|b_t\|_1 = \|A\|_\infty$ because $b_t \in \Delta_n$. Similarly, $\left\|A^\top a_t\right\|_\infty \leqslant \|A\|_\infty$. Therefore, similarly to Proposition 3.4.4, we obtain regret bounds

$$\max_{a\in\Delta_m}\sum_{t=1}^T \langle Ab_t, a - a_t\rangle \leqslant 2\|A\|_\infty \sqrt{T\log m}$$
$$\max_{b\in\Delta_n}\sum_{t=1}^T \left\langle -A^\top a_t, b - b_t\right\rangle \leqslant 2\|A\|_\infty \sqrt{T\log n}.$$

Summing and applying Lemma 9.3.1 gives the result. □

We can use the above result to now prove von Neumann's minmax theorem.

*Proof of Theorem 9.2.5.* Let us first prove that $\delta_A$ is continuous on $\Delta_m\times\Delta_n$.

$$(a,b) \mapsto \max_{a'\in\Delta_m}\left\langle a', Ab\right\rangle = \max_{a'\in\Delta_m}\left\langle A^\top a', b\right\rangle,$$

is defined on $\mathbb{R}^m \times \mathbb{R}^n$ and is convex as the point-wise maximum of linear functions. It is therefore continuous. Similarly, $(a,b)\mapsto \max_{b\in\Delta_n}\langle a, Ab'\rangle$ is continuous. Hence, $\delta_A$ is continuous.

For all $T \geqslant 1$, let $\bar{a}_T \in \Delta_m$ and $\bar{b}_T \in \Delta_n$ from Proposition 9.3.2. Because they belong to simplexes which are compact sets, there exists subsequences so that they converge to $a_* \in \Delta_m$ and $b_* \in \Delta_n$ respectively. Because the bound from Proposition 9.3.2 is vanishing as $T \to \infty$, by continuity of $\delta_A$, $\delta_A(a_*, b_*) = 0$, in other words, $(a_*, b_*)$ is a Nash equilibrium by Proposition 9.2.3.

Moreover, for all $(a,b) \in \Delta_m \times \Delta_n$, in addition to

$$\max_{a'\in\Delta_m}\left\langle a', Ab\right\rangle \geqslant \langle a, Ab\rangle \geqslant \min_{b'\in\Delta_n}\left\langle a, Ab'\right\rangle,$$

it is elementary to prove that

$$\max_{a'} \langle a', Ab \rangle \geqslant \min_{b' \in \Delta_n} \max_{a' \in \Delta_m} \langle a', Ab' \rangle \geqslant \max_{a' \in \Delta_m} \min_{b' \in \Delta_n} \langle a', Ab' \rangle \geqslant \min_{b' \in \Delta_n} \langle a, Ab' \rangle .$$

For $a = a_*$ and $b = b_*$ in particular, because $\delta_A(a_*, b_*) = 0$, the left-most and right-most quantities of both above displays are equal. Therefore, all inequalities are equalities and

$$\langle a_*, Ab_* \rangle = \min_{b' \in \Delta_n} \max_{a' \in \Delta_m} \langle a', Ab' \rangle = \max_{a' \in \Delta_m} \min_{b' \in \Delta_n} \langle a', Ab' \rangle .$$

□

**Proposition 9.3.3** (RM for two-player zero-sum games)**.** *Let $a_1 \in \Delta_m$ and $b_1 \in \Delta_n$. For $t \geqslant 1$, let*

$$a_t = \begin{cases} \dfrac{x_{t,+}}{\|x_{t,+}\|_1} & \text{if } x_{t,+} := \left( \displaystyle\sum_{s=1}^{t-1} (Ab_s - \langle a_s, Ab_s \rangle \mathbb{1}) \right)_+ \neq 0 \\ a_1 & \text{otherwise,} \end{cases}$$

$$b_t = \begin{cases} \dfrac{w_{t,+}}{\|w_{t,+}\|_1} & \text{if } w_{t,+} := \left( \displaystyle\sum_{s=1}^{t-1} \left( \langle a_s, Ab_s \rangle \mathbb{1} - A^\top a_s \right) \right)_+ \neq 0 \\ b_1 & \text{otherwise.} \end{cases}$$

*For all $T \geqslant 1$, let $\bar{a}_T = \frac{1}{T} \sum_{t=1}^T a_t$ and $\bar{b}_T = \frac{1}{T} \sum_{t=1}^T b_t$. Then,*

$$\delta_A \left( \bar{a}_T, \bar{b}_T \right) \leqslant (A_{\max} - A_{\min}) \frac{\sqrt{m} + \sqrt{n}}{\sqrt{T}}.$$

*Proof.* The general regret bound for RM from Proposition 5.5.3 gives

$$\max_{a \in \Delta_m} \sum_{t=1}^T \langle Ab_t, a - a_t \rangle \leqslant \sqrt{\sum_{t=1}^T \|Ab_t - \langle a_t, Ab_t \rangle \mathbb{1}\|_2^2}$$

$$\max_{b \in \Delta_n} \sum_{t=1}^T \left\langle -A^\top a_t, b - b_t \right\rangle \leqslant \sqrt{\sum_{t=1}^T \|\langle a_t, Ab_t \rangle \mathbb{1} - A^\top a_t\|_2^2}$$

For all $t \geqslant 1$, because $a_t \in \Delta_m$ and $b_t \in \Delta_n$,

$$\begin{aligned} \|Ab_t - \langle a_t, Ab_t \rangle \mathbb{1}\|_2^2 &= \sum_{i=1}^m ((Ab_t)_i - \langle a_t, Ab_t \rangle)^2 \\ &\leqslant \sum_{i=1}^m (A_{\max} - A_{\min})^2 \\ &= m \left( A_{\max} - A_{\min} \right)^2 , \end{aligned}$$

and similarly

$$\left\| \langle a_t, Ab_t \rangle \mathbb{1} - A^\top a_t \right\|_2^2 \leqslant n \left( A_{\max} - A_{\min} \right)^2 .$$

The result then follows by applying Lemma 9.3.1. □

**Proposition 9.3.4** (RM+ for two-player zero-sum games)**.** *Let $a_1 \in \Delta_m$, $b_1 \in \Delta_n$, $x_1 = 0$, $w_1 = 0$. For $t \geqslant 1$, let*

$$a_t = \begin{cases} \dfrac{x_t}{\|x_t\|_1} & \text{if } x_t \neq 0 \\ a_1 & \text{otherwise,} \end{cases} \qquad b_t = \begin{cases} \dfrac{w_t}{\|w_t\|_1} & \text{if } w_t \neq 0 \\ b_1 & \text{otherwise.} \end{cases}$$

$$x_{t+1} = \left( x_t + Ab_t - \langle a_t, Ab_t \rangle \mathbb{1} \right)_+ \qquad w_{t+1} = \left( w_t + \langle a_t, Ab_t \rangle \mathbb{1} - A^\top a_t \right)_+ .$$

*Then, the same guarantee as for RM (Proposition 9.3.3) holds.*

*Proof.* Identical to Proposition 9.3.3. □

## 9.4 Optimistic regret learning

Section 8.3 provided an approach for solving a class of problems containing two-player zero-sum games with fast convergence rate $1/T$. We here provide an alternative approach to obtain such a rate.

Informally, an optimistic variant of a regret minimization algorithm uses the last observed vector twice, the second time as a guess for the next vector. For instance, an optimistic exponential weights algorithm against $(u_t)_{t\geqslant 1}$ would write

$$x_t = \left( \frac{\exp\left( \eta \sum_{s=1}^{t-1} u_{s,i} + u_{t-1,i} \right)}{\sum_{i'=1}^{d} \exp\left( \eta \sum_{s=1}^{t-1} u_{s,i'} + u_{t-1,i'} \right)} \right)_{1\leqslant i \leqslant d} \qquad t \geqslant 1.$$

**Lemma 9.4.1** (Optimistic UMD)**.** *Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$, $K > 0$, $h : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ a regularizer which is $K$-strongly convex for $\|\cdot\|$, $(u_t)_{t\geqslant 1}$ a sequence in $\mathbb{R}^d$, and $((x_t, y_t))_{t\geqslant 1}$ be a sequence of strict UMD iterates associated with regularizer $h$ and dual increments $(2u_t - u_{t-1})_{t\geqslant 1}$ (with convention $u_0 = 0$). Then for all $T \geqslant 1$, $\alpha > 0$, and $x \in \operatorname{dom} h$,*

$$\sum_{t=1}^{T} \langle u_t, x - x_t \rangle \leqslant D_h(x, x_1;\ y_1) + \frac{\alpha}{2} \|x - x_T\|^2 + \frac{1}{2\alpha} \|u_1\|_*^2 + \frac{1}{2\alpha} \sum_{t=1}^{T-1} \|u_{t+1} - u_t\|_*^2 + \left( \frac{\alpha}{2} - \frac{K}{2} \right) \sum_{t=1}^{T-1} \|x_{t+1} - x_t\|^2 .$$

*Proof.*

$$\begin{aligned}\sum_{t=1}^{T}\langle u_t, x-x_t\rangle &= \langle u_1, x-x_1\rangle + \sum_{t=1}^{T-1}\langle u_{t+1}-2u_t+u_{t-1}, x-x_{t+1}\rangle \\ &\quad + \sum_{t=1}^{T-1}\langle 2u_t-u_{t-1}, x-x_{t+1}\rangle\,. \end{aligned}\tag{9.1}$$

The above last sum can be bounded from above using Lemma 2.4.1 and the strong convexity of $h$:

$$\begin{aligned}\sum_{t=1}^{T-1}\langle 2u_t-u_{t-1}, x-x_{t+1}\rangle &\leqslant D_h(x,x_1;\ y_1) - \sum_{t=1}^{T-1} D_h(x_{t+1},x_t;\ y_t) \\ &\leqslant D_h(x,x_1;\ y_1) - \frac{K}{2}\sum_{t=1}^{T-1}\left\|x_{t+1}-x_t\right\|^2\,. \end{aligned}\tag{9.2}$$

The other sum is bounded as follows.

$$\begin{aligned}\sum_{t=1}^{T-1}\langle u_{t+1}-2u_t+u_{t-1}, x-x_{t+1}\rangle &= \sum_{t=1}^{T-1}\langle u_{t+1}-u_t, x-x_{t+1}\rangle \\ &\quad - \sum_{t=1}^{T-1}\langle u_t-u_{t-1}, x-x_{t+1}\rangle \\ &= \sum_{t=1}^{T-1}\langle u_{t+1}-u_t, x-x_{t+1}\rangle - \sum_{t=1}^{T-2}\langle u_{t+1}-u_t, x-x_{t+2}\rangle - \langle u_1, x-x_2\rangle \\ &= \sum_{t=1}^{T-2}\langle u_{t+1}-u_t, x_{t+2}-x_{t+1}\rangle + \langle u_T-u_{T-1}, x-x_T\rangle - \langle u_1, x-x_2\rangle \\ &\leqslant \sum_{t=1}^{T-2}\left(\frac{\alpha}{2}\left\|x_{t+2}-x_{t+1}\right\|^2 + \frac{1}{2\alpha}\left\|u_{t+1}-u_t\right\|_*^2\right) \\ &\quad + \frac{1}{2\alpha}\left\|u_T-u_{T-1}\right\|_*^2 + \frac{\alpha}{2}\left\|x-x_T\right\|^2 - \langle u_1, x-x_2\rangle\,. \end{aligned}\tag{9.3}$$

Combining (9.1), (9.2) and (9.3) gives

$$\begin{aligned}\sum_{t=1}^{T}\langle u_t, x-x_t\rangle &\leqslant D_h(x,x_1;\ y_1) + \langle u_1, x_2-x_1\rangle + \frac{\alpha}{2}\sum_{t=1}^{T-2}\left\|x_{t+2}-x_{t+1}\right\|^2 \\ &\quad - \frac{K}{2}\sum_{t=1}^{T-1}\left\|x_{t+1}-x_t\right\|^2 + \frac{\alpha}{2}\left\|x-x_T\right\|^2 \\ &\quad + \frac{1}{2\alpha}\sum_{t=1}^{T-1}\left\|u_{t+1}-u_t\right\|_*^2\,. \end{aligned}$$

Regarding the second term in the above upper bound, we write

$$\langle u_1, x_2 - x_1 \rangle \leqslant \frac{1}{2\alpha} \|u_1\|_*^2 + \frac{\alpha}{2} \|x_2 - x_1\|^2 .$$

The result follows. □

**Proposition 9.4.2** (Optimistic exponential weights for two-player zero-sum games)**.** *Let $\eta > 0$ and for $t \geqslant 1$, let*

$$y_t = \eta \left( \sum_{s=1}^{t-1} A b_s + A b_{t-1} \right), \qquad z_t = -\eta \left( \sum_{s=1}^{t-1} A^\top a_s + A^\top a_{t-1} \right),$$

*with convention $a_0 = 0$ and $b_0 = 0$, and*

$$a_t = \left( \frac{\exp(y_{t,i})}{\sum_{i'=1}^m \exp(y_{t,i'})} \right)_{1 \leqslant i \leqslant m}, \qquad b_t = \left( \frac{\exp(z_{t,j})}{\sum_{j'=1}^m \exp(z_{t,j'})} \right)_{1 \leqslant j \leqslant n}.$$

*(i) $(a_t)_{t \geqslant 1}$ (resp. $(b_t)_{t \geqslant 1}$) corresponds to the exponential weights algorithm associated with dual increments $(\eta(2Ab_t - Ab_{t-1}))_{t \geqslant 1}$ (resp. $(\eta(A^\top a_{t-1}) - 2A^\top a_t)_{t \geqslant 1}$).*

*(ii) For $T \geqslant 1$, let $\bar{a}_T = \frac{1}{T}\sum_{t=1}^T a_t$ and $\bar{b}_T = \frac{1}{T}\sum_{t=1}^T b_t$. Then, if $\eta = 1/(2\|A\|_\infty)$,*

$$\delta_A \left( \bar{a}_T, \bar{b}_T \right) \leqslant \frac{2\|A\|_\infty (\log m + \log n + 2)}{T}.$$

*Proof.* It is easy to verify that for all $t \geqslant 1$, $y_t = \eta \sum_{s=1}^{t-1}(2Ab_t - Ab_{t-1})$. Similarly $z_t = -\eta \sum_{s=1}^{t-1}(2A^\top a_t - A^\top a_{t-1})$. Hence (i) holds.

Because $y_1 = 0$, if $h$ is the entropic regularizer on $\Delta_m$,

$$D_h(a, a_1;\ y_1) = h(a) - h(a_1) \leqslant \log m,$$

by Proposition 3.4.3. Similarly, if $\tilde{h}$ is the entropic regularizer on $\Delta_n$, because $z_1 = 0$,

$$D_{\tilde{h}}(b, b_1;\ z_1) \leqslant \log n.$$

Then, because the entropic regularizer is 1-strongly convex for $\|\cdot\|_1$ by

Proposition 2.2.8, applying Lemma 9.4.1 gives regret bounds

$$\begin{aligned}\eta \sum_{t=1}^{T} \langle Ab_t, a - a_t \rangle &\leqslant \log m + \frac{1}{4} \|a - a_T\|_1^2 + \|2\eta A b_1\|_\infty^2 \\ &\quad + \eta^2 \sum_{t=1}^{T-1} \|Ab_{t+1} - Ab_t\|_\infty^2 - \frac{1}{4} \sum_{t=1}^{T-1} \|a_{t+1} - a_t\|_1^2, \\ \eta \sum_{t=1}^{T} \left\langle -A^\top a_t, b - b_t \right\rangle &\leqslant \log n + \frac{1}{4} \|b - b_T\|_1^2 + \left\|2\eta A^\top a_1\right\|_\infty^2 \\ &\quad + \eta^2 \sum_{t=1}^{T-1} \left\|A^\top a_{t+1} - A^\top a_t\right\|_\infty^2 - \frac{1}{4} \sum_{t=1}^{T-1} \|b_{t+1} - b_t\|_1^2.\end{aligned}$$

Note that for all $t \geqslant 1$,

$$\begin{aligned}\|Ab_{t+1} - Ab_t\|_\infty^2 &\leqslant \|A\|_\infty^2 \|b_{t+1} - b_t\|_1^2, \\ \left\|A^\top a_{t+1} - A^\top a_t\right\|_\infty^2 &\leqslant \|A\|_\infty^2 \|a_{t+1} - a_t\|_1^2.\end{aligned}$$

Besides,

$$\|2Ab_1\|_\infty^2 \leqslant 4 \|A\|_\infty^2 \|b_1\|_1^2 = 4 \|A\|_\infty^2,$$

similarly $\left\|2A^\top a_1\right\|_\infty^2 \leqslant 4 \|A\|_\infty^2$, and

$$\|a - a_T\|_1^2 \leqslant (\|a\|_1 + \|a_T\|_1)^2 = 4,$$

and similarly $\|b - b_T\|_1^2 \leqslant 4$.

Dividing by $\eta$, summing, an applying Lemma 9.3.1 with $\alpha = 1/2$ gives

$$\begin{aligned}T \cdot \delta_A \left(\bar{a}_T, \bar{b}_T\right) &\leqslant \frac{\log m + \log n + 2}{\eta} + 8\eta \|A\|_\infty^2 \\ &\quad + \left(\eta \|A\|_\infty^2 - \frac{1}{4\eta}\right) \sum_{t=1}^{T-1} \left(\|a_{t+1} - a_t\|_1^2 + \|b_{t+1} - b_t\|_1^2\right).\end{aligned}$$

Because $\eta = 1/(2 \|A\|_\infty)$, the above last term is zero and the result follows. $\square$

# Bibliography


[Bub11] Sébastien Bubeck. *Introduction to Online Optimization: Lecture Notes.* Princeton University, 2011.

[CBL06] Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games.* Cambridge University Press, 2006.

[Haz16] Elad Hazan. Introduction to online convex optimization. *Foundations and Trends® in Optimization*, 2(3-4):157–325, 2016.

[JKM23] Anatoli Juditsky, Joon Kwon, and Éric Moulines. Unifying mirror descent and dual averaging. *Mathematical Programming*, (199):793–830, 2023.

[McM17] H. Brendan McMahan. A survey of algorithms and analysis for adaptive online learning. *The Journal of Machine Learning Research*, 18(1):3117–3166, 2017.

[Ora23] Francesco Orabona. A modern introduction to online learning. *arXiv:1912.13213*, 2023.

[Roc70] R. Tyrrell Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

[SS11] Shai Shalev-Shwartz. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2011.