Regret bounds

and their applications to online learning, optimization and games

Joon Kwon

March 1, 2024

Contents

Foreword

As of 2024, 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 discussions with Sylvain Sorin, Gilles Stoltz, Jaouad Mourtada, and the encouragements from Liliane Bel.

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 a set $A \subset \mathbb{R}^d$, $\operatorname{int} A$ and $\operatorname{cl} A$ 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. Besides, one can check that the dual norm is indeed a norm.

Proposition 1.1.4. *Let $\|\cdot\|$ be a norm in $\mathbb{R}^d$. Then, $\|\cdot\|_{**} = \|\cdot\|$.*

Example 1.1.5 (Common dual norms)**.** In $\mathbb{R}^d$, ℓ_2 is its own dual norm, ℓ_p and ℓ_q (with $p, q \geqslant 1$ such that $1/p + 1/q = 1$) are dual of each other, and ℓ_1 and ℓ_∞ are dual of each other. If A is a positive definite matrix, the dual norm of the associated Mahalanobis norm $x \mapsto \sqrt{\langle x, Ax \rangle}$ is the Mahalanobis norm associated with A^{-1}.

Remark 1.1.6. 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 Δ_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 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

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

being convex.

Example 1.2.7. 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 1.2.8 (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 1.2.9 (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 1.2.10. *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 1.2.11 (Lower semicontinuity)**.** A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous if for all $a \in \mathbb{R}^d$, the set $\left\{x \in \mathbb{R}^d,\ f(x) \leqslant a\right\}$ is closed.

Example 1.2.12. 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 1.2.13 (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*.

Proposition 1.2.14. *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 $x' \in \mathbb{R}^d$,

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

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

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

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 $0 \in \partial f(x_*)$. □

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 in $\operatorname{int}\operatorname{dom} f$ 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 $x, y \in \mathbb{R}^d$, $\langle y, x \rangle \leqslant f(x) + f^*(y)$.

The above definition is somewhat abstract, and it may be insightful to informally examine a simple example in dimension 1 by decomposing the transformation into simpler steps. Consider $f(x) = \frac{a}{2}(x - b)^2$ for some $a, b \neq 0$, which is a differentiable convex function with finite values on $\mathbb{R}$. Its derivative is given by $f'(x) = a(x - b)$, which is (strictly) increasing (as would be the case as soon as f is differentiable and strictly convex), and is a bijection from its domain to its range (from $\mathbb{R}$ to $\mathbb{R}$ in this case). Then, the inverse $(f')^{-1}$ is also an increasing function: $(f')^{-1}(y) = y/a + b$. We then consider the following primitive function:

$$f^*(y) = \int_0^y (f')^{-1} - \min_{x \in \mathbb{R}} f(x), \quad y \in \mathbb{R},$$

which can be proved to be an alternative definition of the Legendre–Fenchel transform in this special case (although the proof is somewhat involved). As the primitive of an increasing function, f^* is also convex. Because of the inverse relation between the derivatives, this transformation $f \mapsto f^*$ is involutional up a constant: $f^{**} = f + a$, $(a \in \mathbb{R})$. Then, it can be proved that $a = 0$, which can also be noticed graphically. An intuition that appears in this example is that the more f is curved, the less f^* is so, and vice-versa. The derivatives being inverses of each other can be interpreted as follows: f has slope y at point x if, and only if, f^* has slope x at point y. In higher dimension, this builds on the usual duality between points and hyperplanes. The Legendre–Fenchel transform indeed generalizes the above to higher dimensions, and for nondifferentiable functions. The neatest properties of e.g. Propositions 1.4.5 and 1.4.6 below are obtained for the class of proper lower semicontinuous convex functions. The derivative, which is a *function*, is then replaced by the subdifferential, which is a *correspondence*.

Proposition 1.4.4. *If* $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ *is a proper function,* f^* *is lower semicontinuous and convex.*

Theorem 1.4.5 (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 1.4.6. *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) $y \in \operatorname{Arg\,max}_{y' \in \mathbb{R}^d} \{\langle y', x \rangle - f^*(y')\}$.

Example 1.4.7 (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.

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

where for the second equality, we applied Fenchel's identity from property (iii) in Proposition 1.4.6. □

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 $x, x' \in \mathbb{R}^d$, $D_h(x', x) = \frac{1}{2} \|x' - x\|_2^2$.

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 $x, x' \in \mathbb{R}^d$, $|D_f(x', x)| \leqslant \frac{L}{2}\left\|x' - x\right\|^2$, 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 $D_f(x', x) \leqslant \frac{L}{2}\left\|x' - x\right\|_2^2$ $(x, x' \in \mathbb{R}^d)$.

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 $y \in \partial f(x)$, $D_f(x', x;\ y) \geqslant \frac{K}{2}\left\|x' - x\right\|^2$, 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

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

(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 $y, y' \in \mathbb{R}^d$,*

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

Proposition 1.6.6 (Second order characterization of strong convexity)**.** *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a twice differentiable function, $\|\cdot\|$ a norm in $\mathbb{R}^d$ and $K > 0$. Then, f is K-strongly convex with respect to $\|\cdot\|$ if, and only if,*

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

Proposition 1.6.7 (First and second order characterizations of smoothness)**.** *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a differentiable function, $\|\cdot\|$ a norm in $\mathbb{R}^d$ and $L > 0$.*

(i) f is L-smooth if, and only if, ∇f is L-Lipschitz with respect to $\|\cdot\|$.

(ii) Moreover, if f is twice differentiable, f is L-smooth with respect to $\|\cdot\|$ if, and only if,

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

Corollary 1.6.8. *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 $\|\cdot\|_2$.*

Proposition 1.6.9. *For $p \in (1,2)$, the squared ℓ_p norm $h_p : x \mapsto \frac{1}{2}\left\|x\right\|_p^2$ is $(p-1)$-strongly convex with respect to ℓ_p.*

Proposition 1.6.10. *The negative entropy h_{ent} is 1-strongly convex with respect to ℓ_1.*

Proposition 1.6.11. *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 $(x_t)_{t\geqslant 0}$ in $\mathcal{X}$ based on another sequence $(u_t)_{t\geqslant 0}$ in $\mathbb{R}^d$, where for each $t \geqslant 0$, 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

$$x_{t+1} = x_t + u_t, \quad t \geqslant 0,$$

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 $t \geqslant 0$ 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 $T \geqslant 0$,

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle = \frac{1}{2}\left\| x - x_0 \right\|_2^2 - \frac{1}{2}\left\| x - x_{T+1} \right\|_2^2 + \frac{1}{2}\sum_{t=0}^{T}\left\| u_t \right\|_2^2.$$

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. $\square$

For instance, the classical convergence guarantees about (stochastic) gradient descent in various settings, are consequences of the above identity. The quantity $\sum_{t=0}^{T} \langle u_t, x - x_t \rangle$ 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 $t \geqslant 0$, $x_{t+1} = x_t + u_t$ if, and only if,*

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

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:

$$x_{t+1} = x_0 + \sum_{s=0}^{t} u_s,$$

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_0). 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 $t \geqslant 0$, $x_{t+1} = \arg\min_{x \in \mathcal{X}} \left\| x_t + u_t - x \right\|_2^2$ if, and only if,*

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

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 $(x_t)_{t\geqslant 0}$ and $(y_t)_{t\geqslant 0}$ satisfy*

$$y_{t+1} = y_t + u_t \quad \textit{and} \quad x_{t+1} = \operatorname*{arg\,min}_{x\in\mathcal{X}} \left\| y_{t+1} - x \right\|_2^2, \qquad t \geqslant 0,$$

then for all $t \geqslant 0$,

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

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

Proposition 2.1.5. *For a positive definite matrix $A \in \mathbb{R}^{d\times d}$ and $t \geqslant 0$, $x_{t+1} = x_t + A^{-1}u_t$ 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 $x \in \mathbb{R}^d$,

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

where $\left\| x \right\|_A = \sqrt{\langle x, Ax \rangle}$ and $\left\| y \right\|_{A^{-1}} = \sqrt{\langle y, A^{-1}y \rangle}$.

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 $t \geqslant 0$,*

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

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* on $\mathcal{X}$ if it is strictly convex, lower semicontinuous, and if $\operatorname{cl}\operatorname{dom} h = \mathcal{X}$. Moreover, if $\operatorname{dom} h^* = \mathbb{R}^d$, then h is said to be an *regularizer* on $\mathcal{X}$.

Remark 2.2.2. A regularizer is proper (because $\mathcal{X}$ is nonempty), convex, and lower semicontinuous. In particular, Proposition 1.4.6 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 on $\mathcal{X}$.*

(i) If $\mathcal{X}$ is compact, then h is a regularizer on $\mathcal{X}$.

(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 on $\mathcal{X}$.

(iii) If h is strongly convex, then h is a regularizer on $\mathcal{X}$.

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) Because $\operatorname{cl}\operatorname{dom} h = \mathcal{X}$ by definition of a pre-regularizer, we have:

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

Besides, the function $x \mapsto \langle y, x\rangle - h(x)$ is upper semicontinuous and therefore, according to Proposition 1.2.14, attains a maximum on $\mathcal{X}$ because $\mathcal{X}$ is assumed to be compact. 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 1.4.6, $h^*(y) = \langle y, x\rangle - h(x) < +\infty$.

(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 1.6.11. Therefore, $h^*(y) < +\infty$.

□

Proposition 2.2.4 (Differentiability of h^*)**.** *Let h be a regularizer on $\mathcal{X}$. Then, h^* is differentiable on $\mathbb{R}^d$, ∇h^* takes values in $\operatorname{dom} h \subset \mathcal{X}$, and*

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

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 2.2.5 (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 on $\mathcal{X}$ and ∇h^* is the Euclidean projection onto $\mathcal{X}$, in other words:

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

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 2.2.6 (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 on Δ_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 on $\mathcal{X}$ and $(u_t)_{t \geqslant 0}$ a sequence in $\mathbb{R}^d$. A sequence $((x_t, y_t))_{t \geqslant 0}$ in $\mathbb{R}^d \times \mathbb{R}^d$ is a sequence of *UMD iterates* associated with regularizer h and *dual increments* $(u_t)_{t \geqslant 0}$ if for all $t \geqslant 0$,

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

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

Remark 2.3.2. By Proposition 1.4.6, property (ii) is equivalent to $y_t + u_t \in \partial h(x_{t+1})$.

Remark 2.3.3. For each $t \geqslant 0$, $x_t \in \operatorname{dom} h \subset \mathcal{X}$, 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 on $\mathcal{X}$ and $(u_t)_{t \geqslant 0}$ a sequence in $\mathbb{R}^d$. A sequence $((x_t, y_t))_{t \geqslant 0}$ in $\mathbb{R}^d \times \mathbb{R}^d$ is a sequence of *strict UMD iterates* associated with h and $(u_t)_{t \geqslant 0}$ if for all $t \geqslant 0$,

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

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

Proposition 2.3.5. *Let $((x_t, y_t))_{t\geqslant 0}$ be a sequence of strict UMD iterates defined as above. Then for all $t \geqslant 0$, $x_{t+1} = \nabla h^*(y_t + u_t)$ and thus $((x_t, y_t))_{t\geqslant 0}$ are UMD iterates.*

Proof. Let $t \geqslant 0$ 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 $x_{t+1} = \nabla h^*(y_t + u_t)$. □

Example 2.3.6 (Euclidean regularizer)**.** Denote $\Pi_{\mathcal{X}}$ the Euclidean projection onto $\mathcal{X}$ and consider the Euclidean regularizer on $\mathcal{X}$: $h = \frac{1}{2}\left\| \cdot \right\|_2^2 + I_{\mathcal{X}}$ and $x_0 \in \mathbb{R}^d$.

- If $x_{t+1} = \Pi_{\mathcal{X}}(x_t + u_t)$ for all $t \geqslant 0$, then $((x_t, x_t))_{t\geqslant 0}$ can be proved to be a sequence of strict UMD iterates.
- If $y_{t+1} = y_t + u_t$ and $x_{t+1} = \Pi_{\mathcal{X}}(y_{t+1})$ for all $t \geqslant 0$, then $((x_t, x_t))_{t\geqslant 0}$ can also be proved to be a sequence of strict UMD iterates.

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_0, y_0) such that $y_0 \in \partial h(x_0)$, there may be several possible strict UMD iterates. Here is a simple and explicit example. Consider $d = 1$, $\mathcal{X} = [0, 1]$, $h(x) = \frac{1}{2}x^2 + I_{\mathcal{X}}(x)$, $(x_0, y_0) = (1, 1)$ and $u_t = (-1)^t$ for $t \geqslant 0$. Then, one can verify that $((1, \frac{3+(-1)^t}{2}))_{t\geqslant 0}$ is a strict UMD sequence, and so is $((x_t, y_t))_{t\geqslant 0}$ 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 $(u)_{t\geqslant 0}$ 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 $y_0 \in \mathbb{R}^d$ and consider $x_0 := \nabla h^*(y_0)$, which satisfies $y_0 \in \partial h(x_0)$. Then, for $t \geqslant 0$, 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 $y_0 \in \mathbb{R}^d$, $\Pi_h(y_0)$ is the set of couples (x, y) satisfying

$$x = \nabla h^*(y_0), \quad y \in \partial h(x), \quad \text{and} \quad \forall x' \in \mathcal{X}, \quad \left\langle y_0 - y, x' - x \right\rangle \leqslant 0.$$

Then, one can verify that $((x_t, y_t))_{t\geqslant 0}$ is a strict UMD sequence associated with h and given sequence $(u_t)_{t\geqslant 0}$ if, and only if, $y_0 \in \partial h(x_0)$ and

$$(x_{t+1}, y_{t+1}) \in \Pi_h(y_t + u_t), \quad t \geqslant 0.$$

2.4 Regret bounds

Lemma 2.4.1 (UMD lemma)**.** *Let h be a regularizer on $\mathcal{X}$, $(u_t)_{t\geqslant 0}$ a sequence in $\mathbb{R}^d$, and $((x_t,y_t))_{t\geqslant 0}$ a sequence of UMD iterates associated with regularizer h and* dual increments *$(u_t)_{t\geqslant 0}$ and $x\in \operatorname{dom} h$. Consider notation*

$$D_t = D_h(x,x_t;\ y_t),\quad D'_t = D_h(x_{t+1},x_t;\ y_t),\quad D^*_t = D_{h^*}(y_t{+}u_t;\ y_t),\quad t\geqslant 0.$$

(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, $((x_t,y_t))_{t\geqslant 0}$ are strict UMD iterates, then for all $t\geqslant 0$,

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

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 $t\geqslant 0$,

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

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 $t\geqslant 0$,

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

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 $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ be a sequence of regularizers and $(u_t)_{t\geqslant 0}$ a sequence in $\mathbb{R}^d$. An associated sequence $((x_t,y_t))_{t\geqslant 0}$ in $\mathbb{R}^d\times\mathbb{R}^d$ of *UMD iterates* satisfy for all $t\in\mathbb{N}$,

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

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

Lemma 2.5.2 (UMD lemma with time-dependent regularizers). *Let $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ be a sequence of regularizers, $(u_t)_{t\geqslant 0}$ a sequence in $\mathbb{R}^d$, $((x_t, y_t))_{t\geqslant 1}$ associated UMD iterates, and $x \in \bigcap_{t\in\frac{1}{2}\mathbb{N}} \operatorname{dom} h_t$. For each $t \in \mathbb{N}$, 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 $t \in \mathbb{N}$,

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

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 $t \in \mathbb{N}$, 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 $t \in \mathbb{N}$, 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 $t \in \mathbb{N}$, $h_{t+1/2} \geqslant h_t$ 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 1.4.6. 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 the definition of the Legendre–Fenchel transform immediately implies $h_{t+1/2}^* \leqslant h_t^*$. Therefore,

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

□

Chapter 3

Online linear optimization

This chapter 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 $\{1, \ldots, d\}$, possibly at random. At step $t \geqslant 0$,

- 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 $\{1, \ldots, d\}$ 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 $(x_0, u_0, i_0, \ldots, x_{t-1}, u_{t-1}, i_{t-1})$. Similarly, the choice of u_t by Nature may depend on all past information including x_t: $(x_0, u_0, i_0, \ldots, x_{t-1}, u_{t-1}, i_{t-1}, x_t)$.

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

The Decision Maker wishes to maximize its cumulative payoff $\sum_{t=0}^{T} u_{t,i_t}$. 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 which holds *for all possible sequence* $(u_t)_{t\geqslant 0}$ chosen by Nature. Therefore, the guarantee must be relative to the sequence of payoff vectors $(u_t)_{t\geqslant 0}$. One possible type of guarantee is an upper bound on the *regret*, defined as

$$\max_{1\leqslant i\leqslant d} \sum_{t=0}^{T} u_{t,i} - \sum_{t=0}^{T} u_{t,i_t},$$

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 $t \geqslant 0$.

For a given sequence of payoff vectors $(u_t)_{t\geqslant 0}$, the above regret is a random variable (because for each $t \geqslant 0$, i_t is a random variable), and we are interested in analyzing its expectation. Using the law of total expectation, we can write

$$\begin{aligned}\mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=0}^{T} u_{t,i} - \sum_{t=0}^{T} u_{t,i_t}\right] &= \mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=0}^{T} u_{t,i} - \sum_{t=0}^{T} \mathbb{E}\left[u_{t,i_t}\middle| x_t\right]\right] \\ &= \mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=0}^{T} u_{t,i} - \sum_{t=0}^{T} \langle u_t, x_t\rangle\right]. \end{aligned} \tag{3.1}$$

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

$$\max_{1\leqslant i\leqslant d} \sum_{t=0}^{T} u_{t,i} = \max_{1\leqslant i\leqslant d} \left\langle \sum_{t=0}^{T} u_t, e_i \right\rangle = \max_{x\in\Delta_d} \sum_{t=0}^{T} \langle u_t, x_t\rangle .$$

Therefore, upper bounds on the quantity

$$\max_{x\in\Delta_d} \sum_{t=0}^{T} \langle u_t, x\rangle - \sum_{t=0}^{T} \langle u_t, x_t\rangle = \max_{x\in\Delta_d} \sum_{t=0}^{T} \langle u_t, x - x_t\rangle , \tag{3.2}$$

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 $(u_t)_{t\geqslant 0}$ 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 $(u_t)_{t\geqslant 0}$ 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 will 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 $t \geqslant 0$, 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:

$$i_{t+1} = \operatorname*{arg\,max}_{1\leqslant i\leqslant d} \sum_{s=1}^{t} u_{s,i}, \quad t \geqslant 0,$$

which can be equivalently written

$$x_{t+1} = \operatorname*{arg\,max}_{x\in\Delta_d} \left\langle \sum_{s=1}^{t} u_s, x \right\rangle, \quad t \geqslant 0. \tag{3.3}$$

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

$$u_0 = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}, \quad u_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \dots$$

yields for all $T \geqslant 1$,

$$\max_{1\leqslant i\leqslant d} \sum_{t=0}^{T} u_{t,i} - \sum_{t=0}^{T} u_{t,i_t} \geqslant \frac{T-1}{2} - \frac{1}{2}.$$

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 (4.2), 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 $t \geqslant 0$,

- 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

$$\max_{x \in \mathcal{X}} \sum_{t=0}^{T} \langle u_t, x - x_t \rangle .$$

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

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle ,$$

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 $t \geqslant 0$,

- 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

$$\max_{x \in \mathcal{X}} \sum_{t=0}^{T} \left(\ell_t(x_t) - \ell_t(x) \right) .$$

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 (4.2). It is a special case of UMD iterates where for all $t \geqslant 0$, the regularizers satisfy $h_{t+1/2} = h_{t+1}$ 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 $(h_t)_{t\geqslant 0}$ be a sequence of regularizers on $\mathcal{X}$, $(u_t)_{t\geqslant 0}$ a sequence in $\mathbb{R}^d$, and $y_0 \in \mathbb{R}^d$. The associated sequence $(x_t)_{t\geqslant 0}$ of *dual averaging* (DA) iterates is defined for $t \geqslant 0$ 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

$$\begin{aligned} x_{t+1} &= \operatorname*{arg\,max}_{x\in\mathbb{R}^d} \left\{ \left\langle y_0 + \sum_{s=0}^{t} u_t, x \right\rangle - h_t(x) \right\} \\ &= \operatorname*{arg\,max}_{x\in\mathcal{X}} \left\{ \langle u_t, x\rangle - D_h(x, x_t;\ y_t) \right\}. \end{aligned}$$

In the case $y_0 = 0$, the above second expression gives the following interpretation: x_{t+1} is the maximizer not of the past cumulative payoff function $x \mapsto \left\langle \sum_{s=0}^{t} u_s, x \right\rangle$ 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.

Example 3.2.3 (Lazy Mirror Descent)**.** Let $y_0 \in \mathbb{R}^d$ and denote $\Pi_{\mathcal{X}}$ the Euclidean projection onto $\mathcal{X}$. For a given sequence $(u_t)_{t\geqslant 0}$, $x_0 = \Pi_{\mathcal{X}}(y_0)$ and

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t + u_t), \quad t \geqslant 0$$

corresponds to dual averaging iterates with constant Euclidean regularizer $h = \frac{1}{2}\left\|\cdot\right\|_2^2 + I_{\mathcal{X}}$ (see Proposition 2.2.5 and Example 2.3.6).

Example 3.2.4 (Exponential weights algorithm)**.** For a given sequence $(u_t)_{t\geqslant 0}$, consider

$$x_{t+1} = \left(\frac{\exp\left(\sum_{s=0}^{t} u_{s,i}\right)}{\sum_{j=1}^{d} \exp\left(\sum_{s=0}^{t} u_{s,j}\right)} \right)_{1\leqslant j\leqslant d}.$$

These correspond to dual averaging iterates with constant entropic regularizer h_{ent} (see Proposition 2.2.6)

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 $(x_t)_{t\geqslant 0}$ and $(y_t)_{t\geqslant 0}$ be defined as in Definition 3.2.1, and $x \in \bigcap_{t\geqslant 0} \operatorname{dom} h_t$. Then,*

(i) *$((x_t, y_t))_{t\geqslant 0}$ is a sequence of UMD iterates associated with regularizers $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ and dual increments $(u_t)_{t\geqslant 0}$, where $h_{t+1/2} := h_{t+1}$ for all $t \geqslant 0$;*

(ii) *for all $T \geqslant 0$,*

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant D_0 + (h_{T+1}(x) - h_0(x)) + \sum_{t=0}^{T} D_t^*,$$

where

$$\begin{aligned} D_0 &= h_0(x) - h_0(x_0) - \langle y_0, x - x_0 \rangle \\ D_t^* &= h_{t+1}^*(y_t + u_t) - h_t^*(y_t) - \langle u_t, x_t \rangle, \quad t \geqslant 0; \end{aligned}$$

(iii) *if for $t \geqslant 0$, $h_{t+1} \geqslant h_t$ 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. □

Dual averaging with nonincreasing parameter Let h be a regularizer on $\mathcal{X}$, $(u_t)_{t\geqslant 0}$ a sequence in $\mathbb{R}^d$, $y_0 = 0$ and $(\eta_t)_{t\geqslant 0}$ a positive nonincreasing sequence. Consider iterates $(x_t)_{t\geqslant 0}$ defined for $t \geqslant 0$ as

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

Then, $((x_t, y_t))_{t\geqslant 0}$ are UMD iterates associated with dual increments $(u_t)_{t\geqslant 0}$ and regularizers $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ where

$$h_t(x) = \frac{h(x) - \min h}{\eta_t}, \quad t \geqslant 0, \tag{3.5}$$

and $h_{t+1/2} = h_{t+1}$ for $t \geqslant 1$. 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 $T \geqslant 0$ *and* $x \in \operatorname{dom} h$*,*

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{h(x) - \min h}{\eta_T} + \sum_{t=0}^{T} \frac{D_{h^*}(\eta_t(y_t + u_t), \eta_t y_t)}{\eta_t}.$$

(ii) Moreover, if h *is* K*-strongly convex for some norm* $\|\cdot\|$*, for all* $T \geqslant 0$*,*

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{h(x) - \min h}{\eta_T} + \frac{1}{2K} \sum_{t=0}^{T} \eta_t \|u_t\|_*^2 .$$

(iii) Moreover, if there exists $L > 0$ *such that* $\|u_t\|_* \leqslant L$*, then the choice* $\eta_t = \eta\sqrt{2K}/(L\sqrt{t+1})$ *for some* $\eta > 0$ *yields*

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant \left(\frac{h(x) - \min h}{\eta} + \eta \right) L \sqrt{\frac{T+1}{2K}}.$$

(iv) Moreover, if $\sup_{x \in \mathcal{X}} h(x) < +\infty$*, then* $\eta = \sqrt{\max_{\mathcal{X}} h - \min h}$ *yields*

$$\max_{x \in \mathcal{X}} \sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant L \sqrt{\frac{2(\max_{\mathcal{X}} h - \min h)(T+1)}{K}}.$$

Proof. (i) Applying Proposition 3.2.5, using $y_0 = 0$, and simplifying gives

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{h(x) - \min h}{\eta_{T+1}} + \sum_{t=0}^{T} (h_{t+1}^*(y_t + u_t) - h_t^*(y_t) - \langle u_t, x_t \rangle).$$

For a given $T \geqslant 0$, iterates $x_0, \ldots, x_T$ do not depend on η_{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), $h_{t+1}^*(y_t + u_t) \leqslant h_t^*(y_t + u_t)$, 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 $T \geqslant 0$, without prior knowledge of the latter: the algorithm and the guarantee are said to be *horizon-free*. In the case of a bounded domain such that $\sup_{\mathcal{X}} h < +\infty$, the bound (iv) does not depend on the comparison point x. When $\mathcal{X}$ is unbounded, or when $\sup_{\mathcal{X}} h = +\infty$, 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:

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

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 $t \geqslant 0$, $h_t = h_{t+1/2}$ 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 (ℓ^p norms)**.** For $p > 1$, $x \mapsto \frac{1}{2} \left\| x \right\|_p^2$ is a mirror map compatible with all nonempty closed convex sets.

Example 3.3.3 (Generalized negative entropy)**.** The generalized negative entropy on the closed positive orthant, defined as

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

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 3.3.4 (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 3.3.5. *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* $y \in \mathbb{R}^d$, $\nabla F^*(\nabla F(x)) = x$ *and* $\nabla F(\nabla F^*(y)) = y$.

Proof. Let $y_t \in \mathbb{R}^n$. By property (ii) from Definition 3.3.1, there exists $x_0 \in \mathcal{D}_H$ such that $\nabla H(x_0) = y$. Therefore, function $\varphi_y : x \mapsto \langle y|x\rangle - H(x)$ is differentiable at x_0 and $\nabla \varphi_y(x_0) = 0$. Moreover, φ_y is strictly concave as a consequence of property (i) from Definition 3.3.1. Therefore, x_0 is the unique maximizer of φ_y and:

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

which proves property (i).

Besides, we have

$$x_0 \in \partial H^*(y) \quad \iff \quad y = \nabla H(x_0) \quad \iff \quad x_0 \text{ minimizer of } \phi_y, \tag{3.6}$$

where the first equivalence comes from Proposition 1.4.6. Point x_0 being the unique maximizer of φ_y, we have that $\partial H^*(y)$ is a singleton. In other words, H^* is differentiable in y and

$$\nabla H^*(y) = x_0 \in \mathcal{D}_H. \tag{3.7}$$

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, $x_0 = x$ by uniqueness of x_0. 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 3.3.6 (OMD iteration)**.** *Let* H *be a mirror map compatible with* $\mathcal{X}$, $x \in \mathcal{D}_H$, $u \in \mathbb{R}^d$ *and consider*

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

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

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

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 3.3.7. A sequence $(H_t)_{t \geqslant 0}$ of mirror maps has *nondecreasing domains* if $\operatorname{dom} H_t \subset \operatorname{dom} H_{t+1}$ for all $t \geqslant 0$.

Definition 3.3.8. Let $(H_t)_{t \geqslant 0}$ be a sequence of mirror maps compatible with $\mathcal{X}$ with nondecreasing domains and $(u_t)_{t \geqslant 0}$ a sequence in $\mathbb{R}^d$. A sequence $(x_t)_{t \geqslant 0}$ in $\mathbb{R}^d$ is an sequence of *online mirror descent (OMD)* iterates on $\mathcal{X}$ associated with mirror maps $(H_t)_{t \geqslant 0}$ and *dual increments* $(u_t)_{t \geqslant 0}$ if $x_0 \in \mathcal{X} \cap \operatorname{int} \operatorname{dom} H_0$ and

$$x_{t+1} = \operatorname*{arg\,max}_{x \in \mathcal{X}} \left\{ \langle u_t, x \rangle - D_{H_t}(x, x_t) \right\}, \quad t \geqslant 0.$$

Remark 3.3.9. The above iterates are well-defined. Indeed, we first note that $x \in \mathcal{D}_{H_0}$, and then by induction, as soon as $x_t \in \mathcal{D}_{H_t}$, Proposition 3.3.6 ensures that x_{t+1} is well-defined an unique, and that it belongs to $\mathcal{D}_{H_t}$. Because the mirror maps have nondecreasing domains, x_{t+1} also belongs to $\mathcal{D}_{H_{t+1}}$.

Example 3.3.10 (Online (projected) gradient descent)**.** Denote $\Pi_{\mathcal{X}}$ the Euclidean projection onto $\mathcal{X}$. Let $y_0 \in \mathbb{R}^d$ and a sequence $(u_t)_{t \geqslant 0}$ in $\mathbb{R}^d$. Then, $x_0 = \Pi_{\mathcal{X}}(y_0)$ and

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t + u_t), \quad t \geqslant 0,$$

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

Example 3.3.11 (Exponential weights algorithm)**.** The exponential weights algorithm from Example 3.2.4 gives OMD iterates on the simplex Δ_d associated with generalized entropic mirror map from Example 3.3.3.

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 on $\mathcal{X}$ 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 $(H_t)_{t\geqslant 0}$ be a sequence of mirror maps compatible with $\mathcal{X}$ with nondecreasing domains, $(u_t)_{t\geqslant 0}$ a sequence in $\mathbb{R}^d$, $(x_t)_{t\geqslant 0}$ associated OMD iterates in $\mathcal{X}$ and $x \in \mathcal{X} \cap \operatorname{dom} H_0$. Then,*

(i) $((x_t, \nabla H_t(x_t)))_{t\geqslant 0}$ is a sequence of UMD iterates associated with regularizers $(h_t)_{t\geqslant 0}$ defined as

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

and dual increments $(u_t)_{t\geqslant 0}$;

(ii) for all $T \geqslant 0$,

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant D_{H_0}(x, x_0) - D_{H_{T+1}}(x, x_{T+1}) + \sum_{t=0}^{T} (\tilde{D}^{\Delta}_{t+1/2} + \tilde{D}^*_t),$$

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 $t \geqslant 0$, 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) $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ are indeed regularizers thanks to Proposition 3.3.12. Let us prove that $((x_t, \nabla H_t(x_t)))_{t\geqslant 0}$ satisfy the definition of UMD iterates. Let $t \geqslant 0$ and $x' \in \mathbb{R}^d$. By definition of h_t, $h_t \geqslant H_t$. Moreover, $h_t(x_t) = H(x_t)$ 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 3.3.6 and the definition of $h_{t+1/2}$,

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

which proves the second condition, and $((x_t, \nabla H_t(x_t)))_{t\geqslant 0}$ is indeed a sequence of UMD iterates associated with $(H_t)_{t\geqslant 0}$ and $(u_t)_{t\geqslant 0}$.

(ii) For all $t \in \frac{1}{2}\mathbb{N}$, the definition of h_t imply that $\operatorname{dom} h_t = \mathcal{X} \cap \operatorname{dom} H_t$. Because $(H_t)_{t\geqslant 0}$ has nondecreasing domains,

$$\begin{aligned} x \in \mathcal{X} \cap \operatorname{dom} H_0 &= \mathcal{X} \cap \bigcap_{t\in\mathbb{N}} \operatorname{dom} H_t = \bigcap_{t\in\mathbb{N}} (\mathcal{X} \cap \operatorname{dom} H_t) \\ &= \bigcap_{t\in\mathbb{N}} \operatorname{dom} h_t = \bigcap_{t\in\frac{1}{2}\mathbb{N}} \operatorname{dom} h_t. \end{aligned}$$

Therefore, Lemma 2.5.2 can be applied with x and with notation therein, we get for all $T \geqslant 0$,

$$\sum_{t=0}^{T} \langle u_t, x - x_t\rangle = D_0 - D_{T+1} + \sum_{t=0}^{T} (D^{\Delta}_{t+1/2} + D^*_t).$$

First, because $x, x_0 \in \mathcal{X} \cap \operatorname{dom} H_0$,

$$\begin{aligned} D_0 &= h_0(x) - h_0(x_0) - \langle \nabla H_0(x_0), x - x_0\rangle \\ &= H_0(x) - H_0(x_0) - \langle \nabla H_0(x_0), x - x_0\rangle \\ &= D_{H_0}(x, x_0). \end{aligned}$$

Similarly, x belongs to $\mathcal{X} \cap \operatorname{dom} H_{T+1}$ and so does x_{T+1} as a consequence of $(x_t)_{t\geqslant 0}$ 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 $t \geqslant 0$, denote $y_t = \nabla H_t(x_t)$, then Proposition 3.3.6 ensures that

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

in other words x_{t+1} is a minimizer of a differentiable convex function over $\mathcal{X}$. Therefore, Proposition 1.2.10 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 3.3.5, 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 1.4.6. The latter proposition also gives a characterization of both above gradients as maximizers of concave functions, which yields

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

Hence

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

□

OMD with nonincreasing step-sizes An important special case of OMD iterates is the following. Let H be a mirror map compatible with $\mathcal{X}$, $(\gamma_t)_{t\geqslant 0}$ a positive and nonincreasing sequence, and $(u_t)_{t\geqslant 0}$ a sequence in $\mathbb{R}^d$. Then, associated OMD iterates on $\mathcal{X}$ with time-dependent *step-sizes* (or *learning rate*) are defined as

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

This corresponds to OMD iterates from Definition 3.3.8 with time-dependent mirror maps $(H_t)_{t\geqslant 0}$ defined as $H_t = \gamma_t^{-1} H$ for all $t \geqslant 0$, which is indeed a sequence of mirror maps compatible with $\mathcal{X}$ with nondecreasing domains.

Corollary 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 $T \geqslant 0$ *and* $x \in \mathcal{X} \cap \operatorname{dom} H$*,*

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{\max\limits_{0 \leqslant t \leqslant T} D_H(x, x_t)}{\gamma_T} + \sum_{t=0}^{T} \tilde{D}_t^*,$$

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* $T \geqslant 0$*,*

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{\max\limits_{0 \leqslant t \leqslant T} D_H(x, x_t)}{\gamma_T} + \frac{1}{2K} \sum_{t=0}^{T} \gamma_t \|u_t\|_*^2 .$$

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

$$\max_{x \in \mathcal{X}} \sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant RL\sqrt{\frac{2(T+1)}{K}}.$$

Proof. (i) Applying Proposition 3.3.13,

$$\begin{aligned}
\sum_{t=0}^{T} \langle u_t, x - x_t \rangle &= \frac{D_H(x, x_0)}{\gamma_0} - \frac{D_H(x, x_{T+1})}{\gamma_{T+1}} \\
&\qquad + \sum_{t=0}^{T} \left(\frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \right) D_H(x, x_{t+1}) + \sum_{t=0}^{T} \tilde{D}_t^* \\
&= \frac{D_H(x, x_0)}{\gamma_0} + \sum_{t=0}^{T-1} \left(\frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \right) D_H(x, x_{t+1}) \\
&\qquad - \frac{D_H(x, x_{T+1})}{\gamma_T} + \sum_{t=0}^{T} \tilde{D}_t^* \\
&\leqslant \left(\max_{1 \leqslant t \leqslant T} D_H(x, x_t) \right) \left(\frac{1}{\gamma_0} + \sum_{t=0}^{T-1} \left(\frac{1}{\gamma_{t+1}} - \frac{1}{\gamma_t} \right) \right) + \sum_{t=0}^{T} \tilde{D}_t^* \\
&= \frac{\max\limits_{1 \leqslant t \leqslant T} D_H(x, x_t)}{\gamma_T} + \sum_{t=0}^{T} \tilde{D}_t^*.
\end{aligned}$$

Besides, for all $t \geqslant 0$ 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 $\max_{t \geqslant 0} D_H(x, x_t)$, for instance in the case where $\mathcal{X}$ is bounded, there is no known sublinear regret bound for OMD. This is contrast with DA which, with the right nonincreasing parameters from Proposition 3.2.5, guarantees a regret bound that grows as $\sqrt{T}$ with time, as soon as a bound on the norms of vectors $(u_t)_{t \geqslant 0}$ is known.

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

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t + u_t), \quad t \geqslant 0.$$

Then, for all $T \geqslant 0$ and $x \in \mathcal{X}$,

$$\sum_{t=0}^{T} \langle u_t, x - x_t \rangle \leqslant \frac{\max_{0 \leqslant t \leqslant T} \|x - x_t\|_2^2}{2\gamma_T} + \sum_{t=0}^{T} \frac{\gamma_t \|u_t\|_2^2}{2}.$$

Proof. Combine Corollaries 3.3.16 and 1.6.8. □

3.4 Finite action set

We recall the regret minimization problem where the Decision Maker chooses at each step an element in $\{1, \ldots, d\}$ possibly at random, described in Section 3.1. At step $t \geqslant 0$,

- 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 $\{1, \ldots, d\}$ 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 Δ_d, as formalized in the following lemma.

Lemma 3.4.1. *For all $T \geqslant 0$,*

$$\mathbb{E}\left[\max_{1\leqslant i\leqslant d}\sum_{t=0}^{T}(u_{t,i}-u_{t,i_t})\right]=\mathbb{E}\left[\max_{x\in\Delta_d}\sum_{t=0}^{T}\langle u_t, x-x_t\rangle\right].$$

Therefore, any upper bound on $\max_{x\in\Delta_d}\sum_{t=0}^{T}\langle u_t, x-x_t\rangle$ will also be an upper bound on $\mathbb{E}\left[\max_{1\leqslant i\leqslant d}\sum_{t=0}^{T}(u_{t,i}-u_{t,i_t})\right]$. A common assumption in this setting is that there is a known bound on the ℓ_∞ norm of the payoff vectors. We establish the important $\sqrt{T\log d}$ regret bound in this case.

Definition 3.4.2 (Exponential weights algorithm)**.** Let $(u_t)_{t\geqslant 0}$ be a sequence in $\mathbb{R}^d$ and $(\eta_t)_{t\geqslant 0}$ a positive sequence. The associated iterates of *exponential weights algorithm* (EW) iterates are defined in the simplex Δ_d as

$$x_t=\left(\frac{\exp\left(\eta_t\sum_{s=0}^{t-1}u_{s,i}\right)}{\sum_{j=1}^{d}\exp\left(\eta_t\sum_{s=0}^{t-1}u_{s,j}\right)}\right)_{1\leqslant i\leqslant d},\quad t\geqslant 0.$$

We recall the definition of the entropic regularizer on the simplex:

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

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_{ent} *is* 1*-strongly convex for* $\|\cdot\|_1$.

Proof. (iii) h_{ent} being convex, its maximum on Δ_d is attained at one of the extreme points. At each extreme point, the value of h_{ent} is zero. Therefore, $\max_{\Delta_d}h_{\mathrm{ent}}=0$. As for the minimum, h_{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 Δ_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 (4.3) is proved. F is therefore 1-strongly convex with respect to $\|\cdot\|_1$ and so is h_{ent}. □

Proposition 3.4.4 (Regret bound for EW). *Let $(x_t)_{t\geqslant 0}$ be defined as in Definition 3.4.2. Then,*

(i) $(x_t)_{t\geqslant 0}$ is a sequence of DA iterates associated with regularizer h_{ent}, parameters $(\eta_t)_{t\geqslant 0}$ and dual increments $(u_t)_{t\geqslant 0}$.

(ii) Moreover, if $(\eta_t)_{t\geqslant 0}$ is nonincreasing, for all $T \geqslant 0$,

$$\mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=0}^T (u_{t,i} - u_{t,i_t})\right] \leqslant \frac{\log d}{\eta_T} + \sum_{t=0}^T \frac{D_{h_{\mathrm{ent}}^*}(\eta_t(y_t + u_t), \eta_t y_t)}{\eta_t}.$$

(iii) If there exists $L > 0$ such that $\|u_t\|_\infty \leqslant L$ for all $t \geqslant 0$, and if $\eta_t = \sqrt{2\log d}/(L\sqrt{t+1})$, then for all $T \geqslant 0$,

$$\mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=0}^T (u_{t,i} - u_{t,i_t})\right] \leqslant L\sqrt{2(\log d)(T+1)}.$$

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. □

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 $t \geqslant 0$,

- 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

$$\mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=0}^{T} (u_{t,i} - u_{t,i_t})\right].$$

Since the whole vector u_t is unknown to the Decision Maker, one possible approach is to construct an unbiased estimator of u_t as follows

$$\hat{u}_t = \left(\mathbb{1}_{\{i=i_t\}} \frac{u_{t,i_t}}{x_{t,i_t}}\right)_{1\leqslant i\leqslant d}, \quad t \geqslant 0, \tag{3.13}$$

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. 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 $T \geqslant 0$,*

$$\mathbb{E}\left[\max_{1\leqslant i\leqslant d} \sum_{t=0}^{T} (u_{t,i} - u_{t,i_t})\right] \leqslant \mathbb{E}\left[\max_{x\in\Delta_d} \langle \hat{u}_t, x - x_t\rangle\right].$$

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

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

here 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^*_{ent}.

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

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

Proof. Using the explicit expressions from Proposition 3.4.3

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

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

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

which gives the result. □

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

$$x_t = \left(\frac{\eta_t \sum_{s=0}^{t-1} \hat{u}_{s,i}}{\sum_{j=1}^d \eta_t \sum_{s=0}^{t-1} \hat{u}_{s,j}}\right)_{1 \leqslant i \leqslant d}, \quad t \geqslant 0,$$

where $\hat{u}_t$ is defined as in (3.13) and $\eta_t = \sqrt{2(\log d)/(dL^2(t+1))}$. Then, for all $T \geqslant 0$,

$$\mathbb{E}\left[\max_{1 \leqslant i \leqslant d}\sum_{t=0}^T (u_{t,i} - u_{t,i_t})\right] \leqslant L\sqrt{2(T+1)d\log d}.$$

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

$$\hat{u}_t = \left(\mathbb{1}_{\{i = i_t\}}\frac{u_{t,i_t} - L}{x_{t,i_t}}\right)_{1 \leqslant i \leqslant d}, \quad t \geqslant 0,$$

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 $t \geqslant 0$,

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

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

$$\sum_{t=0}^{T} \left(\ell_t(x_t) - \ell_t(x) \right), \quad T \geqslant 0,\ x \in \mathcal{X}.$$

Example 4.1.3. *Online linear optimization* is a special case of online convex optimization where for all $t \geqslant 0$, 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 $t \geqslant 0$, loss functions are of the form $\ell_t(x) = \frac{1}{2} \|x - z_t\|_2^2$ 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 $t \geqslant 0$, 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 $i \in \{1, \dots, d\}$ 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 $t \geqslant 0$, 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 $(-g_t)_{t\geqslant 0}$ is also an upper bound on the initial regret corresponding to loss functions $(\ell_t)_{t\geqslant 0}$. The approach of using an online linear optimization algorithm with $(-g_t)_{t\geqslant 0}$ 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 $\|\cdot\|$, $(\gamma_t)_{t\geqslant 0}$ a positive and nonincreasing sequence, and $x_0 \in \mathcal{X} \cap \mathcal{D}_H$. Consider associated OMD iterates:*

$$x_{t+1} = \operatorname*{arg\,min}_{x \in \mathcal{X}} \left\{ \langle \gamma_t g_t, x \rangle + D_H(x, x_t) \right\}, \quad t \geqslant 0.$$

Then for all $T \geqslant 0$ and $x \in \mathcal{X} \cap \operatorname{dom} H$,

$$\sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{\max_{0 \leqslant t \leqslant T} D_H(x, x_t)}{\gamma_T} + \frac{1}{2K} \sum_{t=0}^{T} \gamma_t \|g_t\|_*^2 .$$

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

$$\gamma_t = \frac{R\sqrt{K}}{L\sqrt{t+1}}, \quad t \geqslant 0$$

guarantee for all $T \geqslant 0$,

$$\max_{x \in \mathcal{X}} \sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant RL\sqrt{\frac{2(T+1)}{K}}.$$

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

$$x_t = \operatorname*{arg\,min}_{x\in\mathcal{X}} \left\{ \left\langle \eta_t \sum_{s=0}^{t-1} g_s, x \right\rangle + h(x) \right\}, \quad t \geqslant 0.$$

(i) Then for all $T \geqslant 0$ and $x \in \operatorname{dom} h$,

$$\sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{h(x) - \min h}{\eta_T} + \frac{1}{2K} \sum_{t=0}^{T} \eta_t \left\| g_t \right\|_*^2 .$$

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

$$\eta_t = \frac{\eta\sqrt{2K}}{L\sqrt{t+1}}, \quad t \geqslant 0,$$

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

$$\sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \left(\frac{h(x) - \min h}{\eta} + \eta \right) L \sqrt{\frac{T+1}{2K}}.$$

(iii) Moreover, if $\sup_{x\in\mathcal{X}} h(x) < +\infty$, then $\eta = \sqrt{\max_{\mathcal{X}} h - \min}$ yields,

$$\max_{x\in\mathcal{X}} \sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant L \sqrt{\frac{2(\max_{\mathcal{X}} h - \min h)(T+1)}{K}}.$$

4.3 Follow the regularized leader

Loss linearization has the 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 ρ a regularizer on $\mathcal{X}$ such that $\operatorname{dom} f \cap \operatorname{dom} \rho \neq \varnothing$. Then, $f + \rho$ admits a unique minimizer on $\mathcal{X}$.*

Proof. □

Definition 4.3.2 (Follow the regularized leader)**.** Let $(\rho_t)_{t\geqslant 0}$ be a sequence of regularizers on $\mathcal{X}$ and $(\ell_t)_{t\geqslant 0}$ a sequence of loss functions on $\mathcal{X}$. Then, the associated *follow the regularized leader* (FTRL) iterates are defined as $x_0 = \arg\min_{x\in\mathcal{X}} \rho_0(x)$ and

$$x_{t+1} = \operatorname*{arg\,min}_{x\in\mathcal{X}} \left\{ \sum_{s=0}^{t} \ell_s(x) + \rho_{t+1}(x) \right\}, \quad t \geqslant 0.$$

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 $(x_t)_{t\geqslant 0}$ be defined as in Definition 4.3.2 and $h_0 = \rho_0 - \min \rho_0$. For all $t \geqslant 0$, let $g_t \in \partial \ell_t(x_t)$, $y_t = -\sum_{s=0}^{t-1} g_s$ and*

$$h_{t+1/2}(x) = h_{t+1}(x) = \sum_{s=0}^{t} D_{\ell_s}(x, x_s;\ g_s) + \rho_{t+1}(x) - \min \rho_{t+1}, \quad x \in \mathbb{R}^d.$$

Then,

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

(ii) $((x_t, y_t))_{t\geqslant 0}$ *is a sequence of UMD iterates associated with regularizers $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ and dual increments $(-g_t)_{t\geqslant 0}$,*

(iii) for all $T \geqslant 0$ and $x \in \bigcap_{t\geqslant 0} \operatorname{dom} \rho_t$,

$$\sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \rho_T(x) - \min \rho_T + \sum_{t=0}^{T} D_t^*,$$

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

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

$$\sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{\rho(x) - \min_{\mathcal{X}} \rho}{\eta_T} + \frac{1}{2K} \sum_{t=0}^{T} \eta_t \|g_t\|_*^2.$$

(v) Moreover, if there exists $L > 0$ such that loss functions $(\ell_t)_{t\geqslant 0}$ are L-Lipschitz continuous for $\|\cdot\|$, and if $\max_{\mathcal{X}} \rho_T < +\infty$, then the choice

$$\eta_t = \frac{1}{L}\sqrt{\frac{2K(\max_{\mathcal{X}} \rho - \min \rho)}{t+1}}, \quad t \geqslant 0$$

guarantees

$$\max_{x\in\mathcal{X}}\sum_{t=0}^{T}(\ell_t(x_t)-\ell_t(x))\leqslant L\sqrt{\frac{2\,(\max_{\mathcal{X}}\rho-\min\rho)\,(T+1)}{K}}.$$

Proof. (i) $h_0=\rho_0-\min\rho_0$ is a regularizer on $\mathcal{X}$ by assumption. For all $t\geqslant 0$, 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 ρ_{t+1}) and convex functions. Besides, for all $s\geqslant 0$, because ℓ_s has nonempty subdifferential on $\mathcal{X}$ by assumption, its domain contains $\mathcal{X}$. Therefore,

$$\operatorname{cl}\operatorname{dom}h_{t+1}=\operatorname{cl}\left(\operatorname{dom}\rho_{t+1}\cap\bigcap_{0\leqslant s\leqslant t}\operatorname{dom}\ell_s\right)=\operatorname{cl}\operatorname{dom}\rho_{t+1}=\mathcal{X}.$$

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 on $\mathcal{X}$.

(ii) For all $t\geqslant 0$,

$$\begin{aligned}x_t&=\operatorname*{arg\,min}_{x\in\mathcal{X}}\left\{\sum_{s=0}^{t-1}\ell_s(x)+\rho_t(x)\right\}\\&=\operatorname*{arg\,min}_{x\in\mathcal{X}}\left\{\sum_{s=0}^{t-1}(\ell_s(x_s)+\langle g_s,x-x_s\rangle+D_{\ell_s}(x,x_s;\ g_s))+\rho_t(x)\right\}\\&=\operatorname*{arg\,min}_{x\in\mathcal{X}}\left\{\sum_{s=0}^{t-1}(\langle g_s,x\rangle+D_{\ell_s}(x,x_s;\ g_s))+\rho_t(x)\right\}\\&=\operatorname*{arg\,min}_{x\in\mathcal{X}}\left\{\langle -y_t,x\rangle+h_t(x)\right\}\\&=\operatorname*{arg\,min}_{x\in\mathbb{R}^d}\left\{\langle -y_t,x\rangle+h_t(x)\right\}\\&=\nabla h_t^*\left(y_t\right),\end{aligned}$$

where the penultimate equality holds because for all $t\geqslant 0$, $\operatorname{dom}h_t\subset\mathcal{X}$. The above is equivalent to $y_t\in\partial h_t^*(x_t)$ by Proposition 1.4.6. Because $h_{t+1/2}=h_{t+1}$, the above also imply for all $t\geqslant 0$,

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

hence the result.

(iii) Applying Lemma 2.5.2 with notation therein gives for all $t \geqslant 0$,

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

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

$$\sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant D_0 + \sum_{t=0}^{T} D_t^* + (\rho_{T+1}(x) - \min \rho_{T+1}) - (\rho_0(x) - \min \rho_0).$$

Note that for a given $T \geqslant 0$, iterates $x_0, \dots, x_T$ do not depend on ρ_{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,

$$D_0 = h_0(x) - h_0(x_0) - \langle 0, x - x_0\rangle = \rho_0(x) - \min \rho_0.$$

Simplifying gives the result.

(iv) Continuing the above analysis (fixing $T \geqslant 0$ and considering $\rho_{T+1} = \rho_T$, which here corresponds to $\eta_{T+1} = \eta_T$), for all $0 \leqslant t \leqslant T$, h_t is K/η_t-strongly convex for $\|\cdot\|$ as the sum of convex functions and a K/η_t-strongly convex function (ρ_t); besides, $h_{t+1} \geqslant h_t$ because parameters $(\eta_t)_{t\geqslant 0}$ 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.

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

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 1.6.8.

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 ℓ_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 $(K_t)_{t\geqslant 0}$ be a positive sequence, $(\ell_t)_{t\geqslant 0}$ loss functions such that for all $t \geqslant 0$, ℓ_t is K_t-strongly convex for $\|\cdot\|_2$ and $x_0 \in \mathcal{X}$. Consider online gradient descent iterates:*

$$x_{t+1} = \Pi_{\mathcal{X}}(x_t - \gamma_t g_t), \quad t \geqslant 0,$$

where $g_t \in \partial\ell_t(x_t)$ and $\gamma_t = \left(\sum_{s=0}^t K_s\right)^{-1}$. Then, for all $T \geqslant 0$,

$$\max_{x\in\mathcal{X}} \sum_{t=0}^T (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{1}{2}\sum_{t=0}^T \frac{\|g_t\|_2^2}{\sum_{s=0}^t K_s}.$$

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

$$\max_{x\in\mathcal{X}} \sum_{t=0}^T (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{L^2}{2K}(1 + \log(T+1)).$$

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

$$\sum_{t=0}^T (\ell_t(x_t) - \ell_t(x)) \leqslant \sum_{t=0}^T \left(\langle g_t, x_t - x\rangle - \sum_{t=0}^T \frac{K_t}{2}\|x - x_t\|_2^2\right).$$

We then conclude by applying Proposition 3.3.13 and simplifying. □

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 $\mathbb{R}^d$, $(K_t)_{t\geqslant 0}$ a positive sequence, and $(\ell_t)_{t\geqslant 0}$ loss functions on $\mathcal{X}$ such that for all $t \geqslant 0$, ℓ_t is K_t-strongly convex for $\|\cdot\|$ and $x_0 \in \mathcal{X}$. Consider*

$$x_{t+1} = \operatorname*{arg\,min}_{x\in\mathcal{X}} \left\{ \sum_{s=0}^{t} \ell_s(x) \right\}, \quad t \geqslant 0. \tag{4.2}$$

Let $h_0 = \varepsilon \|x - x_0\|_2^2 + I_{\mathcal{X}}(x)$ (for some $\varepsilon > 0$) and for all $t \geqslant 0$,

$$h_{t+1/2}(x) = h_{t+1}(x) = \sum_{s=0}^{t} D_{\ell_s}(x, x_s;\ g_s) + I_{\mathcal{X}}(x),$$

where for all $s \geqslant 0$, $g_s \in \partial \ell_s(x_s)$. Then,

(i) $(h_t)_{\frac{1}{2}\mathbb{N}}$ *is a sequence of regularizers on $\mathcal{X}$,*

(ii) $((x_t, -\sum_{s=1}^{t-1} g_s))_{t\geqslant 0}$ *is a sequence of UMD iterates associated with regularizers $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ and dual increments $(-g_t)_{t\geqslant 0}$,*

(iii) *for all $T \geqslant 0$,*

$$\max_{x\in\mathcal{X}} \sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{1}{2} \sum_{t=0}^{T} \frac{\|g_t\|_*^2}{\sum_{s=0}^{t} K_s}.$$

Proof. The FTL iterates are well-defined because for $t \geqslant 0$, $\sum_{s=0}^{t} \ell_s + I_{\mathcal{X}}$ is a proper lower semicontinuous strongly convex function, and thus admits a unique minimizer by Proposition 1.6.11.

(i) Let $t \in \frac{1}{2}\mathbb{N}$. h_t is lower semicontinuous as the sum of lower semicontinuous functions. It is also strictly convex because it is strongly convex. For $t = 0$, $\operatorname{dom} h_0 = \mathcal{X}$ as an immediate consequence of the definition and for $t \geqslant 1$, because the domain of the convex losses contain $\mathcal{X}$,

$$\mathcal{X} \subset \operatorname{dom} \left(\sum_{s=0}^{t-1} D_{\ell_s}(\,\cdot\,, x_s;\ g_s) \right),$$

and therefore

$$\operatorname{dom} h_t = \operatorname{dom} \left(\sum_{s=0}^{t-1} D_{\ell_s}(\,\cdot\,, x_s;\ g_s) + I_{\mathcal{X}} \right) = \mathcal{X}.$$

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 $t \geqslant 0$, 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

$$\sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \varepsilon \left\| x - x_0 \right\|_2^2 + \sum_{t=0}^{T} D_t^*.$$

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 $t \geqslant 0$. 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 $y_t = -\sum_{s=0}^{t-1} g_s$ we can write

$$\begin{aligned}
x_t = \nabla h_t^*\left(y_t\right) &= \operatorname*{arg\,min}_{x \in \mathbb{R}^d} \left\{ -\left\langle y_t, x \right\rangle + \sum_{s=0}^{t-1} D_{\ell_s}(x, x_s;\ g_s) \right\} \\
&= \operatorname*{arg\,min}_{x \in \mathbb{R}^d} \left\{ -\left\langle y_t, x \right\rangle + \sum_{s=0}^{t} D_{\ell_s}(x, x_s;\ g_s) \right\} \\
&= \nabla h_{t+1}^*(y_t) = \nabla h_{t+1/2}^*(y_t),
\end{aligned}$$

and similarly $h_t^*(y_t) = h_{t+1/2}^*(y_t) = h_{t+1}^*(y_t)$. Besides, because ℓ_t (and therefore $D_{\ell_t}(\,\cdot\,, x_t;\ g_t)$) is K_t-strongly convex for $\|\cdot\|$ by assumption, $h_{t+1} = \sum_{s=0}^{t} D_{\ell_t}(\,\cdot\,, x_t;\ g_t)$ is $\sum_{s=0}^{t} K_s$-strongly convex. Hence,

$$\begin{aligned}
D_t^* &= h_{t+1/2}^*(y_t - g_t) - h_t^*(y_t) + \langle g_t, x_t \rangle \\
&= h_{t+1}^*(y_t - g_t) - h_{t+1}^*(y_t) + \left\langle g_t, \nabla h_{t+1}^*(y_t) \right\rangle \\
&= D_{h_{t+1}^*}(y_t - g_t, y_t) \\
&\leqslant \frac{\|g_t\|_*^2}{2 \sum_{s=0}^{t} K_s},
\end{aligned}$$

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 $t \geqslant 0$,

- 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

$$\sum_{t=0}^{T}\left(\frac{1}{2}(\langle w_t, x_t\rangle - z_t)^2 - \frac{1}{2}(\langle w_t, x\rangle - z_t)^2\right), \quad T \geqslant 0,\ x \in \mathbb{R}^d.$$

The following algorithm is a variant of FTRL which at step $t \geqslant 0$ uses the knowledge of w_t to choose x_t. Since the loss function at step $t \geqslant 0$ writes

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

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

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

$$x_t = \operatorname*{arg\,min}_{x\in\mathbb{R}^d}\left\{\frac{1}{2}\sum_{s=0}^{t-1}(z_s - \langle w_s, x\rangle)^2 + \frac{1}{2}\left\langle w_t, x\right\rangle^2 + \frac{\lambda}{2}\left\|x\right\|_2^2\right\}, \quad t \geqslant 0.$$

Lemma 4.5.2 (Lemma 1.11 and Theorem 11.7 in [CBL06])**.** *Let $T \geqslant 0$, $w_0, \dots, w_T \in \mathbb{R}^d$ and $\lambda > 0$. For all $0 \leqslant t \leqslant T$, denote $S_t = \lambda I + \sum_{s=0}^{t} w_s w_s^\top$. Then,*

$$\sum_{t=0}^{T} w_t^\top S_t^{-1} w_t \leqslant \sum_{i=1}^{d}\log\left(1 + \frac{\lambda_i}{\lambda}\right),$$

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

Proposition 4.5.3. *Consider the iterates $(x_t)_{t\geqslant 0}$ defined as Definition 4.5.1, $h_0(x) = \frac{\lambda}{2}\left\|x\right\|_2^2$ and*

$$h_{t+1/2}(x) = h_{t+1}(x) = \frac{1}{2}x^\top\left(\frac{\lambda}{2}I + \sum_{s=0}^{t+1} w_s w_s^\top\right)x, \quad t \geqslant 0.$$

Then,

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

(ii) $((x_t, \sum_{s=0}^{t-1} z_s w_s))_{t\geqslant 0}$ *is a sequence of UMD iterates associated with* $(h_t)_{t\in\frac{1}{2}\mathbb{N}}$ *and dual increments* $(z_t w_t)_{t\geqslant 0}$*,*

(iii) for all $T \geqslant 0$*,*

$$\sum_{t=0}^{T} (\ell_t(x_t) - \ell_t(x)) \leqslant \frac{\lambda}{2} \|x\|_2^2 + \frac{dZ_T^2}{2} \log\left(1 + \frac{W_T^2 T}{\lambda d}\right).$$

where

$$Z_T = \max_{0\leqslant t\leqslant T} |z_t| \quad \textit{and} \quad W_T = \max_{0\leqslant t\leqslant T} \|w_t\|_2 .$$

Proof. (i) For $t \geqslant 0$, $\frac{\lambda}{2} I + \sum_{s=0}^{t} w_s w_s^\top$ 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 $t \geqslant 0$, denote $y_t = \sum_{s=0}^{t-1} z_s w_s$. Then,

$$\begin{aligned}
x_t &= \operatorname*{arg\,min}_{x'\in\mathbb{R}^d} \left\{ \frac{1}{2} \sum_{s=0}^{t-1} \left(\langle w_s, x'\rangle - z_s\right)^2 + \frac{1}{2} \langle w_t, x\rangle^2 + \frac{\lambda}{2} \|x\|_2^2 \right\} \\
&= \operatorname*{arg\,min}_{x'\in\mathbb{R}^d} \left\{ \frac{1}{2} \sum_{s=0}^{t-1} \left(\langle w_s, x'\rangle^2 - 2z_s \langle w_s, x'\rangle + z_s^2\right) + \frac{1}{2} \langle w_t, x'\rangle^2 + \frac{\lambda}{2} \|x\|_2^2 \right\} \\
&= \operatorname*{arg\,min}_{x'\in\mathbb{R}^d} \left\{ -\sum_{s=0}^{t-1} \langle z_s w_s, x\rangle + \frac{1}{2} \sum_{s=0}^{t} \langle w_s, x\rangle^2 + \frac{\lambda}{2} \|x\|_2^2 \right\} \\
&= \nabla h_t^*(y_t),
\end{aligned}$$

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

$$\langle z_t w_t, x - x_t\rangle \leqslant D_t - D_{t+1} + D_t^* + h_{t+1}(x) - h_t(x), \quad t \geqslant 0.$$

Summing gives

$$\sum_{t=0}^{T} \langle z_t w_t, x - x_t\rangle \leqslant D_0 + \sum_{t=0}^{T} D_t^* + h_{T+1}(x) - h_0(x).$$

Because iterates $x_0, \dots, x_T$ do not depend on h_{T+1}, the above analysis can be carried by h_{T+1} replaced by h_T. Besides, $x_0 = y_0 = 0$ by definition, which implies $D_0 = h_0(x)$. Therefore,

$$\sum_{t=0}^{T} \langle z_t w_t, x - x_t\rangle \leqslant \sum_{t=0}^{T} D_t^* + \frac{\lambda}{2} \|x\|_2^2 + \frac{1}{2} x^\top \left(\sum_{t=0}^{T} w_t w_t^\top\right) x. \tag{4.3}$$

For each $t \geqslant 0$, denote $S_t = \frac{\lambda}{2} I + \sum_{s=0}^{t} w_s w_s^\top$ and using Fenchel's inequality from Remark 1.4.3,

$$\begin{aligned}
D_t^* &= h_{t+1}^*(y_{t+1}) - h_t^*(y_t) - \langle z_t w_t, x_t \rangle \\
&= D_{h_t^*}(y_{t+1}, y_t) + h_{t+1}^*(y_{t+1}) - h_t^*(y_{t+1}) \\
&\leqslant D_{h_t^*}(y_{t+1}, y_t) + \langle y_{t+1}, x_{t+1} \rangle - h_{t+1}(x_{t+1}) - \langle y_{t+1}, x_{t+1} \rangle + h_t(x_{t+1}) \\
&= \frac{1}{2}(z_t w_t)^\top S_t^{-1}(z_t w_t) - \langle w_{t+1}, x_{t+1} \rangle^2 ,
\end{aligned} \tag{4.4}$$

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:

$$\begin{aligned}
&\frac{1}{2}\left((\langle w_t, x_t \rangle - z_t)^2 - (\langle w_t, x \rangle - z_t)^2 \right) \\
&\qquad = \frac{1}{2} \langle w_t, x_t \rangle^2 - \frac{1}{2} \langle w_t, x \rangle^2 + z_t \langle w_t, x - x_t \rangle . \end{aligned} \tag{4.5}$$

Combining (4.3), (4.4) and (4.5) together with Lemma 4.5.2 gives the result.

□

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.