Optimality of the Subgradient Algorithm in the Stochastic Setting

Daron Anderson
Department of Computer Science and Statistics
Trinity College Dublin
Ireland

Douglas Leith
Department of Computer Science and Statistics
Trinity College Dublin
Ireland

September 2019

Abstract
Recently Jaouad Mourtada and Stéphane Gaïffas showed the anytime hedge algorithm has pseudo-regret $O(\log(d)/\Delta)$ if the cost vectors are generated by an i.i.d sequence in the cube $[0,1]^d$. This is remarkable because the Hedge algorithm was designed for the antagonistic setting. We prove a similar result for the anytime subgradient algorithm on the simplex. Given i.i.d cost vectors in the unit ball our pseudo-regret bound is $O(1/\Delta)$ and does not depend on the dimension of the problem.

1 Introduction and Related Work

Online convex optimisation comes from [Zinkevich 2003] who gave the following toy example: Suppose at the start of the year we must choose which of two crops to plant. Bananas grow best in warm weather and potatoes grow best in cold weather. We may plant all bananas or all potatoes or mix the two in any proportion. The average temperature for the year determines which of the two crops grows better and thus the profit for that year.

Suppose also we have two neighbors, one who plants only potatoes every year and one who plants only bananas. Zinkevich gave algorithms whereby after $N$ seasons the more successful of the two neighboring will be only $O(\sqrt{N})$ more succesful than us. Taking averages the difference is $O(1/\sqrt{N})$ which decays to zero.

Formally suppose $a_1, a_2, \ldots \in \mathbb{R}^d$ are cost vectors. On turn $n$ we know only $a_1, \ldots, a_{n-1}$ and must select an action $x_n$ in the $d$-simplex $\mathcal{S}$ with a mind to minimising the sum $\sum_{i=1}^{N} a_i \cdot x_i$. There exist algorithms whereby the regret $\sum_{i=1}^{N} a_i \cdot x_i - \sum_{i=1}^{N} a_i \cdot x^*$ has order $O(\sqrt{N})$ for all $x^* \in \mathcal{S}$ simultaneously. Remarkably these algorithms work when the only assumption is a uniform bound for the cost vectors. To extend the metaphor, there is a planting strategy that is effective even if the weather is selected antagonistically by nature.
In reality the weather does not change based on which crops we plant. One popular model for predictable weather is to assume $a_1, a_2, \ldots$ are generated from some i.i.d sequence of random variables with expectation $a$. In this case $a_1, a_2, \ldots, a_{n-1}$ provide estimates for $a_n$ and there are strategies where for $\mathbb{E}[a_i] = a$ the pseudo-regret $\sum_{i=1}^{N} a \cdot x_i - \sum_{i=1}^{N} a \cdot x^*$ has expectation $O(\log N)$.

Such algorithms exist even in the bandit setting, where on turn $n$ we know only the sequence of costs $a_1 \cdot x_1, a_2 \cdot x_2, \ldots, a_{n-1} \cdot x_{n-1}$ rather than the full information setting where we know the cost vectors $a_1, a_2, \ldots, a_{n-1}$ themselves. The bandit setting comes from Flaxman et al. [2004]. For a recent exposition of i.i.d bandits see Chapter 2 of Bubeck and Cesa-Bianchi [2012].

There has been much interest in universal algorithms for the bandit setting. For example see Seldin and Slivkins 2014, Zimmert and Seldin 2018, Auer and Chiang 2016, Seldin and Lugosi 2017, Wei and Luo 2018. Recently Bubeck and Slivkins 2012 gave an algorithm with expected regret $O(\sqrt{N \log^{3/2} N})$ in the antagonistic setting that adapts to give pseudo-regret $O(\log^{2} N)$ in the stochastic setting.

Back in the full-information setting Mourtada and Gáifás 2019 have recently shown the anytime Hedge algorithm (also called Exp3) is already universal. It is well known to have $O(L\sqrt{N})$ worst case regret. For stochastic costs vectors in $[0, L]^d$ they show the pseudo-regret is $O(L^2 \log(d) / \Delta)$.

Next to Hedge, the most simple and familiar algorithm for online optimisation is the subgradient algorithm. Here we show the subgradient is also universal. It is known to have $O(L\sqrt{N})$ worst case regret. In Section 3 our main result (Theorem 2) says the pseudo-regret is $O(L^2 / \Delta)$ in the stochastic setting. Here $L$ bounds the 2-norms of the cost vectors. So the preference of one algorithm over the other should depend on the dimension and the geometry of the particular problem. If the cost vectors naturally come from a sphere rather than a cube we can achieve pseudo regret independent of the dimension.

In Section 4 we consider tuning the subgradient step size based on some prior knowledge of the probability the cost vectors are i.i.d or antagonistic. While the optimal step size has no closed form Theorem 3 gives an upper bound on the associated (pseudo-)regret. In Section 5 we replace the simplex with a more general domain in $\mathbb{R}^d$. We give examples of domains and i.i.d cost vectors where the pseudo-regret is $\Omega(N^{1/2-\varepsilon})$ for any $\varepsilon > 0$. Thus the pseudo-regret can be almost as bad as the $O(\sqrt{N})$ worst-case regret. In Section 6 we return to the simplex and prove some tail bounds. For $\mathcal{R}$ the pseudo-regret our Theorem 13 shows the subgradient algorithm has

$$P(\mathcal{R} > \delta + O(\sqrt{\delta})) \leq O(\delta e^{-C\delta})$$

for some constant $C > 0$. The constant does not not depend on the step size but the

The most important concentration results for i.i.d cost vectors generalise immediately to when $a_i - \mathbb{E}[a_i]$ form a martingale difference sequence. The elements of a martingale difference sequence are not required to be either identical or independent. Our results hold for martingales but to avoid lengthy definitions we stick to the i.i.d case.
$O(\sqrt{\delta})$ function does. Section 7 contains plots of some simulations for varying dimension and noisiness of the i.i.d sequence. The plots suggest the constants in Theorem 2 are two orders of magnitude higher than the theoretical best constants.

The authors are unaware of any universal subgradient-type analysis that appears explicitly in the literature. The methods of Sani et al. [2014] can be used to combine the worst-case subgradient algorithm with a second algorithm tailored for the i.i.d setting, for example follow-the-leader. Their bounds have better constants than our Theorem 2 at the cost of added complexity. Another drawback is their method can combine only a pair of algorithms, rather than an arbitrary collection.

2 Terminology and Notation

Throughout $d$ is the dimension of the online optimisation problem. We write $S$ for the $d$-simplex $S = \{(x_1, x_2, \ldots, x_d) \in \mathbb{R}^d : \text{all } x_j \geq 0 \text{ and } x_1 + \ldots + x_d = 1\}$. We call $S$ the action set and elements of $S$ are called actions. We write $1$ for the vector $(1, \ldots, 1) \in \mathbb{R}^d$.

For any function $f : X \rightarrow \mathbb{R}$ we write $\text{argmin}\{f(x) : x \in X\}$ for the set of minimisers. Each linear function on the simplex is minimised on some vertex. Hence $\min\{a \cdot x : x \in S\} = \min\{a \cdot e_j : j \leq d\}$. We write $\|\cdot\|$ for the Euclidean norm and for any convex $X \subset \mathbb{R}^d$ we write $P_X(x) = \text{argmin}\{\|y - x\|^2 : y \in X\}$ for the projection of $x$ onto $\Omega$.

Thoughout the cost vectors $a_1, a_2, \ldots \in \mathbb{R}^d$ are realisations of a sequence of i.i.d random variables with each $E[a_i] = a$. When we write $b_1, b_2, \ldots$ we make no assumptions on whether the cost vectors are i.i.d or otherwise. We assume bounds of the form $\|a_i - a\| \leq R$ and $\|a_i\| \leq L$.

For cost vectors $b_1, b_2, \ldots$ the regret of an action sequence $x_1, \ldots, x_N$ is defined as $\sum_{i=1}^{N} b_i \cdot x_i - \sum_{i=1}^{N} b_i \cdot x^* \text{ for } x^* \in \text{argmin}\sum_{i=1}^{N} b_i \cdot x$. For stochastic cost vectors $a_1, a_2, \ldots$ the pseudo-regret of the action sequence is $E\left[\sum_{i=1}^{N} a_i \cdot x_i\right] - Na \cdot x^*$ where the expectation is taken over the domain of $a_1, \ldots, a_N$.

By permuting the coordinates if neccesary we assume $e_1$ is a minimiser of $a$ and that the differences $\Delta_j = a \cdot (e_j - e_1)$ satisfy $0 = \Delta_1 \leq \Delta_2 \leq \ldots \leq \Delta_d$. Note the permutation is part of the analysis only, and our algorithm do not require access to it. We write $\Delta = \Delta_2 = \min\{\Delta_j : \Delta_j > 0\}$ and $\Delta = \Delta_d = \max\{\Delta_j : \Delta_j > 0\}$.

3 Pseudo-Regret

The subgradient algorithm is one of the simplest and most familiar algorithms for online convex optimisation. The anytime version Algorithm 1 does not need the time horizon in advance. In this algorithm the step size on turn $n$ is $\eta / \sqrt{n-1}$ where $\eta > 0$ is a design parameter.

The subgradient algorithm is known to have $O(L \sqrt{N})$ regret. See Shalev-Shwartz [2012], Zinkevich [2003].
Algorithm 1: Anytime Subgradient Algorithm

**Data:** Compact convex subset $\mathcal{X} \subset \mathbb{R}^d$. Parameter $\eta > 0$.

1. select action $x_1 = P_{\mathcal{X}}(0)$
2. pay cost $a_1 \cdot x_1$
3. for $n = 2, 3, \ldots$ do
   4. receive $a_{n-1}$
   5. $y_n = -\eta \left( \frac{a_1 + \ldots + a_{n-1}}{\sqrt{n-1}} \right)$
   6. select action $x_n = P_{\mathcal{X}}(y_n)$
   7. pay cost $a_n \cdot x_n$

**Theorem 1.** For cost vectors $b_1, b_2, \ldots, b_N$ with all $\|b_i\| \leq L$ Algorithm 1 with parameter $\eta > 0$ has regret satisfying

$$\sum_{i=1}^{N} b_i \cdot (x_i - x^*) \leq (L\|\mathcal{X}\| - \eta L^2) + \left( \frac{1}{2\eta}\|\mathcal{X}\|^2 + 2\eta L^2 \right) \sqrt{N}$$

for $\|\mathcal{X}\| = \max \{ \|x\| : x \in \mathcal{X} \}$. In particular for $\mathcal{X} = S$ and $\eta = 1/L$ we have

$$\sum_{i=1}^{N} b_i \cdot (x_i - x^*) \leq \frac{5L}{4} \sqrt{N}.$$ 

**Proof.** See Appendix A. \qed

Our main result is that, in addition to the above bound, the algorithm adapts to the stochastic case to have $O(L^2/\Delta)$ pseudo-regret. In particular the bound is independent of the dimension of the problem.

**Theorem 2.** Suppose the cost vectors $a_1, a_2, \ldots$ are i.i.d with all $\|a_i\| \leq L$ and $\|a_i - a\| \leq R$ for $\mathbb{E}[a_i] = a$. Suppose we run Algorithm 1 on the simplex with parameter $\eta$. The pseudo-regret is no more than

$$\Delta + \frac{18}{\Delta} \left( \frac{1}{\eta^2} + 4R^2 e^{-1/2R^2\eta^2} \right)$$

for $\Delta = \max \{ \Delta_j : i \leq d \}$ and $\Delta = \min \{ \Delta_j : \Delta_j > 0 \}$.

For ease of notation assume the suboptimality gaps $0 = \Delta_1 < \Delta_2 < \ldots < \Delta_d$ are distinct. We later explain how to modify the proofs to when some of the gaps coincide. The strategy is to obtain separate bounds over the intervals

$I(d) = \{1, 2, \ldots, N_d\} \quad I(j) = \{N_{j+1} + 1, \ldots, N_j\} \quad I(1) = \{N_2 + 1, \ldots\}$

where we define $N_j = \lceil (3/\eta \Delta_j)^2 \rceil$ for $j = d - 1, \ldots, 2$. The first bound is easy and corresponds to the first term of $[1]$. 

4
Lemma 1. The pseudo-regret over $I(d)$ is at most $\Delta_d N_d$.

Proof. The pseudo-regret for any given round is at most $\Delta_d$. Since there are $N_d$ rounds the pseudo-regret is at most $\Delta_d N_d$. \qed

For the intervals $I(d-1), \ldots, I(1)$ we use probabilistic bounds. The term $y_{n+1}$ from Algorithm 1 should be thought of as a noisy version of $-\eta \sqrt{n} a$. We can isolate the noise by writing

$$y_{n+1} = -\frac{\eta}{\sqrt{n}} \sum_{i=1}^{n} a_i = -\eta \sqrt{n} a + \frac{\eta}{\sqrt{n}} \sum_{i=1}^{n} (a - a_i) = -\eta \sqrt{n} a + \eta \varepsilon$$

for $\varepsilon = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (a - a_i)$ \hspace{1cm} (2)

The central idea behind our proof is that for $n$ large $-\eta \sqrt{n} a$ is inside the normal cone at $e_1$ (shown here in green) hence projects onto $e_1$ and the pseudo-regret for that round is zero. Moreover as $n$ grows $-\eta \sqrt{n} a$ moves deeper inside the normal cone and a larger noise term $\eta \varepsilon$ is required to push it back out.

Concentration inequalities ensure the noise terms are on expectation the same size each round. Hence the chance of not selecting $e_1$ shrinks with time. Indeed it shrinks fast enough to give a convergent series.

More generally we will derive conditions that make $y_{n+1}$ project into the convex hull of $\{e_1, \ldots, e_j\}$ and make the pseudo-regret for that round at most $\Delta_j$. For example in the schematic $-\eta \sqrt{n} a$ travels first into the normal cone (blue) of the face $e_1 e_2$ and later into the cone of the optimal vertex. So we get a better bound by considering the decreasing sequence of normal cones.

Lemma 2. Let $S \subset \mathbb{R}^d$ be the $d$-simplex. Suppose $w \in \mathbb{R}^d$ has two coordinates $k, \ell$ with $w_k - w_\ell \geq 1$. Then $P_S(w)$ has $\ell$-coordinate zero.

Proof. See Appendix B. \qed

Lemma 2 combined with the following lemma, says a suitable definition of $n$ being large and $\varepsilon$ small is $n > N_j$ and $\|\varepsilon\| \leq (\Delta_j/3) \sqrt{n}$. 

Lemma 3. Let $j \in \{2, \ldots, d\}$ and $n > N_j$. Suppose on turn $n + 1$ the error $\varepsilon$ from (2) has $\|\varepsilon\| \leq (\Delta_j/3)\sqrt{n}$. Then the pseudo-regret for turn $n + 1$ is no more than $\Delta_j - 1$.

Proof. In the notation of Lemma 2 take $k = 1$ and $\ell \geq j$ and $w = -\eta\sqrt{n}a$. We have

$$w_k - w_{\ell} = \eta\sqrt{n}(a_{\ell} - a_1) = \eta\sqrt{n}\Delta_{\ell} \geq \eta\sqrt{n}\Delta_j$$

since $\ell \geq j$ implies $\Delta_{\ell} \geq \Delta_j$. Now take $w = \eta\varepsilon$. We have

$$w_k - w_{\ell} = \eta(\varepsilon_k - \varepsilon_{\ell}) \geq -2\eta\|\varepsilon\| \geq -2\frac{2\eta\Delta_j}{3}\sqrt{n}$$

Hence for $w = -\eta\sqrt{n}a + \eta\varepsilon$ and $k = 1$ and all $\ell \geq j$ we have

$$w_k - w_{\ell} \geq \eta\Delta_j\sqrt{n} - \frac{2\eta\Delta_j}{3}\sqrt{n} = \frac{\eta\Delta_j}{3}\sqrt{n} \geq \frac{\eta\Delta_j}{3}\sqrt{N_j} \geq 1$$

where the last inequality uses the definition $N_j = \lceil (\Delta_j/\eta)^2 \rceil$. Since $w = -\eta\sqrt{n}a + \eta\varepsilon = y_{n+1}$ and $\ell \geq j$ is arbitrary Lemma 2 says $x_{n+1}$ has coordinates $j, j + 1, \ldots, d$ equal zero. Hence the regret on turn $n + 1$ is at most $\Delta_j - 1$. □

Lemma 3 says small error leads to small pseudo-regret. Next we use concentration results to force the error to be small.

Lemma 4. Let $j \in \{2, \ldots, d\}$ and $n \geq N_j$. Then we have

$$P\left(a \cdot (x_{n+1} - e_1) \leq \Delta_{j-1}\right) \geq 1 - 2\exp\left(-\frac{\Delta_j^2}{18R^2n}\right).$$

(3)

In particular for $n > N_2$ we have

$$P\left(a \cdot (x_{n+1} - e_1) = 0\right) \geq 1 - 2\exp\left(-\frac{\Delta_j^2}{18R^2n}\right).$$

(4)

Proof. By Theorem 7 Appendix C we have for each $r > 0$ the bound

$$P\left(\|\sum_{i=1}^n (a - a_i)\| \leq r\sqrt{n}\right) \geq 1 - 2\exp\left(-\frac{r^2}{2R^2}\right).$$

Recall the definition (2) of $\varepsilon$ to see

$$P(\|\varepsilon\| \leq r) = P\left(\frac{1}{\sqrt{n}}\left\|\sum_{i=1}^n (a - a_i)\right\| \leq r\right) \geq 1 - 2\exp\left(-\frac{r^2}{2R^2}\right).$$

For $r = (\Delta_j/3)\sqrt{n}$ the right side is $1 - 2\exp\left(-\frac{\Delta_j^2}{18R^2n}\right)$. Thus $\|\varepsilon\| \leq (\Delta_j/3)\sqrt{n}$ with probability at least $1 - 2\exp\left(-\frac{\Delta_j^2}{18R^2n}\right)$. Using Lemma 3 completes the proof. □
Recall $N_d < N_{d-1} < \ldots < N_2$. Hence over each interval $I(j)$ the bound (3) holds with $j$ replaced by each $k \geq j$. For ease of notation write the exponents as $\Gamma(j) = \frac{\Delta_j^2}{18R^2}$. The next lemma bounds the pseudo-regret over each $I(d-1), \ldots, I(1)$.

**Lemma 5.** Let $j \in \{2, \ldots, d-1\}$. The pseudo-regret over $I(j)$ is at most

$$(N_j - N_{j+1})\Delta_j + 2 \sum_{n>N_{j+1}}^{N_j} \sum_{k>j}^d (\Delta_k - \Delta_{k-1})e^{-\Gamma(k)n}.$$ 

The pseudo-regret over $I(1)$ is at most

$$2 \sum_{n>N_2}^{\infty} \Delta_2e^{-\Gamma(2)n} + 2 \sum_{n>N_2}^{\infty} \sum_{k>2}^d (\Delta_k - \Delta_{k-1})e^{-\Gamma(k)n}.$$ 

**Proof.** Let $j \in \{2, \ldots, d-1\}$ and suppose $n \in I(j)$. Then $n > N_{j+1}$. From the previous lemma we have

$$P(a \cdot (x_{n+1} - e_1) > 0) \leq 2e^{-\Gamma(k)n}$$

for all $k \geq j$. Thus over $[0, \infty)$ the complementary CDF $F(t) = P(a \cdot (x_{n+1} - e_1) > t)$ is dominated by the piecewise function

$$f(x) = \begin{cases} 
1 & x \leq \Delta_j \\
2e^{-\Gamma(k)n} & \Delta_{k-1} < x \leq \Delta_k \\n0 & \Delta_d < x 
\end{cases}$$

By Lemma 16 we have

$$\mathbb{E}[a \cdot (x_{n+1} - e_1)] \leq \int_0^{\infty} f(t)dt = \int_0^{\Delta_d} f(t)dt = \Delta_j + 2 \sum_{k>j}^d (\Delta_k - \Delta_{k-1})e^{-\Gamma(k)n}.$$ 

Summing over $n \in I(j) = \{N_{j+1} + 1, \ldots, N_j\}$ we see the pseudo-regret is at most

$$(N_j - N_{j+1})\Delta_j + 2 \sum_{n>N_{j+1}}^{N_j} \sum_{k>j}^d (\Delta_k - \Delta_{k-1})e^{-\Gamma(k)n}.$$ 

This proves the first part of the conclusion. For the second part suppose $n > N_2$. From the previous lemma (3) holds for $j \in \{2, \ldots, d\}$. By (4) we have $P(a \cdot (x_{n+1} - e_1) > 0) \leq 2e^{-\Gamma(2)n}$. Hence over $[0, \infty)$ the complementary CDF $F(t) = P(a \cdot (x_{n+1} - e_1) > t)$ is dominated by the piecewise function
\[
f(x) = \begin{cases} 
2e^{-\Gamma(2)n} & 0 < x \leq \Delta_2 \\
2e^{-\Gamma(k)n} & \Delta_{k-1} < x \leq \Delta_k \text{ with } k > 3 \\
0 & \Delta_d < x 
\end{cases}
\]

Like before we integrate to get

\[
E[a \cdot (x_{n+1} - e_1)] \leq \int_0^\infty f(t)dt = \int_0^{\Delta_d} f(t)dt = 2\Delta_2e^{-\Gamma(2)n} + 2 \sum_{k>3}^{d} (\Delta_k - \Delta_{k-1})e^{-\Gamma(k)n}.
\]

Summing over \( n \in I(1) = \{N_2 + 1, \ldots\} \) we see the pseudo-regret is at most

\[
2 \sum_{n>N_2}^{\infty} \Delta_2e^{-\Gamma(2)n} + 2 \sum_{n>N_2}^{\infty} \sum_{k>3}^{d} (\Delta_k - \Delta_{k-1})e^{-\Gamma(k)n}.
\]

This proves the second part of the conclusion.

Next we add up our bounds over the separate intervals and simplify.

**Lemma 6.** The pseudo-regret is at most \( \Delta_d + \frac{18}{\Delta_2} \left( \frac{1}{\eta^2} + 4R^2e^{-1/2n^2R^2} \right) \).

**Proof.** Define each \( A(k, n) = (\Delta_k - \Delta_{k-1})e^{-\Gamma(k)n} \). Summing the bounds from Lemmas 1 and 5 and gathering terms we see the total pseudo-regret is at most

\[
N_d\Delta_d + \sum_{j=2}^{d-1} (N_j - N_{j+1})\Delta_j + 2 \sum_{n>N_2}^{\infty} \Delta_2e^{-\Gamma(2)n}
\]

\[
+ 2 \sum_{j=2}^{d-1} \sum_{n>N_{j+1}} \sum_{k>j}^{d} A(k, n) + 2 \sum_{n>N_2}^{\infty} \sum_{j>2}^{d} A(k, n).
\]

By Lemma 19 the second line equals \( 2 \sum_{j=3}^{d} \sum_{n>N_j} A(j, n) \). Thus we have

\[
N_d\Delta_d + \sum_{j=2}^{d-1} (N_j - N_{j+1})\Delta_j + 2 \sum_{j=3}^{d} \sum_{n>N_j} (\Delta_j - \Delta_{j-1})e^{-\Gamma(j)n} + 2 \sum_{n>N_2}^{\infty} \Delta_2e^{-\Gamma(2)n}.
\]

For the first two terms recall the definition \( N_j = \left\lceil (3/\eta \Delta_j)^2 \right\rceil \). Hence by Lemma 18 we have

\[
N_d\Delta_d + \sum_{j=2}^{d-1} (N_j - N_{j+1})\Delta_j \leq \Delta_d + \left( \frac{3}{\eta \Delta_d} \right)^2 \Delta_d + \sum_{j=2}^{d-1} \left( \left( \frac{3}{\eta \Delta_j} \right)^2 - \left( \frac{3}{\eta \Delta_{j+1}} \right)^2 \right) \Delta_j
\]

\[
= \Delta_d + \frac{9}{\eta^2} \left( \frac{1}{\Delta_d^2} \right) \Delta_d + \frac{9}{\eta^2} \sum_{j=2}^{d-1} \left( \frac{1}{\Delta_j^2} - \frac{1}{\Delta_{j+1}^2} \right).
\]

8
Grouping like \( \frac{1}{\Delta_j} \) terms we get

\[
\Delta_d + \frac{9}{\eta^2} \left( \frac{1}{\Delta_2} + \frac{\Delta_3 - \Delta_2}{\Delta_3^2} + \ldots + \frac{\Delta_d - \Delta_{d-1}}{\Delta_d^2} \right).
\]

Lemma 17 says the above is no more than \( \Delta_d + \frac{18}{\eta^2} \Delta_2 \).

For the second sum in (5) the terms are decreasing in \( n \). Hence we can estimate the sums using integrals:

\[
\sum_{n>N_j} (\Delta_j - \Delta_{j-1}) e^{-\Gamma(j)n} \leq \int_{N_j}^{\infty} (\Delta_j - \Delta_{j-1}) e^{-\Gamma(j)x} dx = \frac{\Delta_j - \Delta_{j-1}}{\Gamma(j)} e^{-\Gamma(j)N_j}
\]

\[
\leq \frac{\Delta_j - \Delta_{j-1}}{\Gamma(j)} e^{-\Gamma(j)(3/\eta \Delta_j)^2} = 18 R^2 \left( \frac{\Delta_j - \Delta_{j-1}}{\Delta_j^2} \right) e^{-1/2 \eta^2 R^2}
\]

Likewise for the final sum.

\[
\sum_{n>N_2} \Delta_2 e^{-\Gamma(2)n} \leq \int_{N_2}^{\infty} \Delta_2 e^{-\Gamma(2)x} dx = \frac{\Delta_2}{\Gamma(2)} e^{-\Gamma(2)N_2}
\]

\[
\leq \frac{\Delta_2}{\Gamma(2)} e^{-\Gamma(2)(3/\eta \Delta_2)^2} = 18 R^2 \frac{e^{-1/2 \eta^2 R^2}}{\Delta_2}.
\]

Now combine lines (6) and (7) to get

\[
18 R^2 \left( \frac{1}{\Delta_2} + \frac{\Delta_3 - \Delta_2}{\Delta_3^2} + \ldots + \frac{\Delta_d - \Delta_{d-1}}{\Delta_d^2} \right) e^{1/2 \eta^2 R^2} \leq \frac{36 R^2}{\Delta_2} e^{-1/2 \eta^2 R^2}.
\]

where we have used Lemma 17. We conclude the second two sums in (5) are no more than twice the above.

Lemma 5 proves the main theorem under the assumption \( 0 = \Delta_1 < \Delta_2 < \ldots < \Delta_d \). For the more general theorem the proof is the same with \( d \) replaced with \( d' \) and \( \Delta_j \) replaced with \( \Delta(j) \) where \( 0 = \Delta(1) < \Delta(2) < \ldots < \Delta(d') \) are the distinct elements of \( \{ \Delta_j : j \leq d \} \).

**Theorem 2** Suppose the cost vectors \( a_1, a_2, \ldots \) are i.i.d with all \( \|a_i\| \leq L \) and \( \|a_i - a\| \leq R \) for \( \mathbb{E}[a_i] = a \). Suppose we run Algorithm 1 on the simplex with parameter \( \eta \). The pseudo-regret is no more than

\[
\Delta + \frac{18}{\Delta} \left( \frac{1}{\eta^2} + 4 R^2 e^{-1/2 \eta^2 R^2} \right)
\]

for \( \Delta = \max\{\Delta_j : i \leq d\} \) and \( \Delta = \min\{\Delta_j : \Delta_j > 0\} \).

The parameter \( \eta = 1/L \) gives an almost-optimal bound for the antagonistic setting. For that parameter the exponent in (1) is \( \frac{1}{2 \eta^2 R^2} = \frac{L^2}{2 R^2} \). Since Theorem 1 holds for any \( R \) with all \( \|a_i - a\| \leq R \) it holds for \( R = 2L \). Hence we can assume the exponent \( \frac{2L^2}{R^2} \geq \frac{1}{8} \) and we have the weaker but simpler bound.
Corollary 1. Under the hypotheses of Theorem 2 with parameter $\eta = 1/L$ the pseudo-regret satisfies

$$E \left[ \sum_{i=1}^{\infty} a \cdot (x_i - x^*) \right] \leq \Delta + \frac{18L^2 + 67R^2}{\Delta}$$

Since we can take $R = 2L$ and the first term has $\Delta = \max \{a \cdot (e_j - e_1) : j \leq d\} \leq a \cdot (e_j - e_1) \leq 2L$ we have the corollary.

Corollary 2. Suppose the cost vectors have all $\|a_i\| \leq L$. The anytime subgradient algorithm on the simplex with parameter $\eta = 1/L$ has pseudo-regret $O(L^2/\Delta)$ in the stochastic setting and regret $O(L\sqrt{N})$ in the stochastic setting.

Our main result is similar to that of Mourtada and Gaïffas [2019] for the Hedge (also called Exp3) algorithm. They assume the cost vectors lie in $[0,1]^d$ and show the pseudo-regret is $O(\log(d)/\Delta)$. While our bound is dimension-free it involves a bound on the 2-norm rather than the $\infty$-norm. To apply Theorem 2 to cost vectors $[0,1]^d$ we can only take $L = \sqrt{d}$ and get bound $O(dL^2/\Delta)$. This has a much stronger dependence than $\log(d)$. Hence the subgradient algorithm is more appropriate only if the cost vectors naturally arise from a sphere rather than a cube. In that case we can achieve pseudo-regret independent of the dimension.

4 Choosing the Parameter

Rather than choosing $\eta > 0$ to optimise the antagonistic bound in Theorem 1 we could try to optimise the stochastic bound in Theorem 2. We claim the the right-hand-side of (1) is decreasing in $\eta$ hence the bound is minimised as $\eta \to \infty$ and the algorithm becomes follow-the-leader. To prove this differentiate $\frac{1}{x} + 2R^2 e^{-1/2xR^2}$ to get

$$-\frac{1}{x^2} + \frac{e^{-1/2xR^2}}{x^2} = \frac{1}{x^2} \left(e^{-1/2xR^2} - 1\right).$$

Since the exponent is negative so is the derivative. Hence the function is decreasing and the bound (1) is no better than $72R^2/\Delta$.

Of course as $\eta \to \infty$ the worst-case bound in Theorem 1 becomes meaningless. Thus some compromise is required between the two bounds. One idea is to choose $\eta$ based on some estimate of how often the cost vectors will be stochastic and how often they will be antagonistic. This approach requires some bound $N$ on the time horizon since the increasing $O(L\sqrt{N})$ regret of an antagonistic problem will eventually outpace any finite number of stochastic problems with fixed $\Delta$. Likewise we need some lower bound on $\Delta$ since – for any fixed time horizon – there is a small enough $\Delta > 0$ that the $O(L^2/\Delta)$ stochastic bound exceeds the $O(L\sqrt{N})$ regret of any finite number of antagonistic problems.
In case these bounds are available we can choose $\eta$ to minimise the expression (8). The minimiser and minimum appear to have no simple closed form, and so $\eta$ must be computed numerically. Even dropping the exponential term, the minimiser is the solution to a cubic and the resulting bound is unenlightening. Theorem 3 gives a more readable bound on what performance we can expect after minimising (8).

**Theorem 3.** Suppose with probability $P$ the cost vectors $b_1, b_2, \ldots$ are generated antagonistically; and with probability $1 - P$ they come from an i.i.d sequence with $E[b_i] = b$ and $\|b_i - b\| \leq R$ and all $\Delta_j \geq \Delta$. Suppose we select $\eta > 0$ to minimise the expression (8) and run Algorithm 1 for $N$ turns. Define $R$ as the regret in the antagonistic case and pseudo-regret in the antagonistic case. Then we have

$$
\mathbb{E}[R] \leq \max \left\{ 4PL^2\sqrt{N}, \frac{72(1 - P)}{\Delta} + 2P\sqrt{N} \right\} + (2 - P)L + \frac{72}{\Delta} R^2 (1 - P) R^2.
$$

**Proof.** Using the bounds from Theorems 1 and 2 we see $\mathbb{E}[R]$ is at most

$$
P(L - \eta L^2) + P \left( \frac{1}{2\eta} + 2\eta L^2 \right) \sqrt{N} + (1 - P) \left( 2L + \frac{18}{\Delta} \left( \frac{1}{\eta^2} + 4R^2 e^{-1/2\eta^2 R^2} \right) \right)
$$

(8)

Drop the terms $PL + (1 - P)2L = (2 - P)L$ without a factor of $\eta$ and group the remaining terms by order of $\eta$ to get

$$
\frac{18(1 - P)/\Delta}{\eta^2} + \frac{P\sqrt{N}/2}{\eta} + 2PL^2(\sqrt{N} - 1)\eta + \frac{72R^2(1 - P)}{\Delta} e^{-1/2\eta^2 R^2}.
$$

This can of course be optimised numerically. To analyse algebraically round up the exponential term to 1 and then drop it since it doesn’t depend on $\eta$. Then we get

$$
\frac{18(1 - P)/\Delta}{\eta^2} + \frac{P\sqrt{N}/2}{\eta} + 2PL^2(\sqrt{N} - 1)\eta = \frac{A}{\eta^2} + \frac{B}{\eta} + C\eta = p(\eta).
$$

(9)

for the coefficients $A, B, C \geq 0$ determined by the left-hand-side. The derivative is

$$
p'(\eta) = \frac{C\eta^3 - B\eta - 2A}{\eta^3}
$$

which vanishes on the solution to a cubic equation. Since the cubic formula is hard to parse we consider some special cases. In the special case $\frac{A+B}{C} \geq 1$ then for $x \geq 1$ we have

$$
p(x) = \frac{A}{x^2} + \frac{B}{x} + Cx \leq \frac{A+B}{x} + Cx = q(x).
$$

The polynomial $q(x)$ is minimised at $\eta_1 = \sqrt{\frac{A+B}{C}}$ which is at least 1 by assumption. Since $p(1) = q(1)$ and $p'(x) \leq q'(x)$ for $x \geq 1$ we have $p \leq q$ over $[1, \infty)$. In particular

\[11\]
\[ p(\eta_0) \leq p(\eta_1) \leq q(\eta_1) = 2\sqrt{C(A+B)}. \] Since \( \frac{A+B}{C} \geq 1 \) we have \( C \leq A + B \) and \( p(\eta_0) \leq 2(A+B). \)

In the special case \( \frac{A+B}{C} \leq 1/2 \) then for \( x \leq 1 \) we have

\[ p(x) = \frac{A}{x^2} + \frac{B}{x} + Cx \leq \frac{A+B}{x^2} + Cx = r(x). \]

The polynomial \( r(x) \) is minimised at \( \eta_2 = \sqrt[3]{2(A+B)} \) which is at most 1 by assumption. Similar to before we have \( p(x) \leq r(x) \) over \([0,1]\). In particular

\[ p(\eta_0) \leq 2(\eta_0 + B). \]

To get the two terms in the max expans \( 2C \) and \( 4(A+B) \) respectively. For simplicity replace \( 2C = 4PL^2(\sqrt{N} - 1) \) with \( 4PL^2\sqrt{N} \). The two terms outside the max are the two constants we dropped earlier.

\[ \text{The only remaining case is } 1/2 \leq \frac{A+B}{C} \leq 1 \text{ observe } p(\eta) \leq \frac{\tilde{A}}{\eta} + \frac{\tilde{B}}{\eta} + C\eta \text{ for } \tilde{A} = 2A \text{ and } \tilde{B} = 2B. \] Since \( \frac{\tilde{A}+\tilde{B}}{C} \geq 1 \) the above says \( p(\eta_0) \leq 2(\tilde{A} + \tilde{B}) = 4(A + B) \).

The \( O(P\sqrt{N}) \) bound in Theorem 3 is the best we can expect if the cost vectors are generated antagonistically with probability \( P \).

The simulations in Section 7 indicate the coefficients 18 and 72 in Theorem 2 are far from optimal. Ideally the step size \( \eta \) should be chosen based instead on the smallest constants \( C_1, C_2 \) such that the pseudo-regret satisfies

\[ \mathbb{E} \left[ \sum_{i=1}^{\infty} a \cdot (x_i - x^*) \right] \leq \Delta + \frac{1}{\Delta} \left( \frac{C_1}{\eta^2} + C_2R^2e^{-1/2\eta^2R^2} \right) \]

Suppose we have obtained some \( D_1 \geq C_2 \) and \( D_2 \geq C_2 \) experimentally. For example Section 7 suggests \( C_1 = C_2 = 1 \) is reasonable to believe. We can repeat the proof of Theorem 11 with \( (8) \) replaced by

\[ P(L - \eta L^2) + P \left( \frac{1}{2\eta} + 2\eta L^2 \right) \sqrt{N} + (1 - P) \left( 2L + \frac{1}{\Delta} \left( \frac{D_1}{\eta^2} + D_2R^2e^{-1/2\eta^2R^2} \right) \right) \]

(10) to get the following.

**Theorem 4.** Let the hypotheses be the same as Theorem 3. Suppose \( D_1 \geq C_1 \) and \( D_2 \geq C_2 \) and we select \( \eta > 0 \) to minimise the expression (10). Then we have

\[ \mathbb{E}[\mathcal{R}] \leq \max \left\{ 4PL^2\sqrt{N}, \frac{4D_1(1-P)}{\Delta} + 2P\sqrt{N} \right\} + (2 - P)L + \frac{D_2(1-P)R^2}{\Delta}. \]
5 Beyond the Simplex

While the simplex is the natural setting for the Hedge algorithm, the advantage of subgradient methods is they can be applied to arbitrary compact convex domains in $\mathbb{R}^d$. This raises the question of what kinds of domains we can use to replace the simplex while keeping the bound from Theorem 2.

One natural counterpart to the simplex is the sphere. We are able to prove similar bound to Theorem 2 when the simplex is replaced with the sphere. However the proof is long and will be included in a later paper.

That said, Algorithm 1 certainly does not work for arbitrary domains. For example we will show the cubic domain $Y = \{(x, y) \in \mathbb{R}^2 : y \geq |x|^3 \text{ and } y \leq 1\}$ can give pseudo-regret $\Omega(\sqrt{N})$.

The crucial difference between the sphere and the cubic seems to be a curvature condition at the minimiser. The following example with the domain $Y_1 = \{(x, y) \in \mathbb{R}^2 : y \geq 0\}$ which is uncurved at the minimiser, gives $O(1)$ pseudo-regret. At the other end of the spectrum the domain $Y_2 = \{(x, y) \in \mathbb{R}^2 : y \geq |x| \text{ and } y \leq 1\}$ which is infinitely curved at the minimiser, gives $O(1)$ pseudo-regret by a similar argument to the simplex. Thus any suitable generalization of the curvature of the sphere will not be straightforward.

Example 1. Suppose we apply Algorithm 1 with parameter $\eta$ to the domain $Y$. There is a sequence $a_1, a_2, \ldots$ of i.i.d cost vectors with $\mathbb{E}[a_i] = a$ and

$$\sum_{i=1}^{N} a \cdot (x_i - x^*) \geq \Omega(\sqrt[4]{N}).$$

Proof. Let $B_1, B_2, \ldots$ be independent random variables with each $P(B_i = 1) = P(B_i = -1) = 1/2$. Then $B(n) = B_1 + \ldots + B_n$ is the position of a one dimensional random walk after $n$ steps. It is known that $\mathbb{E}|B(n)| \sim \sqrt{2n/\pi}$ for large $n$. Hence for some $M$ we have $\sqrt{n/\pi} \leq \mathbb{E}|B(n)| \leq \sqrt{3n/\pi}$ for all $n \geq M$.

Define the i.i.d sequence by $a = (0, 1)$ and each $a_n = (B_n, 1)$. According to the algorithm By definition $x_{n+1} = P_Y(-\eta a_1 + \ldots + a_n)$. Write the argument as

$$-\frac{a_1 + \ldots + a_n}{\sqrt{n}} = -\eta \sqrt{n} e_2 - \frac{\eta B(n)}{\sqrt{n}} e_1 = (\varepsilon, -c)$$

for $\varepsilon = -\eta B(n)\sqrt{n}$ and $c = \eta \sqrt{n}$. Write $x$ for the first component of $x_{n+1}$. The pseudo-regret $R_{n+1}$ for round $n+1$ is $|x|^2$. Lemma 8 says either $x^2 = 1$ or

$$x^2 \geq \frac{\varepsilon}{3c} - \frac{1}{3c} \left( 3 \left( \frac{\varepsilon}{3c} \right)^{5/2} + \sqrt{\frac{\varepsilon}{3c}} \right)$$

$$= \frac{|B(n)|}{3n} - \frac{1}{\eta \sqrt{n}} \left( \left( \frac{|B(n)|}{3n} \right)^{5/2} + 3 \sqrt{\frac{|B(n)|}{3n}} \right)$$

13
Since $|B(n)| \leq n$ the first term is no more than $1/3$. Hence the above holds for $x^2 = 1$ and so always holds.

The Jensen inequality says $E[R_{n+1}] = E|x^3| \geq E[x^2]^{3/2}$. Hence it is enough to lower-bound $E[x^2]$. For all $n > M$ we have for the first negative term

$$\left( \frac{|B(n)|}{3n} \right)^{5/2} \leq \frac{|B(n)|}{3n} \leq \frac{E|B(n)|}{3n} \leq \frac{\sqrt{3n/\pi}}{3n} = \frac{1}{\sqrt{3\pi n}} \leq \frac{1}{\sqrt{n}}$$

For the second negative term the Jensen inequality gives

$$E\left[ \sqrt{|B(n)|/3n} \right] \leq \sqrt{E|B(n)|/3n} \leq \sqrt{\sqrt{3n/\pi}/3n} \leq \frac{1}{\sqrt{3\pi n}} \leq \frac{1}{\sqrt{n}}$$

Hence for $n > M$ we have

$$E[x^2] \geq \frac{1}{3\eta \sqrt{\pi n}} - \frac{1}{\eta} \left( \frac{1}{n} \frac{3}{n^{3/4}} \right)$$

Since the first term dominates we can, increasing $M$ if necessary, assume $E[x^2] \geq \frac{1}{4\eta \sqrt{\pi n}}$ for all $n > M$. Then we have

$$E[R_{n+1}] = E|x^3| \geq E[x^2]^{3/2} \geq \left( \frac{1}{4\eta \sqrt{\pi n}} \right)^{3/2} \geq \frac{1}{23\eta} \frac{1}{n^{3/4}}.$$ 

Since

$$\sum_{i=M}^{N} \frac{1}{n^{3/4}} \geq \int_{M}^{N} \frac{1}{x^{3/4}} dx = 4(\sqrt{N} - \sqrt{M})$$

we see the total pseudo-regret has order $\Omega(\sqrt{N})$.

More generally we can take the domain $\mathcal{Y}_\alpha = \{(x,y) \in \mathbb{R}^2 : y \geq |x|^\alpha \text{ and } y \leq 1\}$ for any $\alpha > 2$. Then an analogous proof to the above shows the pseudo-regret has order $\Omega(N^{1-\frac{\alpha}{2(\alpha-1)}})$. Hence we get the following

**Lemma 7.** For each $\varepsilon > 0$ there exists a compact convex domain $\mathcal{Y} \subset \mathbb{R}^2$ and i.i.d sequence $a_1, a_2, \ldots$ of cost vectors with $E[a_i] = a$ such that running Algorithm 1 with any parameter $\eta > 0$ gives

$$\sum_{i=1}^{N} a \cdot (x_i - x^*) \geq \Omega(N^{1/2-\varepsilon}).$$

Thus the order of pseudo-regret can be almost as bad as the antagonistic regret.

Finally we prove the lemma mentioned in Example 1.
Lemma 8. Suppose $|\varepsilon| \in \mathbb{R}$ and $c > 0$. For $Y = \{(x, y) \in \mathbb{R}^2 : y \geq |x|^3$ and $y \leq 1\}$ we have $P_Y(\varepsilon, -c) = (x, |x|^3)$ for some $x \in [-1, 1]$ with either $x^2 = 1$ or

$$\frac{|\varepsilon|}{3c} - \frac{1}{3c} \left(3 \left(\frac{|\varepsilon|}{3c}\right)^{5/2} + \sqrt{\frac{|\varepsilon|}{3c}}\right) \leq x^2 \leq \frac{|\varepsilon|}{3c}.$$

Proof. For $|\varepsilon| = 0$ we have $P_Y(0, -c) = (0, 0)$ and the bound is trivial. Otherwise we can assume by symmetry that $\varepsilon > 0$. If $x^2 < 1$ then the line from $(\varepsilon, -c)$ to $P_Y(\varepsilon, -c) = (x, x^3)$ is normal to the graph. Since $\varepsilon > 0$ the projection has $x > 0$ and the normal is $\{(x, x^3 + t(3, -x^{-2}) : t \geq 0\}$. Since the line passes through $(\varepsilon, -c)$ we have for some $t \geq 0$ that $x + 3t = \varepsilon$ and $x^3 - t/x^2 = -c$. Solve to get

$$x^5 + cx^2 + x/3 = \varepsilon/3.$$

Since $x > 0$ the above says $x^2$ is at most $\varepsilon/3c$. Plugging in the value gives

$$(\varepsilon/3c)^{5/2} + cx^3 + (\varepsilon/27c)^{1/2} \geq \varepsilon/3$$

$$x^2 \geq \frac{1}{3c} \left(\varepsilon - 3 \left(\frac{\varepsilon}{3c}\right)^{5/2} - \sqrt{\frac{\varepsilon}{3c}}\right).$$

6 Tail Bounds

In this section we show the value $\sum_{i=1}^{\infty} a \cdot (x_i - e_1)$ is unlikely to stray too far from the expectation. Tail bounds are not a given as for example the original bandit algorithm of Flaxman et al. [2004] uses a sequence of random variables with fixed expectations with variances increasing to infinity. Thus the tail bounds get worse with time.

Throughout $0 = \Delta(1) < \Delta(2) < \ldots < \Delta(d')$ are the distinct elements of $\{\Delta_j : j \leq d\}$ in ascending order. We write $\Delta = \Delta(2) = \min\{\Delta_j : \Delta_j > 0\}$ and $\Delta = \Delta(d') = \max\{\Delta_j : j \leq d\}$. For our main theorem and throughout the section, our order bounds are with respect to the deviation $\delta$. The parameters $\eta, \Delta, \Delta$ are kept fixed. The main theorem follows.

Theorem 5. Suppose the cost vectors $a_1, a_2, \ldots$ are i.i.d with all $\|a_i\| \leq L$ and $\|a_i - a\| \leq R$ for $\mathbb{E}[a_i] = a$. Suppose we run Algorithm 1 with parameter $\eta$. For $C = \frac{\Delta}{36R^2} \left(1 - \left(\frac{\Delta}{4\Delta}\right)^2\right)$ there is a function $F$ of order $O(\delta e^{-C\delta})$ such that for each $\delta > 0$ we have

$$P \left(\sum_{i=1}^{\infty} a \cdot (x_n - e_1) > \delta + O(\sqrt{\delta})\right) \leq F(\delta)$$

In particular since $\Delta \geq \Delta$ we have

$$P \left(\sum_{i=1}^{\infty} a \cdot (x_n - e_1) > \delta + O(\sqrt{\delta})\right) \leq O \left(\delta \exp \left(\frac{\Delta}{41R^2}\delta\right)\right)$$
The remainder of the section is a proof of Theorem 5. To begin recall the formula from Lemma 4:

\[ P(a \cdot (x_{n+1} - e_1) > \Delta(j - 1)) \leq 2e^{-\Gamma(j)n} = 2 \exp \left( -\frac{\Delta(j)^2 n}{18R^2} \right) \]  

(11)

for all \( n > \left( \frac{3}{\eta \Delta(j)} \right)^2 \).

For any \( \alpha \geq 1 \) and \( j = d', \ldots, 3 \) define turns \( M(j) = \left\lceil \alpha \left( \frac{3}{\eta \Delta(j)} \right)^2 \right\rceil + 1 \) and the events

\[
E_j = \{a \cdot (x_{n+1} - e_1) \leq \Delta(j - 1) \text{ for } n = M(j), \ldots, M(j - 1) - 1\}
\]

\[
E_2 = \{a \cdot (x_{n+1} - e_1) = 0 \text{ for } n = M(2), M(2) + 1, \ldots\}
\]

These events force the pseudo-regret to be small.

**Lemma 9.** If the events \( E_{d'}, \ldots, E_2 \) occur the pseudo-regret over \( \{M(d') + 1, M(d') + 2, \ldots\} \) is at most

\[ \Delta + \frac{18\alpha}{\eta^2} \frac{1}{\Delta}. \]

**Proof.** If all the events hold the pseudo-regret on each turn in \( \{M(j) + 1, \ldots, M(j - 1)\} \) is at most \( \Delta(j - 1) \) and the pseudo-regret over \( \{M(2) + 1, M(2) + 2, \ldots\} \) is zero. It follows the pseudo-regret over \( \{M(d') + 1, M(d') + 2, \ldots\} \) is at most

\[
\sum_{j=2}^{d'-1} \Delta(j)(M(j) - M(j + 1))
\]

Lemma 18 says the above is no more than

\[
\Delta(d' - 1) + \frac{9\alpha}{\eta^2} \left\{ \sum_{j=2}^{d'-1} \Delta(j) \left( \frac{1}{\Delta(j)^2} - \frac{1}{\Delta(j + 1)^2} \right) \right\}
\]

\[
= \Delta(d' - 1) + \frac{9\alpha}{\eta^2} \left\{ \sum_{j=2}^{d'-1} \Delta(j) \left( \frac{1}{\Delta(j)^2} - \frac{1}{\Delta(j + 1)^2} \right) \right\}
\]

\[
= \Delta(d' - 1) + \frac{9\alpha}{\eta^2} \left( \frac{1}{\Delta(2)} + \frac{\Delta(3) - \Delta(2)}{\Delta(3)^2} + \ldots + \frac{\Delta(d' - 1) - \Delta(d' - 2)}{\Delta(d' - 1)^2} \right)
\]

\[
\leq \Delta(d' - 1) + \frac{9\alpha}{\eta^2} \left( \frac{1}{\Delta(2)} + \frac{\Delta(3) - \Delta(2)}{\Delta(3)^2} + \ldots + \frac{\Delta(d' - 1) - \Delta(d' - 2)}{\Delta(d' - 1)^2} \right)
\]

\[
\leq \Delta(d' - 1) + \frac{18\alpha}{\eta^2} \frac{1}{\Delta(2)} \leq \Delta + \frac{18\alpha}{\eta^2} \Delta.
\]

where the last line uses Lemma 17.

Next we bound the probability that \( E_{d'} \cap \ldots \cap E_2 \) occurs.
Lemma 10. For each \( \alpha \geq 1 \) we have

\[
P(E_{d'} \cap \ldots \cap E_2) \geq 1 - \left( \frac{18\alpha}{\eta^2 \Delta^2} + \frac{36R^2}{\Delta^2} \right) e^{-\alpha/2\eta^2 R^2}.
\]

Proof. Fix \( j = d', \ldots, 3 \) and take a union bound\(^2\) over \( n \) in formula (11) to see \( E_j \) fails with probability at most

\[
\sum_{i=M(j)}^{M(j)-1} 2e^{-\Gamma(j)n} \leq \sum_{i=M(j)}^{M(j)-1} 2e^{-\Gamma(j)M(j)} \leq \sum_{i=M(j)}^{M(j)-1} 2e^{-\alpha/2\eta^2 R^2} = 2(M(j) - M(j)) e^{-\alpha/2\eta^2 R^2}
\]

Likewise \( E_2 \) fails with probability at most

\[
2 \sum_{i=M(2)}^{\infty} e^{-\Gamma(2)n} \leq 2 \int_{M(2)-1}^{\infty} e^{-\Gamma(2)x} dx \leq \frac{36R^2e}{\Delta(2)^2} e^{-\Gamma(2)(M(2)-1)} \leq \frac{36R^2}{\Delta(2)^2} e^{-\alpha/2\eta^2 R^2}.
\]

Taking a union bound over \( j \) we see \( E = E_{d'} \cap \ldots \cap E_2 \) fails with probability at most

\[
\left( 2 \sum_{j=3}^{d'} (M(j) - M(j)) + \frac{36R^2}{\Delta(2)^2} \right) e^{-\alpha/2\eta^2 R^2} = \left( 2(M(2) - M(d')) + \frac{36R^2}{\Delta(2)^2} \right) e^{-\alpha/2\eta^2 R^2}
\]

Since \( M(d') \geq 2 \) and \( M(2) \leq \alpha \left( \frac{3}{\eta^2 \Delta^2} \right)^2 + 2 \) the above is no more than

\[
\left( \frac{18\alpha}{\eta^2 \Delta^2} + \frac{36R^2}{\Delta^2} \right) e^{-\alpha/2\eta^2 R^2}
\]

\( \square \)

Next we simplify one side of the bound from the previous two lemmas to get a bound over a final segment of the turns.

Lemma 11. For each \( \delta > 0 \) define \( \alpha = \frac{\delta \eta^2 \Delta}{18} + 1 \) and \( C = \Delta + \frac{18}{\eta^2 \Delta} \). Then we have

\[
P\left( \sum_{i=M(d')}^{\infty} a \cdot (x_n - e_1) > C + \delta \right) \leq \left( \frac{18}{\eta^2 \Delta^2} + \frac{\delta}{\Delta} + \frac{36R^2}{\Delta^2} \right) \exp \left( -\frac{\delta \Delta}{36R^2} - \frac{1}{2\eta^2 R^2} \right).
\]

Proof. For \( \alpha = \frac{\delta \eta^2 \Delta}{18} + 1 \) the bound in Lemma 9 is

\[
\Delta + \frac{18}{\eta^2 \Delta} + \delta = C + \delta.
\]

\(^2\)This union bound approach seems rather crude since the sequence of events that make up each \( E_j \) are not independent. For example if \( a \cdot (x_{n+1} - e_1) \leq \Delta(j - 1) \) occurs then since \( a \cdot \frac{x_{n+1} - e_1}{\sqrt{n+1}} \) is close to \( a \cdot \frac{x_n - e_1}{\sqrt{n+1}} \) we know \( x_{n+1} \) is close to \( x_n \) and the next inequality will be approximately true. We are unable to take advantage of this mathematically, but suspect the constant \( R^2/41\Delta \) in Theorem 6 is far from optimal. This is supported by Section 7.
The bound in Lemma 10 is
\[
\left( \frac{18}{\eta^2 \Delta^2} + \frac{\delta}{\Delta} + \frac{36R^2}{\Delta^2} \right) \exp \left( -\frac{\delta \Delta}{36R^2} - \frac{1}{2\eta^2 R^2} \right).
\]

Next we get a bound over the initial segment of the turns.

**Lemma 12.** For each \( \delta > 0 \) define \( \alpha = \delta \frac{\eta^2 \Delta}{18} + 1 \). We have
\[
P \left( \sum_{i=1}^{M(d')} a_i \cdot (x_i - e_1) > \delta + O(\sqrt{\delta}) \right) \leq O \left( \exp \left( -\frac{\Delta^2}{4\Delta R^2} \delta \right) \right)
\]

**Proof.** Theorem 1 says
\[
\sum_{i=1}^{M(d')} a_i \cdot (x_i - e_1) \leq \left( \frac{1}{2\eta} + 2\eta L^2 \right) \sqrt{M(d')}. \tag{12}
\]

Since \( M(d') \leq \alpha \frac{9}{\eta^2 \Delta^2} + 2 \) we have
\[
M(d') \leq \left( \delta \frac{\eta^2 \Delta}{18} + 1 \right) \frac{9}{\eta^2 \Delta^2} + 2 = \frac{\Delta}{2\Delta^2} \delta + \frac{9}{\eta^2 \Delta^2} + 2 \tag{13}
\]

Combining the above with (12) we have
\[
\sum_{i=1}^{M(d')} a_i \cdot (x_i - e_1) \leq \left( \frac{1}{2\eta} + 2\eta L^2 \right) \sqrt{\frac{\Delta}{2\Delta^2} \delta + \frac{9}{\eta^2 \Delta^2} + 2} = O(\sqrt{\delta}).
\]

The above is a regret bound and not a pseudo-regret bound. To put \( a \cdot (x_i - e_1) \) on the left-hand-side write
\[
\sum_{i=1}^{M(d')} a \cdot (x_i - e_1) \leq O(\sqrt{\delta}) + \sum_{i=1}^{M(d')} (a - a_i) \cdot (x_i - e_1).
\]

Since each \( x_i \) is decided based on \( a_1, \ldots, a_{i-1} \), the terms in the sum are independent. The terms are bounded by \( (a - a_i) \cdot (x_i - e_1) \leq \|a - a_i\|\|x_i - e_1\| \leq 2R \). By Hoeffding’s Inequality (Theorem 8) we have
\[
P \left( \sum_{i=1}^{M(d')} (a - a_i) \cdot (x_i - e_1) > t \right) \leq \exp \left( -\frac{t^2}{8M(d') R^2} \right). \tag{14}
\]
From (13) we know $M(d') \leq \frac{\Delta}{2\Delta^2} \delta + C$ for some constant $C$. By convexity of $1/x$ we have

$$\frac{1}{M(d')} = \frac{1}{\Delta \delta / 2 \Delta^2 + C} \geq \frac{1}{\Delta \delta / 2 \Delta^2} - \frac{C}{\Delta^2 \delta^2 / 4 \Delta^4} = \frac{2 \Delta^2}{\Delta \delta} - \frac{4 C \Delta^4}{\Delta^2 \delta^2}$$

For $t = \delta$ the exponent in (14) is

$$\frac{\delta^2}{8M(d')R^2} \geq \frac{\delta^2}{8R^2} \left( \frac{2 \Delta^2}{\Delta \delta} - \frac{4 C \Delta^4}{\Delta^2 \delta^2} \right) = \frac{\Delta^2 \delta}{4 \Delta R^2} - \frac{C \Delta^4}{2 \Delta^2 R^2}$$

giving

$$P \left( \sum_{i=1}^{M(d')} a_i \cdot (x_i - e_1) > \delta + O(\sqrt{\delta}) \right) \leq \exp \left( \frac{C \Delta^4}{2 \Delta^2 R^2} \right) \exp \left( -\frac{\Delta^2}{4 \Delta R^2} \delta \right)$$

$$\leq O \left( \exp \left( -\frac{\Delta^2}{4 \Delta R^2} \delta \right) \right).$$

\[ \square \]

Combining the bounds over the initial and final segments we get the main theorem.

**Theorem 5** Suppose the cost vectors $a_1, a_2, \ldots$ are i.i.d with all $\|a_i\| \leq L$ and $\|a_i - a\| \leq R$ for $E[a_i] = a$. Suppose we run Algorithm 1 with parameter $\eta$. For $C = \frac{\Delta}{36 R^2} \left( 1 - \left( \frac{\Delta}{3 \Delta} \right)^2 \right)$ there is a function $F$ of order $O(\delta e^{-C\delta})$ such that for each $\delta > 0$ we have

$$P \left( \sum_{i=1}^{\infty} a \cdot (x_n - e_1) > \delta + O(\sqrt{\delta}) \right) \leq F(\delta)$$

In particular since $\Delta \geq \Delta$ we have

$$P \left( \sum_{i=1}^{\infty} a \cdot (x_n - e_1) > \delta + O(\sqrt{\delta}) \right) \leq O \left( \delta \exp \left( \frac{\Delta}{41 R^2} \delta \right) \right)$$

**Proof.** Combining Lemma 11 and 12 we have for all $\delta, c > 0$ that

$$P \left( \sum_{i=1}^{\infty} a \cdot (x_n - e_1) > \delta + c\delta + O(\sqrt{\delta}) \right) \leq O \left( \delta \exp \left( -\frac{\Delta}{36 R^2} \delta \right) + \exp \left( -\frac{\Delta^2}{4 \Delta R^2} c\delta \right) \right)$$

For $c = \Delta^2/9 \Delta^2$ the right-hand-side becomes
\[ \delta \exp \left( -\frac{\Delta}{36R^2} \right) + \exp \left( -\frac{\Delta}{36R^2} \right) \leq O \left( \delta \exp \left( -\frac{\Delta}{36R^2} \right) \right). \]

Thus we have

\[ P \left( \sum_{i=1}^{\infty} a \cdot (x_n - e_1) > \left( 1 + \frac{\Delta^2}{9\Delta^2} \right) \delta + O(\sqrt{\delta}) \right) \leq O \left( \delta \exp \left( -\frac{\Delta}{36R^2} \right) \right) \]

Redefining \( \delta \) as \( 1 + \frac{\Delta^2}{9\Delta^2} \delta = \left( \frac{9\Delta^2 + \Delta^2}{9\Delta^2} \right) \delta \) we have

\[ P \left( \sum_{i=1}^{\infty} a \cdot (x_n - e_1) > \delta + O(\sqrt{\delta}) \right) \leq O \left( \delta \exp \left( -\frac{\Delta}{36R^2} \frac{9\Delta^2}{9\Delta^2 + \Delta^2} \right) \right). \]

To simplify the exponent use convexity to get

\[ \frac{1}{9\Delta^2 + \Delta^2} \geq \frac{1}{9\Delta^2} - \frac{\Delta^2}{81\Delta^4} \quad \frac{9\Delta^2}{9\Delta^2 + \Delta^2} \geq 1 - \frac{9\Delta^2 \Delta^2}{81\Delta^4} = 1 - \left( \frac{\Delta}{3\Delta} \right)^2. \]

Plugging back into the above we have

\[ P \left( \sum_{i=1}^{\infty} a \cdot (x_n - e_1) > \delta + O(\sqrt{\delta}) \right) \leq O \left( \delta e^{-C\delta} \right) \quad C = \frac{\Delta}{36R^2} \left( 1 - \left( \frac{\Delta}{3\Delta} \right)^2 \right) \]

\[ \square \]

### 7 Simulations

Here we plot the results of some simulations. We compare the constant coefficients in Theorem 2 to those observed empirically. Our simulations suggest the true constants are two orders of magnitude smaller than our theoretical bounds.

For each simulation we fix \( \Delta = \eta = 1 \). The i.i.d sequence \( a_1, a_2, \ldots, \in \mathbb{R}^d \) was generated as \( a_n = a + RN_n \) for \( N_1, N_2, \ldots \) drawn uniformly from the \((d-1)\)-dimensional unit sphere. Sampling on the unit sphere was done by drawing independent standard normals \( U_1, \ldots, U_d \) and normalising the vector \( (U_1, \ldots, U_d) \). See [Muller 1959](#) Section 4 for a proof of this method.

In Section 4 we showed the bound in Theorem 2 is optimised as \( \eta \to \infty \). Then the algorithm becomes follow-the-leader and the bound becomes \( \Delta + 72R^2/\Delta \). Even for \( \eta = 1 \) Figures 2 and 3 suggest \( \overline{\Delta} + 0.4R^2/\Delta \) for \( \overline{\Delta} = \frac{1}{d} \sum_{i=1}^{d} a_i \) is a more realistic bound.

Figures 2 and 3 also suggests higher dimensions regularise the data, lowering the mean and significantly lowering the variance.
Another observation is that — even for large noise levels — the behaviour seems to stabilise faster than the equations suggest. In Figures 2 and 3 we are interested in the quantity \( \sum_{i=1}^{\infty} a \cdot (x_i - e_1) \) but can only plot \( \sum_{i=1}^{N} a \cdot (x_i - e_1) \) for some finite \( N \). For large \( R \) more turns are required to approximate the infinite sum. For example suppose \( R = 10 \) and \( N > N_2 \). Lemma 5 gives the pseudo-regret bound over \( \{N_2 + 1, \ldots\} \) as

\[
2 \sum_{n>N_2}^{N} \Delta_2 e^{-\Gamma(2)n} + 2 \sum_{n>N_2}^{N} \sum_{k>2}^{d} (\Delta_k - \Delta_{k-1}) e^{-\Gamma(k)n}.
\]
In Lemma 6 line (7) we showed the first term is approximately
\[
\frac{36R^2}{\Delta}e^{-1/2\eta^2R^2} \simeq \frac{36R^2}{\Delta} = 3600.
\]
Consider the remainder beyond turn \(N\):
\[
2 \sum_{n>N}^\infty \Delta_2 e^{-\Gamma(2)n} = 2 \sum_{n>N}^\infty \exp\left(-\frac{n}{18R^2}\right) \simeq 2 \int_N^\infty \exp\left(-\frac{x}{18R^2}\right) \, dx = 36R^2 \exp\left(-\frac{N}{18}\right)
\]
\[
= 3600 \exp\left(-\frac{N}{1800}\right).
\]
To make the remainder less than 1% of the total the above suggests we must run for at least \(N = 8000\) turns. However Figures 4 and 5 suggest \(N = 500\) turns is enough for low dimensions and \(N = 100\) for higher dimensions.

The above simulations use \(a = (0,1,\ldots,1)\) because all other expectations we tried gave better performance. Two extreme cases are \(a = (0,1,2,\ldots,d-1)\) and \(a = (0,\ldots,0,1)\). The first gives moderately better performance in the long-run: The large cost on turn 1 and differences between arms makes the pseudo-regret stabilise faster and gives a steeper shoulder to the graph. The second gives significantly better performance.

Acknowledgements

This work was supported by Science Foundation Ireland grant 16/IA/4610.
Appendix A: Regret in the General Setting

Here we give the proof the subgradient algorithm with suitable parameter has regret $O(L\sqrt{N})$. The proof uses the techniques from Shalev-Shwartz [2012] modified slightly to not mention the time horizon.

**Theorem 1** For cost vectors $b_1, b_2, \ldots, b_N$ with all $\|b_i\| \leq L$ Algorithm 1 with parameter $\eta > 0$ has regret satisfying

$$\sum_{i=1}^{N} b_i \cdot (x_i - x^*) \leq (L\|\mathcal{X}\| - \eta L^2) + \left(\frac{1}{2\eta} \|\mathcal{X}\|^2 + 2\eta L^2\right) \sqrt{N}$$
for $\|\mathcal{X}\| = \max\{\|x\| : x \in \mathcal{X}\}$. In particular for $\mathcal{X} = \mathcal{S}$ and $\eta = 1/L$ we have

$$
\sum_{i=1}^{N} b_i \cdot (x_i - x^*) \leq \frac{5L}{4} \sqrt{N}.
$$

Proof. Define the functions $R_n(x) = \frac{\sqrt{n}}{2\eta} \|x\|^2$ for $n \geq 1$. First we show each $x_{n+1}$ is the unique minimiser of $\sum_{i=1}^{n} b_i + R_n(x)$. Since rescaling by a positive constant does not change the minimisers the function has the same minimisers as

$$
\|x\|^2 + 2\eta \frac{\sqrt{n}}{n} \sum_{i=1}^{n} b_i \cdot x = \left\| x + \frac{\eta}{\sqrt{n}} \sum_{i=1}^{n} b_i \right\|^2 - \frac{\eta^2}{n} \left( \sum_{i=1}^{n} b_i \right)^2.
$$

Since the last term is constant the above has global minimum at $x = -\frac{n}{\sqrt{n}} \sum_{i=1}^{n} b_i$. This is the point $y_{n+1}$ in the algorithm description. Lemma 13 says the minimum on $\mathcal{X}$ is the projection of the global minimum. Namely the point $x_{n+1} = P(y_{n+1})$ as required.

Now define $Q_2(x) = R_1(x) + b_1 \cdot x$ and $Q_{n+1}(x) = R_n(x) - R_{n-1}(x)$ for $n \geq 2$. Clearly each $Q_2 + \ldots + Q_n = b_1 \cdot x + R_{n-1}$. Lemma 3.1 of [Cesa-Bianchi and Lugosi 2006] says $\sum_{i=2}^{N} (b_i + Q_i) z_i \leq \sum_{i=2}^{N} (b_i + Q_i) x^*$ where $z_i$ are any minimisers of $\sum_{i=2}^{n} (b_i + Q_i)$ over $\mathcal{X}$ and $x^* \in \mathcal{X}$ is arbitrary. Unpack both sides, bring the $b_i$ terms to the left and telescope $Q_2 + \ldots + Q_n$ on the right to get

$$
\sum_{i=2}^{N} b_i \cdot (z_i - x^*) + R_1(x_1) + b_1 \cdot x_1 + \sum_{i=3}^{N} Q_i(z_i) \leq R_{N-1}(x^*)
$$

The $R_1$ and $Q_i$ terms on the left are positive so we can remove them while preserving the inequality to get

$$
b_1 \cdot x_1 + \sum_{i=2}^{N} b_i \cdot (z_i - x^*) \leq \frac{\sqrt{N-1}}{2\eta} \|x^*\|^2 \leq \frac{\sqrt{N}}{2\eta} \|\mathcal{X}\|^2
$$

To get the regret on the left-hand-side add $\sum_{i=2}^{N} b_i \cdot (x_i - z_i) - b_1 \cdot x^*$ to both sides to get

$$
\sum_{i=1}^{N} b_i \cdot (x_i - x^*) \leq \frac{\sqrt{N}}{2\eta} \|\mathcal{X}\|^2 + \sum_{i=2}^{N} b_i \cdot (x_i - z_i) - b_1 \cdot x^*. \tag{15}
$$

By Cauchy-Schwarz the last term is at most $\|b_1\| \|x^*\| \leq L \|\mathcal{X}\|$. To bound the sum on the right recall $z_n$ minimises $\sum_{i=2}^{n} (b_i + Q_i) = \sum_{i=1}^{n} b_i \cdot x + R_{n-1}$. Again since positive rescaling preserves the minimisers $z_n$ minimises

$$
\frac{2\eta}{\sqrt{n-1}} \sum_{i=2}^{n} (b_i + Q_i)(x) = \|x\|^2 + \frac{2\eta}{\sqrt{n-1}} \sum_{i=1}^{n} b_i \cdot x = \left\| x + \frac{\eta}{\sqrt{n-1}} \sum_{i=1}^{n} b_i \right\|^2 - \frac{\eta^2}{n-1} \left( \sum_{i=1}^{n} b_i \right)^2.
$$

24
Lemma 13 says $z_n = P_X \left( -\frac{\eta}{\sqrt{n-1}} \sum_{i=1}^{n} b_i \right)$. By definition $x_n = P_X \left( -\frac{\eta}{\sqrt{n-1}} \sum_{i=1}^{n-1} b_i \right)$ and so

$$\|x_n - z_n\| = \left\| P_X \left( -\frac{\eta}{\sqrt{n-1}} \sum_{i=1}^{n} b_i \right) - P_X \left( -\frac{\eta}{\sqrt{n-1}} \sum_{i=1}^{n-1} b_i \right) \right\|$$

$$\leq \left\| \frac{\eta}{\sqrt{n-1}} \sum_{i=1}^{n} b_i - \frac{\eta}{\sqrt{n-1}} \sum_{i=1}^{n-1} b_i \right\| = \frac{\eta}{\sqrt{n-1}} \|b_n\| \leq \frac{\eta L}{\sqrt{n-1}}$$

where the inequality uses Theorem 23 of [Nedic 2008]. Thus (15) becomes

$$\sum_{i=1}^{N} b_i \cdot (x_i - x^*) \leq \frac{\sqrt{N}}{2\eta} \|\mathcal{X}\|^2 + L\|\mathcal{X}\| + \eta L^2 \sum_{i=2}^{N-1} \frac{1}{\sqrt{n-1}}$$

$$= \frac{\sqrt{N}}{2\eta} \|\mathcal{X}\|^2 + L\|\mathcal{X}\| + \eta L^2 + \eta L^2 \sum_{i=2}^{N-1} \frac{1}{\sqrt{n}}. \quad (16)$$

Since the terms are decreasing the sum is bounded by the integral

$$\sum_{n=2}^{N-1} \frac{1}{\sqrt{n}} \leq \int_{1}^{N} \frac{dx}{\sqrt{x}} = 2\sqrt{N} - 2$$

and (16) simplifies to

$$\sum_{i=1}^{N} b_i \cdot (x_i - x^*) \leq (L\|\mathcal{X}\| - \eta L^2) + \left( \frac{1}{2\eta} \|\mathcal{X}\|^2 + 2\eta L^2 \right) \sqrt{N}.$$

For parameter $\eta = \frac{\|\mathcal{X}\|}{L}$ the first term vanishes and the right-hand-side simplifies to $\frac{5}{4}L\|\mathcal{X}\|\sqrt{N}$ For $\mathcal{X}$ the simplex $\|\mathcal{X}\| = 1$ and we get $\frac{5}{4}L\sqrt{N}$. \[\Box\]

**Appendix B: Convex Geometry**

Here we prove the convex geometry lemmas needed for the main analysis. The goal is Lemma 2 which gives a suitable condition for the projection of a point onto the simplex to have a tail of zeros. We prove the lemma in several stages.

**Lemma 13.** Suppose $\alpha \geq 0$ and $F(x) = \alpha \|x - v\|^2 + w$ is a quadratic function on $\mathbb{R}^d$ and $\mathcal{X} \subset \mathbb{R}^d$ convex. Then $\text{argmin}\{F(x) : x \in \mathcal{X}\} = P_X(v)$.

*Proof.* By definition $P_X(v) = \text{argmin}\{\|x - v\|^2 : x \in \mathcal{X}\}$. Since positive rescaling and adding a constant does not change the minimisers we have $P_X(v) = \text{argmin}\{\alpha \|x - v\|^2 + w : x \in \mathcal{X}\} = \text{argmin}\{F(x) : x \in \mathcal{X}\}$. \[\Box\]

Similar formulae to the next few results appear in [Wang and Carreira-Perpiñán 2013] and [Chen and Ye 2011].

25
Lemma 14. Let $S \subset \mathbb{R}^d$ be the $d$-simplex and $v = (v_1, \ldots, v_d) \in \mathbb{R}^d$ have $v_1 \geq v_2 \geq \ldots \geq v_d$. Suppose for some $n \leq d$ we have $\sum_{j=1}^{n} (v_j - v_n) \geq 1$. Then $P_S(v)$ has coordinate $n$ zero.

Proof. We use Lagrange multipliers to minimise $f(x) = \|x - v\|^2$ over $\Omega = \{x_j \geq 0 : j \leq d\}$ subject to $1 \cdot x - 1 = 0$. Call that point $x^*$. The Lagrangian is

$$L(\lambda, x) = \|x - v\|^2 + \lambda(1 \cdot x - 1) = \sum_{j=1}^{d} (x_j - v_j)^2 + \sum_{j=1}^{d} \lambda x_j - \lambda$$

$$= \sum_{j=1}^{d} (x_j - v_j + \lambda/2)^2 + \lambda \sum_{j=1}^{d} v_j - d\lambda^2/4 - \lambda$$

$$= \|x - v + (\lambda/2)1\|^2 + \lambda v - d\lambda^2/4 - \lambda.$$

We want to minimise $L(\lambda, x)$ over $x \in \Omega$. The minimum over $\mathbb{R}^d$ is $x = v - (\lambda/2)1$ and the previous lemma says the minimum $x(\lambda)$ over is $\Omega$ the projection onto $\Omega = \{x_j \geq 0 : j \leq d\}$. It can be checked $x_j(\lambda) = \max\{v_j - \lambda/2, 0\}$.

The method of Lagrange multipliers says $x^* = x(\lambda^*)$ where $\lambda^*$ maximises $L(\lambda, x(\lambda))$ over $\lambda \in \mathbb{R}$. If $\lambda^*/2 \geq v_n$ then $x_j(\lambda^*) = \max\{v_j - \lambda^*/2, 0\} = 0$ as required. To show we cannot have $\lambda^*/2 < v_n$ assume otherwise.

Then $\lambda \in [v_{m+1}, v_m)$ for some $m \geq n$. Observe over $[v_{m+1}, v_m)$ the functions $x_j(\lambda)$ are smooth. In particular $x_j(\lambda) = v_j - \lambda/2$ for $j \leq m$ and $x_j(\lambda) = 0$ for $j > m$. It follows that over $[v_{m+1}, v_m)$ we have

$$L(\lambda, x(\lambda)) = \sum_{j=1}^{d} (x_j(\lambda) - v_j + \lambda/2)^2 + \sum_{j=1}^{d} v_j \lambda - d\lambda^2/4 - \lambda$$

$$= \sum_{j=m+1}^{d} (v_j - \lambda/2)^2 + \sum_{j=1}^{d} v_j \lambda - d\lambda^2/4 - \lambda.$$

The function is differentiable with derivative.

$$\frac{d}{d\lambda} L(\lambda, x(\lambda)) = - \sum_{j=m+1}^{d} (v_j - \lambda/2) + \sum_{j=1}^{d} v_j - d\lambda/2 - 1$$

$$= (d - m)\lambda/2 + \sum_{j=1}^{m} v_j - d\lambda/2 - 1 = \sum_{j=1}^{m} v_j - m(\lambda/2) - 1.$$

Since $\lambda^*/2 \in [v_{n+1}, v_m)$ we have $\lambda^*/2 < v_m$ and

$$\frac{d}{d\lambda} L(\lambda^*, x(\lambda^*)) = \sum_{j=1}^{m} v_j - m(\lambda^*/2) - 1 > \sum_{j=1}^{m} v_j - mv_m - 1 = \sum_{j=1}^{m} (v_j - v_m) - 1.$$

Since $m \geq n$ the Lemma 13 says the right-hand-side is nonnegative. Hence the derivative at $\lambda^*$ is positive. Thus we can perturb $\lambda^*$ slightly to the right while remaining in $[v_{m+1}, v_m)$ and get a larger value of $L(\lambda, x(\lambda))$. This contradicts the definition of $\lambda^*$. \qed
Corollary 3. Let $S \subset \mathbb{R}^d$ be the $d$-simplex and $v = (v_1, \ldots, v_d) \in \mathbb{R}^d$ have $v_1 \geq v_2 \geq \ldots \geq v_d$ with not all coordinates equal. Then $\alpha = \sum_{j=1}^n (v_j - v_n)$ is positive and for all $\beta > 1/\alpha$ the projection $P_S(\beta v)$ has coordinate $n$ zero.

Lemma 15. Let $S \subset \mathbb{R}^d$ be the $d$-simplex and $v = (v_1, \ldots, v_d) \in \mathbb{R}^d$ have $v_1 \geq v_2 \geq \ldots \geq v_d$. Suppose for some $n \leq d$ we have $\sum_{j=1}^n (v_j - v_n) \geq 1$. Then we have $\sum_{j=1}^m (v_j - v_m) \geq 1$ for all $m \geq n$.

Proof. For $m = n + 1$ we can write

$$\sum_{j=1}^{n+1} (v_j - v_{n+1}) = \sum_{j=1}^{n+1} (v_j - v_n) + \sum_{j=1}^{n+1} (v_n - v_{n+1}) = \sum_{j=1}^{n+1} (v_j - v_n) + (n+1)(v_n - v_{n+1})$$

$$= \sum_{j=1}^n (v_j - v_n) + (v_{n+1} - v_n) + (n+1)(v_n - v_{n+1})$$

$$= \sum_{j=1}^n (v_j - v_n) + n(v_n - v_{n+1})$$

By assumption the first sum is at least 1. The second is nonnegative since the components are decreasing. By induction we see $\sum_{j=1}^m (v_j - v_n) \geq 1$ for all $m \geq n$. \qed

Lemma 2 Let $S \subset \mathbb{R}^d$ be the $d$-simplex. Suppose $w \in \mathbb{R}^d$ has two coordinates $k, \ell$ with $w_k - w_\ell \geq 1$. Then $P_S(w)$ has $\ell$-coordinate zero.

Proof. There is a permutation $\sigma$ with $w_{\sigma(1)} \geq w_{\sigma(2)} \geq \ldots \geq w_{\sigma(d)}$. Define the linear operator $\sigma : \mathbb{R}^d \to \mathbb{R}^d$ by $(x_1, \ldots, x_d) \mapsto (x_{\sigma(1)}, \ldots, x_{\sigma(d)})$. For $v = \sigma w$ observe each $w_i = v_{\sigma^{-1}(i)}$. Hence for $j = \sigma^{-1}(k)$ and $n = \sigma^{-1}(\ell)$ we have $v_j = w_k$ and $v_n = w_\ell$.

Since $w_k - w_\ell > 0$ we have $w_\ell < w_k$. Hence $v_n < v_j$ and since the coordinates of $v$ are decreasing $n > j$. Since all terms are nonnegative $\sum_{j=1}^n (v_j - v_n) \geq v_j - v_n = w_k - w_\ell \geq 1$.

Corollary 3 says $P_S(v) = P_S(\sigma w)$ has coordinate $n$ zero.

For $\varsigma = \sigma^{-1}$ the operator $\varsigma : \mathbb{R}^d \to \mathbb{R}^d$ preserves distances hence commutes with the projection. That means each $\varsigma P_S(x) = P_S(\varsigma x)$. Since $\varsigma$ maps the simplex onto itself we have $\varsigma P_S(x) = P_S(\varsigma x)$. For $x = \sigma w$ this gives $\varsigma P_S(\sigma w) = P_S(\varsigma w) = P_S(w)$.

Define $u = P_S(\sigma w)$. By the second paragraph $\varsigma P_S(\sigma w)$ has coordinate $\varsigma^{-1}(n)$ equal to $u_n = 0$. By the third paragraph so does $P_S(w)$. But by definition $\varsigma^{-1}(n) = \sigma(n) = \ell$ as required. \qed

Appendix C: Probability

Our main concentration result is due to [Pinelis 1994].
Theorem 6. (Pinelis Theorem 3.5) Suppose the martingale $f_1, \ldots, f_n$ takes values in the $(2, D)$-smooth Banach space $(E, \| \cdot \|)$. Suppose we have $\|f_1\|_\infty^2 + \sum_{i=2}^n \|f_i - f_{i-1}\|_\infty^2 \leq b^2$ for some constant $b$. Then for all $t \geq 0$ we have

$$P\left(\max\{\|f_1\|, \ldots, \|f_n\|\} \geq t\right) \leq 2 \exp\left(-\frac{t^2}{2D^2b^2}\right).$$

Here $\|f\|_\infty = \max\{\|f(x)\| : x \in \Omega\}$ is the sup norm taken over the probability space. The Banach space $(E, \| \cdot \|)$ is called $(2, D)$-smooth to mean $\|x + y\|_2 + \|x - y\|_2 \leq 2\|x\|_2 + 2D^2\|x\|_2$ for all $x, y \in E$. The fact that $\mathbb{R}^d$ is $(2, D)$-smooth is sometimes called the parallelogram law.

For brevity we omit the technical definition of a martingale. It is well known that if $a_1, a_2, \ldots$ are i.i.d with $\mathbb{E}[a_i] = a$ then $f_n = \sum_{i=1}^n (a_i - a)$ defines a martingale. If $\|a_i - a\| \leq R$ then taking $b^2 = nR^2$ and $t = r\sqrt{n}$ in the Pinelis theorem we have the following.

Theorem 7. Suppose the i.i.d sequence $a_1, a_2, \ldots$ takes values in $\mathbb{R}^d$. Suppose for $\mathbb{E}[a_i] = a$ we have $\|a_i - a\| \leq R$. Then for each $t \geq 0$ we have

$$P\left(\left\|\sum_{i=1}^n (a_i - a)\right\| \geq \sqrt{nr}\right) \leq 2 \exp\left(-\frac{r^2}{2R^2}\right).$$

Several times we use Hoeffding’s inequality for real-valued random variables to get one-sided bounds and avoid the leading factor of 2 in the Pinelis Theorem. See [Rigollet 2015] Theorem 1.9 for proof.

Theorem 8. (Hoeffding’s Inequality) Suppose $X_1, \ldots, X_n$ are independent and real valued with all $\mathbb{E}[X_i] = 0$ and $|X_i| \leq R$. Then we have

$$P\left(\sum_{i=1}^n X_i \geq t\right) \leq \exp\left(-\frac{t^2}{2R^2}\right).$$

The following fact about computing the expectation in terms of the CDF is well-known. But we were unable to find a suitably general proof in the literature.

Lemma 16. Suppose $X$ is a real-valued random variable. Then

$$\mathbb{E}[X] = \int_0^\infty P(X > x)dx - \int_-\infty^0 P(X \leq x)dx.$$  

In particular

$$\mathbb{E}[X] \leq \int_0^\infty P(X > x)dx.$$  

Proof. First assume $X$ takes only positive values. The second integral vanishes and we can write the first as

$$\int_0^\infty P(X > x)dx = \int_0^\infty \mathbb{E}_y[1_{X(y) > x}(y)] \, dx = \mathbb{E}_y\left[\int_0^\infty 1_{X(y) > x}(y) \, dx\right].$$
For fixed $y$ define the function $g(x) = 1_{X(y)>x}(y)$. We have $g(x) = 1$ for all $x > X(y)$ and $g(x) = 0$ elsewhere. Since $X(y)$ is nonnegative that means $g(x)$ is the indicator function of $[0, X(y))$. It follows the inner integral equals $X(y)$ and the above becomes $E_y[X(y)] = E[X]$. Observe the above also holds if we assume $X$ takes only nonnegative values and replace $P(X > x)$ with $P(X \geq x)$.

For a general random variable we can write $X = X^+ + X^-$ where $X^+$ takes only nonnegative values and $X^-$ only nonpositive values, and at each point one of $X^+$ or $X^-$ is zero. Since $-X^-$ is nonnegative we have already shown

$$E[-X^-] = \int_0^\infty P(-X^- \geq x) dx = \int_0^\infty P(X^- \leq -x) dx = \int_{-\infty}^0 P(X^- \leq x) dx$$

The left-hand-side is $-E[X^-]$. By construction $P(X^- \leq x) = P(X \leq x)$ for each $x \leq 0$. Hence the right-hand-side is $\int_{-\infty}^0 P(X \leq x) dx$. Finally write

$$E[X] = E[X^+] + E[X^-] = E[X^+] - E[-X^-] = \int_0^\infty P(X > x) dx - \int_{-\infty}^0 P(X \leq x) dx.$$

\[\Box\]

**Appendix D: Inequalities and Sums**

This section collects some ways to simplify sums that occur midway through the main analysis. In practice all the sums run over the vertices of the simplex.

**Lemma 17.** Consider the function

$$G(X_1, \ldots, X_m) = \frac{1}{X_1} + \frac{X_2 - X_1}{X_2^2} + \ldots + \frac{X_m - X_{m-1}}{X_m^2}$$

defined on \{ $X \in \mathbb{R}^m : X_1 \leq \ldots \leq X_m$ \}. We have $G(X) \leq 2/X_1$ for all allowed $X$.

**Proof.** Observe only the final term depends on $X_m$. Fix $X_1, \ldots, X_{m-1}$ and let $X_m$ vary over $X_m \geq X_{m-1}$. For $X_m = X_{m-1}$ or $X_m \to \infty$ the final term is zero. Since the term takes only positive values the maximum is achieved at some local maximum where the $X_m$-derivative vanishes. The derivative is $\frac{2X_{m-1} - X_m}{X_m^3}$ which vanishes only at $X_m = 2X_{m-1}$. At that point the last term is $1/4X_{m-1}$ which is less than $1/X_{m-1}$. Thus we have

$$G(X) \leq \frac{1}{X_1} + \frac{X_2 - X_1}{X_2^2} + \ldots + \frac{X_{m-1} - X_{m-2}}{X_{m-1}^2} + \frac{1}{X_{m-1}} = \frac{1}{X_1} + \frac{X_2 - X_1}{X_2^2} + \ldots + \frac{2X_{m-1} - X_{m-2}}{X_{m-1}^2}.$$

Like before only the final term depends on $X_{m-1}$. Fix $X_1, \ldots, X_{m-2}$ and let $X_{m-1}$ vary over $X_{m-1} \geq X_{m-2}$. For $X_{m-1} \to \pm \infty$ the last term goes to zero. For $X_{m-1} \to \pm \infty$ the last term goes to zero. Since the term is positive at least one value of $X_{m-1}$ it is
Thus we have shown at \( X \) Lemma 19. Suppose we have whole numbers maximised at a local maximum. The \( X_{m-1} \) derivative is \( \frac{2X_{m-2} - 2X_{m-1}}{X_{m-1}^2} \) which is zero only at \( X_{m-1} = X_{m-2} \). Plugging in that value we see the last term is no more than \( 1/X_{m-2} \). Thus we have shown

\[
G(X) \leq \frac{1}{X_1} + \frac{X_2 - X_1}{X_2^2} + \ldots + \frac{X_{m-2} - X_{m-3}}{X_{m-2}^2} + \frac{1}{X_{m-2}}.
\]

This is the same inequality as earlier but with one fewer terms. Repeating the process \( m-1 \) times we get \( G(X) \leq \frac{1}{X_1} + \frac{1}{X_1} = \frac{2}{X_1} \).

The next lemma simplifies an almost telescoping series that involves ceiling functions. Usually estimates \( \lceil n \rceil \leq n+1 \) are sufficient. However in our case we have a sum of \( d \) terms and want a bound independent of \( d \). This needs a more delicate approach.

**Lemma 18.** Suppose \( \Delta_2 \leq \ldots \leq \Delta_d \) and \( n_d \leq \ldots \leq n_2 \) are real numbers. Then we have

\[
\Delta_d \lceil n_d \rceil + \Delta_{d-1}(\lceil n_{d-1} \rceil - \lceil n_d \rceil) + \ldots + \Delta_2(\lceil n_2 \rceil - \lceil n_3 \rceil) \\
\leq \Delta_d + \Delta d_{n_d} + \Delta_{d-1}(n_{d-1} - n_d) + \ldots + \Delta_2(n_2 - n_3)
\]

**Proof.** Suppose we replace \( \lceil n_d \rceil \) with \( n_d \) on the left-hand-side. Since only the first two terms depend on \( n_d \) the difference is

\[
\left( \Delta_d \lceil n_d \rceil + \Delta_{d-1}(\lceil n_{d-1} \rceil - \lceil n_d \rceil) \right) - \left( \Delta_d n_d + \Delta_{d-1}(n_{d-1} - n_d) \right) \\
= \Delta_d (\lceil n_d \rceil - n_d) + \Delta_{d-1}(n_d - \lceil n_d \rceil) = (\Delta_d - \Delta_{d-1})(\lceil n_d \rceil - n_d).
\]

Since \( \Delta_d \geq \Delta_{d-1} \) and the factor on the right is at most 1 we see the difference above is at most \( \Delta_d - \Delta_{d-1} \).

Similarly replacing any \( \lceil n_i \rceil \) with \( n_i \) increases the value by at most \( \Delta_i - \Delta_{i-1} \). By induction it follows making all the replacements increases the value by at most \( (\Delta_d - \Delta_{d-1}) + (\Delta_{d-1} - \Delta_{d-2}) + \ldots (\Delta_3 - \Delta_2) = \Delta_d - \Delta_2 \geq \Delta_d \).

**Lemma 19.** Suppose we have whole numbers \( N_d < N_{d-1} < \ldots < N_2 \) and real numbers \( A(k, n) \) for each \( n \in \mathbb{N} \) and \( k = 3, \ldots, d \). Then

\[
\sum_{j=2}^{d-1} \sum_{n>N_{j+1}}^{N_j} \sum_{k>j}^d A(k, n) + \sum_{j=3}^d \sum_{n>N_2}^\infty A(j, n) = \sum_{j=3}^d \sum_{n>N_{j+1}}^\infty A(j, n).
\]

**Proof.** Expand the first sum by letting \( j \) vary.

\[
\sum_{j=2}^{d-1} \sum_{n>N_{j+1}}^{N_j} \sum_{k>j}^d A(k, n) = \sum_{n>N_3}^{N_2} (A(3, n) + \ldots + A(d, n)) + \sum_{n>N_4}^{N_3} (A(4, n) + \ldots + A(d, n)) \\
+ \ldots + \sum_{n>N_{d-2}}^{N_{d-1}} (A(d - 1, n) + A(d, n)) + \sum_{n>N_d}^{N_{d-1}} A(d, n)
\]
The term $A(d, n)$ appears only for $n$ in the intervals $\{N_d + 1, \ldots, N_{d-1}\}, \ldots, \{N_2 + 1, \ldots, N_2\}$. The union of these intervals is $\{N_d + 1, \ldots, N_2\}$. Likewise each $A(j, n)$ appears only for $n$ in $\{N_j + 1, \ldots, N_2\}$. Hence the above is $\sum_{d=3}^{d} \sum_{n=N_d}^{N_2} A(j, n)$ and the left-hand-side is exactly

$$
\sum_{d=3}^{d} \sum_{n>N_j}^{N_2} A(j, n) + \sum_{d=3}^{d} \sum_{n>N_2}^{\infty} A(j, n) = \sum_{d=3}^{d} \sum_{n>N_j}^{\infty} A(j, n).
$$

References

Peter Auer and Chao-Kai Chiang. An algorithm with nearly optimal pseudo-regret for both stochastic and adversarial bandits. *CoRR*, abs/1605.08722, 2016. URL [http://arxiv.org/abs/1605.08722](http://arxiv.org/abs/1605.08722).

Sébastien Bubeck and Nicolò Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *CoRR*, abs/1204.5721, 2012. URL [http://arxiv.org/abs/1204.5721](http://arxiv.org/abs/1204.5721).

Sébastien Bubeck and Aleksandrs Slivkins. The best of both worlds: stochastic and adversarial bandits. *CoRR*, abs/1202.4473, 2012. URL [http://arxiv.org/abs/1202.4473](http://arxiv.org/abs/1202.4473).

Nicolò Cesa-Bianchi and Gabor Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, New York, NY, USA, 2006. ISBN 0521841089.

Yunmei Chen and Xiaojing Ye. Projection Onto A Simplex. *arXiv e-prints*, art. arXiv:1101.6081, Jan 2011.

Abraham Flaxman, Adam Tauman Kalai, and H. Brendan McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. *CoRR*, cs.LG/0408007, 2004. URL [http://arxiv.org/abs/cs.LG/0408007](http://arxiv.org/abs/cs.LG/0408007).

Jaouad Mourtada and Stéphane Gaïffas. On the optimality of the Hedge algorithm in the stochastic regime. *Journal of Machine Learning Research*, 20:1–28, 2019.

Mervin E. Muller. A note on a method for generating points uniformly on n-dimensional spheres. *Commun. ACM*, 2(4):19–20, April 1959. ISSN 0001-0782. doi: 10.1145/377939.377946. URL [http://doi.acm.org/10.1145/377939.377946](http://doi.acm.org/10.1145/377939.377946).

Angelia Nedic. Convex Optimisation: Chapter 2. Fundamental Concepts in Convex Optimisation. 2008. URL [http://www.ifp.illinois.edu/~angelia/L5_exist_optimality.pdf](http://www.ifp.illinois.edu/~angelia/L5_exist_optimality.pdf).

Iosif Pinelis. Optimum bounds for the distributions of martingales in Banach spaces. *The Annals of Probability*, 22(4):1679–1706, 10 1994. doi: 10.1214/aop/1176988477. URL [https://doi.org/10.1214/aop/1176988477](https://doi.org/10.1214/aop/1176988477).
