## *Digression to boosting, experts, dense models, and their quantum counterparts*

There is a large collection of results across many fields such as:

- Duality in linear programming (Farkas [1902],Minkowski [1896])
- The Hahn-Banach theorem in functional analysis (Hahn [1927],Banach [1929])
- The minimax theorem in game theory (Neumann [1928])
- Regret minimization and expert learning (Hannan [1957], Littlestone and Warmuth [1989])
- Boosting in machine learning (Schapire [1990], Freund and Schapire [1995])
- The hard core lemma in computational complexity (Impagliazzo [1995])
- The dense model theorem in additive combinatorics (Green and Tao [2008],Tao and Ziegler [2008])

  that all share the following characteristics:

- They appear initially counterintuitive
- They are incredibly useful
- They are not that hard to prove once you gather the nerve to conjecture that they could be true. In fact, they can all proven by some kind of a *local search/improvements* type of algorithm such as *best response*, *multiplicative weights* or *gradient descent*.

To show optimality of sos we will need to use a result in this framework, and specifically the generalization of such results into the *quantum* or *positivesemidefinite* setting.

## *Regret minimization*

Consider the following setting. There is some universe $U$ of assets. An investor strategy can be thought of as a distribution $\mu$ over the assets (which we can think of as either describing the way to partition the portfolio or as describing how to probabilistically sample a single asset to invest in). At each time period $t$, the investor comes up with a distribution $\mu_t$, the universe comes up with a function $f_t \colon U \to [-1,+1]$ and profit to the investor is $\mathbb{E}_{x\sim\mu_t} f_t(x)$. In the setting of *regret minimization* (also known as expert learning) our goal is to come up with an investment strategy that would minimize the loss we suffer compared to the best *fixed* strategy in *hindsight* $\mu^*$. That is, we wish to find a way such that if for $t = 1, \dots, T$ we compute $\mu_t$ based on $f_0, \dots, f_{t-1}$ then we will minimize the maximum of

$$\sum_{t=1}^{t} \mathop{\mathbb{E}}_{\mu^*} f_t \;-\; \sum_{t=1}^{T} \mathop{\mathbb{E}}_{\mu_t} f_t \tag{1}$$

over all distributions $\mu^*$ over $U$.

The basic result in this area is the following:

**1. Theorem (Regret minimization).** *For every parameter $\eta$, and every choice of $f_1, \dots, f_t$ and distribution $\mu^*$ we can choose $\mu_t$ based only on $f_1, \dots, f_{t-1}$ such that*

$$\sum_{t=1}^{T} \mathop{\mathbb{E}}_{\mu_t} f_t \le (1 + O(\eta))\Big[\sum_{t=1}^{T} \mathop{\mathbb{E}}_{\mu^*} f_t\Big] + \tfrac{1}{\eta}\Delta(\mu^* \| \mu_1) \tag{2}$$

*where $\Delta(\mu' \| \mu)$ denotes the KL divergence of $\mu'$ from $\mu$.*

In particular if we set $\mu_1$ to be the uniform distribution, then since $\Delta(\mu^* \| \mu_1) \le \log |U|$ we cam set $\eta$ to be $\sqrt{\log |U|}/T$ and get that the total regret is bounded by $O(\sqrt{T \log |U|})$ which (for $T \gg \log |U|$) is sublinear in $T$.

*Proof.* We are going to simply let $\mu_{t+1}(x)$ be eaual to to $Z_t \mu_t(x) 2^{\eta f_t(x)}$ where $Z_t = \left(\mathbb{E}_{\mu_t} 2^{\eta f_t(x)}\right)^{-1}$ is a normalization factor.

Now let us upper bound the decrease in distance between $\mu^*$ and our current distribution by something related to the loss we suffer compared to the optimum:

$$\Delta(\mu^* \| \mu_{t+1}) - \Delta(\mu^* \| \mu_t) = \mathop{\mathbb{E}}_{x \sim \mu^*} \log \left(\tfrac{\mu^*(x)}{\mu_{t+1}(x)}\right) - \mathop{\mathbb{E}}_{x \sim \mu^*} \left(\tfrac{\mu^*(x)}{\mu_t(x)}\right) \tag{3}$$

which equals

$$\mathop{\mathbb{E}}_{x \sim \mu^*} \left(\tfrac{\mu_t(x)}{\mu_{t+1}(x)}\right) = \mathop{\mathbb{E}}_{x \sim \mu^*} \log\left(\tfrac{1}{Z_t 2^{\eta f_t(x)}}\right) = \log Z_t^{-1} - \eta \mathop{\mathbb{E}}_{\mu^*} f_t \,. \tag{4}$$

but since $Z_t^{-1} = \mathbb{E}_{\mu_t} 2^{\eta f_t} \le (\eta - O(\eta^2))\, \mathbb{E}_{\mu_t} f_t$ we get

$$\Delta(\mu^* \| \mu_{t+1}) - \Delta(\mu^* \| \mu_t) \le \eta\big((1 - \eta) \mathop{\mathbb{E}}_{\mu_t} f_t - \mathop{\mathbb{E}}_{\mu^*} f_t\big) \,. \tag{5}$$

The telescopic sum of Eq. (5) over all $t$ from 1 to $T$ yields the theorem. □

The proof of Theorem 1 actually yields more than just the statement. In particular the following two points are important:

- Our strategies $\mu_1, \ldots, \mu_t$ are *simple* in the sense that they are composed of the initial prior $\mu_1$ reweighed by "few" of the functions $f : U \to [0,1]$.
- The complexity of our strategy can be controlled by the KL distance from our prior to the optimal distribution. That is, if there were only few bits of information that we were missing, then there is a simple strategy that is nearly optimal.

This is a general (and very useful) phenomena that **simple tests can be fooled by simple distributions** (see, e.g., Trevisan et al. [2009]). In particular, the above proof establishes the following theorem:

**2. Theorem (Simple tests can be fooled by simple distributions: classical version).** *Let $\mathcal{F}$ be a collection of test functions mapping some universe $U$ to $[-1,1]$, and let $\mu_{opt}, \mu_{prior}$ be some distribution over $U$. Then there exists a distribution $\mu$ such that*

$$\mathop{\mathbb{E}}_{\mu_{opt}} f - \mathop{\mathbb{E}}_{\mu_{prior}} f < \epsilon \tag{6}$$

*for every $f \in \mathcal{F}$ and $\mu$ is* simple *in the sense that it is obtained by reweighing $\mu_{prior}$ using a function proportional to $\varepsilon^{\sum_{i=1}^{t} f_i}$ where $t = \Delta(\mu_{opt} \| \mu_{prior})\,\mathrm{poly}(1/\epsilon)$.*

## *Quantum version*

We can extend the above observations to the *quantum setting*. Suppose that now the investor strategy is a *quantum state* $\rho$ (i.e., psd matrix of trace 1) on a system with the universe of states $U$, and the gain is now the probability that $\rho$ passes some *measurement* which is a $|U| \times |U|$ matrix $M$ satisfying $0 \preceq M \preceq I$. The same algorithm works where we now use $\rho_{t+1}$ as proportional to $\rho_t e^{\eta M_t}$. This is known as the *matrix multiplicative weights* algorithm (e.g., see Arora et al. [2012]). A matrix exponential can be computed using the power series for the exponential (or by keeping the same eigenbasis and exponentiating the eigenvalues). The same analysis works except that we now replace the KL divergence of $\mu_{opt}$ and $\mu_{prior}$ by the corresponding *quantum relative entropy* which corresponds to the *von Neumann entropy* of the states. In particular we can get the following result:

**3. Theorem (Simple tests can be fooled by simple distributions: classical version).** *Let $\mathcal{F}$ be a collection of* quantum measurements *over a system with universe of states $U$, where for every $f \in \mathcal{F}$, $-\mathrm{Id}_U \preceq f \preceq +\mathrm{Id}_U$ where $\mathrm{Id}_U$ denotes the $|U| \times |U|$ identity matrix and $\preceq$ denotes spectral domination. Let $\rho_{opt}, \rho_{prior}$ be two* density matrices *over $U$ (i.e., $|U| \times |U|$ psd matrices with trace 1). Then there exists a density $\rho$ such that*

$$\mathrm{Tr}(\rho f^*) - \mathrm{Tr}(\rho_{opt} f^*) < \epsilon \quad (7)$$

*for every $f \in \mathcal{F}$ and $\rho$ is* simple *in the sense that it is obtained by reweighing $\rho_{prior}$ using a function proportional to $\varepsilon^{\sum_{i=1}^{t} f_i}$ where $t = \Delta(\rho_{opt} \| \rho_{prior})\,\mathrm{poly}(1/\epsilon)$ and $\Delta(\rho \| \sigma)$ denotes the* quantum relative entropy $\mathrm{Tr}(\rho(\log \rho - \log \sigma))$.

## *References*


Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(6):121–164, 2012. doi: 10.4086/toc.2012.v008a006. URL http://www.theoryofcomputing.org/articles/v008a006.

Stefan Banach. Sur les fonctionnelles linéaires ii. *Studia Mathematica*, 1 (1):223–239, 1929.

Julius Farkas. Theorie der einfachen ungleichungen. *Journal für die reine und angewandte Mathematik*, 124:1–27, 1902.

Yoav Freund and Robert E Schapire. A desicion-theoretic generalization of on-line learning and an application to boosting. In *European conference on computational learning theory*, pages 23–37. Springer, 1995.

Ben Green and Terence Tao. The primes contain arbitrarily long arithmetic progressions. *Annals of Mathematics*, pages 481–547, 2008.

Hans Hahn. Über lineare gleichungssysteme in linearen räumen. *Journal für die reine und angewandte Mathematik*, 157:214–229, 1927.

James Hannan. Approximation to bayes risk in repeated play. *Contributions to the Theory of Games*, 3:97–139, 1957.

Russell Impagliazzo. Hard-core distributions for somewhat hard problems. In *FOCS*, pages 538–545. IEEE Computer Society, 1995.

Nick Littlestone and Manfred K. Warmuth. The weighted majority algorithm. In *FOCS*, pages 256–261. IEEE Computer Society, 1989.

Hermann Minkowski. Geometrie der zahlen. *Teubner, Leipzig*, 1, 1896.

J v Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, 1928.

Robert E Schapire. The strength of weak learnability. *Machine learning*, 5(2):197–227, 1990.

Terence Tao and Tamar Ziegler. The primes contain arbitrarily long polynomial progressions. *Acta Mathematica*, 201(2):213–305, 2008.

Luca Trevisan, Madhur Tulsiani, and Salil P. Vadhan. Regularity, boosting, and efficiently simulating every high-entropy distribution. In *IEEE Conference on Computational Complexity*, pages 126–136. IEEE Computer Society, 2009.