# Introduction to Statistical Learning Theory

## Lecture 11

Last time we looked at FoReL

**Algorithm** `Follow-The-Regularized-Leader (FoReL)`

```
**Input:** Convex set $\mathcal{S}$, regularization $R$.
**for** t=1,...,T **do**
    predict $w_t = \arg\min_{w \in \mathcal{S}} R(w) + \sum_{i=1}^{t-1} f_t(w)$
**end for**
```

and we have seen a regret bound

**Theorem 1.1**

*Let $w_1, \dots, w_T, \dots$ be the outputs of the FoReL algorithm. For all $u \in \mathcal{S}$ we have*

$$Regret_T(u) \leq R(u) - R(w_1) + \sum_{t=1}^{T} \left(f_t(w_t) - f_t(w_{t+1})\right)$$

We will see more concrete regret bounds for strongly convex regularizers.

### Definition 1.2

Let $\mathcal{S}$ be a convex set and $f : \mathcal{S} \to \mathbb{R}$. The function $f$ is $\sigma$-strongly convex over $\mathcal{S}$ with respect to the norm $|| \cdot ||$ if for all $w \in \mathcal{S}$ we have

$$f(u) \geq f(w) + \langle z, u - w \rangle + \frac{\sigma}{2} ||u - w||^2$$

for all $z \in \partial f(w)$.

This is not exactly how we defined it earlier, but is equivalent.

An intimidate corollary of this definition - If $f$ is $\sigma$-strongly convex and $w = \arg\min_{v \in \mathcal{S}} f(v)$ then $f(u) \geq f(w) + \frac{\sigma}{2} ||u - w||^2$.

For twice differentiable functions $f$, it is enough to show $x^T \nabla^2 f(w) x \geq \sigma ||x||^2$ for strong convexity (Taylor).

## Theorem 1.3

*Let $f_1, .., f_t$ be convex functions such that $f_t$ is $L_t$-Lipschitz with respect to a norm $||\cdot||$. Let $L$ satisfy $\frac{1}{T}\sum_{t=1}^{T} L_t^2 \leq L^2$. If the FoReL is run with a $\sigma$-strongly convex regularizer $R$ with respect to the same norm, then for all $u \in \mathcal{S}$,*

$$Regret_T(u) \leq R(u) - \min_{v\in\mathcal{S}} R(v) + TL^2/\sigma$$

Proof - Using Lipschitz-ness of $f_t$ we have
$f_t(w_t) - f_t(w_{t+1}) \leq L_t||w_t - w_{t+1}||$.

We can define $F_t(w) = \sum_{i=1}^{t-1} f_i(w) + R(w)$. So by definition
$w_t = \arg\min_{w\in\mathcal{S}} F_t(w)$.

Using strong convexity we get

$$F_t(w_{t+1}) \geq F_t(w_t) + \frac{\sigma}{2}||w_t - w_{t+1}||^2$$

and similarly

$$F_{t+1}(w_t) \geq F_{t+1}(w_{t+1}) + \frac{\sigma}{2}||w_t - w_{t+1}||^2$$

Summing both inequalities and removing duplicates we get

$$f_t(w_t) - f_t(w_{t+1}) \geq \sigma||w_t - w_{t+1}||^2$$

We can now conclude that $||w_t - w_{t+1}|| \leq L_t/\sigma$ to finish the proof. □

We know that $R(w) = \frac{\sigma}{2}||w||^2$ is $\sigma$-strongly convex with regard to $\ell_2$ norm.

We will now show another regularizer that is strongly convex with respect to $\ell_1$ norm on the probability simplex.

Define $R(w) = \sum_{i=1}^{d} w[i]\log(w[i])$, then it is easy to show that $\frac{\partial^2 R}{\partial w[i]\partial w[j]} = \delta_{ij}\frac{1}{w[i]}$. we now have -

$$\begin{aligned} x^T\nabla^2 R(w)x = & \sum_i \frac{x[i]^2}{w[i]} = \left(\sum_i w[i]\right)\left(\sum_i \frac{x^2[i]}{w[i]}\right) \geq \left\langle \sqrt{w}, \frac{|x|}{\sqrt{w}} \right\rangle^2 \\ = & ||x||_1^2 \end{aligned}$$

We conclude that the entropic regularizer is 1-strongly convex over the $||\cdot||_1$ norm.

The prediction with expert advice framework is the following:

At each round the learner receives advice from each of the $d$ experts, he then chooses an expert $p_t$ and finally receives a loss vector $y_t \in [0,1]^d$ and suffers loss $y_t[p_t]$.

As we have seen before, it is impossible to get sublinear regret bounds. The solution - convexification by randomization. The learner will return a distribution $p_t \in \Delta_d$, and suffer loss $\mathbb{E}_{p_t}[y_t] = \sum_i p_t[i]y_t[i] = \langle p_t, y_t \rangle$.

Consider running FoReL with $R(w) = \frac{\sigma}{2}||w||_2^2$ regularization. We have $f_t(w) = \langle y_t, w \rangle$ is $||y_t||_2$-Lipschitz, so $L = \sqrt{d}$. The regert bound is

$$Regret_T(u) \leq \frac{\sigma}{2}||u||_2^2 + dT/\sigma$$

As $u \in \Delta_d$ then $||u||_2 \leq ||u||_1 \leq 1$ so by fixing the optimal learning rate $\sigma = \sqrt{2dT}$ we get

$$Regret_T(\Delta_d) \leq \sqrt{2dT}$$

Can we do better? What about entropic regularization $R(w) = \sigma \sum w[i] \log(w[i])$?

To get a regret bound we first need to bound $R(u) - \min_{v \in \Delta_d} R(v)$. It is easy to see that $R(u) \leq 0$, and with Lagrange multipliers we can get $\min_{v \in \Delta_d} R(v) = -\log(d)$.

Next we notice that $|f_t(w) - f_t(u)| = | \langle w - u, y_t \rangle | \leq ||w - u||_1 ||y_t||_\infty$ from the Holder inequality. As $||y_t||_\infty = 1$, we get that $f_t$ is 1-Lipschitz with respect to the $|| \cdot ||_1$ norm.

Pluging everything in the formula we get $Regret_T(\Delta_d) \leq \sigma \log(d) + T/\sigma$ and with $\sigma = \sqrt{\frac{T}{\log(d)}}$ we get $Regret_T(\Delta_d) \leq 2\sqrt{\log(d) T}$ which scales much better with $d$.

How does the entropic update look like?

We have $w_t = \arg\min_{w \in \Delta_d} \sum_{i=1}^{t-1} \langle y_i, w \rangle + \sigma \sum_{j=1}^{d} w[i] \log(w[i])$. Consider the constraint $\sum_i w[i] = 1$, then by Lagrange multipliers we have

$$\sum_{i=1}^{t-1} y_i[j] + \sigma \log(w[j]) + 1 = \lambda$$

Which means that $w_t[j] \propto \exp(-\sum_{i=1}^{t-1} y_i[j]/\sigma)$ and we get the simple update rule $w_t[j] = \frac{w_{t-1}[j] \exp(-y_t[j]/\sigma)}{\sum_{i=1}^{d} w_{t-1}[i] \exp(-y_t[i]/\sigma)}$.

This algorithm is called exponentiated gradient descend.

Last week we have seen that the realizable case (binary classification) is controlled by the Littlestone dimension. We also presented an optimal learning algorithm for that case.

**Algorithm** `Standard Optimal Algorithm`

```
Input: Hypothesis space $\mathcal{H}$
Initialize: $V_1 = \mathcal{H}$
for t=1,...,T do
    recieve $x_t$
    for $b \in \{\pm 1\}$ set $V_t^b = \{h \in V_t : h(x_t) = b\}$
    predict $p_t = \arg\max_{b \in \{\pm 1\}} LDim(V_t^b)$
    recieve $y_t$
    Update: $V_{t+1} = \{h \in V_t : h(x_t) = y_t\}$ .
end for
```

We will use the learning with expect advice to show that $\mathcal{O}(\sqrt{Ldim(\mathcal{H})\log(T)T})$ regret can be achieved in the general case.

The idea - we will simulate every $h \in \mathcal{H}$ using a finite set of expects.

Each expert will perform a SOA run which fails at time steps $t_1, ..., t_L$ for $L \leq LDim(\mathcal{H})$

**Algorithm** `Expert`$(t_1, ..., t_L)$

```
Input: Hypothesis space $\mathcal{H}$
Initialize: $V_1 = \mathcal{H}$
for t=1,...,T do
  recieve $x_t$
  for $b \in \{\pm 1\}$ set $V_t^b = \{h \in V_t : h(x_t) = b\}$
  define $\tilde{y}_t = \arg\max_{b \in \{\pm 1\}} LDim(V_t^b)$
  if $t \in \{t_1, ..., t_L\}$ then
    predict $y_t = 1 - \tilde{y}_t$
  else
    predict $y_t = \tilde{y}_t$
  end if
  Update: $V_{t+1} = \{h \in V_t : h(x_t) = y_t\}$ .
end for
```

## Lemma 2.1

*Let $\mathcal{H}$ be a hypothesis class with Ldim$< \infty$. For each $h \in \mathcal{H}$ and inputs $x_1, ..., x_T$ there exists $L \leq Ldim(\mathcal{H})$ and indices $t_1, ..., t_L$ such that $Expert(t_1, ..., t_L)[x_t] = h(x_t)$*

Proof - Consider running the SOA on inputs $(x_1, h(x_1)), ..., (x_T, h(x_T))$. The SOA will err at at most $L \leq Ldim(\mathcal{H})$ indices $t_1, ..., t_L$.

The expert $Expert(t_1, ..., t_L)$ agrees with the SOA except on $t_1, ..., t_L$ so he agrees with $h$ on every input.

The immediate corollary - the regret compared to each expert is the same regret compared to $\mathcal{H}$.

## Littlestone dimension

We need $d = \sum_{L=1}^{Ldim} \binom{T}{L}$ experts to simulated $\mathcal{H}$.

Using the Saur-Shelach lemma we get $d \leq \left(\frac{eT}{Ldim(\mathcal{H})}\right)^{Ldim(\mathcal{H})}$

The regret is bounded by

$$\sqrt{2\log(d)T} \leq \sqrt{2Ldim(\mathcal{H})T\log(T)}$$

for $Ldim(\mathcal{H}) \geq 3$.

We now present a variation of PEA with partial information.

Consider $d$ slot machines, with losses at step $t$ of $y_t[i]$.

At each step we pick machine $p_t$ and suffer loss $y_t[p_t]$.

The difference - we are only given $y_t[p_t]$ afterwards and not the whole $y_t$ vector.

This means we cannot do exact gradient descend.

As before we randomize - we predict a distribution $w_t \in \Delta_d$, and pick $p_t \sim w_t$.

Our basic tool (which we will not prove) will be based on local norms $||z||_t \equiv \sqrt{\sum w_t[i]z[i]^2}$

## Theorem 3.1

*Assume you run the exponentiated gradient algorithm with linear loss* $\langle z_t, w\rangle$ *such that* $z_t[i]/\sigma > -1$*, then the following holds for all* $u \in \Delta_d$

$$\sum_{t=1}^{T} \langle w_t - u, z_t\rangle \leq \sigma \log(d) + \sum_{t=1}^{T}\sum_{i=1}^{d} z_t[i]^2 w_t[i]/\sigma$$

We do not know the real gradient $z_t$, but we can estimate it.

Consider the distribution $w_t$ and $p_t \sim w_t$. We can set $z_t^{p_t}[p_t] = z_t[p_t]/w_t[p_t]$ and zero otherwise. We get that

$$\mathbb{E}[z_t^{p_t}[j] | z_1, ..., z_{t-1}] = \sum_{i=1}^{d} P[p_t = i] z_t^i[j] = w_t[j] z_t^j[j] = w_t[j] \frac{y_t[j]}{w_t[j]} = y_t[j]$$

This means that our $z_t^{p_t}$ is an unbiased estimator.

**Algorithm** `Multi-armed bandit algorithm`

```
**Initialize:** $w_1 = (1/d, ..., 1/d)$
Pick bandit $p_t$ from distribution $w_t$
Recieve loss $y_t[p_t] \in [0, 1]$
**Update:**
For $i \neq p_t$ $\tilde{w}[i] = w_t[i]$
$\tilde{w}[p_t] = w_t[p_t] \exp(-y_t[p_t]/\sigma w_t[p_t])$
$w_{t+1} = \tilde{w}/||\tilde{w}||_1$.
```

To bound the regret we first use the inequality

$$\sum_{t=1}^{T} \langle w_t - u, z_t^{p_t} \rangle \leq \sigma \log(d) + \sum_{t=1}^{T} \sum_{i=1}^{d} z_t^{p_t}[i]^2 w_t[i]/\sigma$$

We can take expectation on both side and notice that

$$\mathbb{E}[\langle w_t - u, z_t^{p_t} \rangle] = \mathbb{E}[\mathbb{E}[\langle w_t - u, z_t^{p_t} \rangle \, | z_1, ..., z_{t-1}]]$$
$$= \mathbb{E}[\langle w_t - u, \mathbb{E}[z_t^{p_t} | z_1, ..., z_{t-1}] \rangle] = \mathbb{E}[\langle w_t - u, z_t \rangle]$$

We can conclude that

$$\mathbb{E}\left[\sum_{t=1}^{T} \langle w_t - u, z_t \rangle\right] \leq \sigma \log(d) + \mathbb{E}\left[\sum_{t=1}^{T} \sum_{i=1}^{d} z_t^{p_t}[i]^2 w_t[i]\right] /\sigma$$

So we get, as $z_t$ are the subgradients, that

$$\mathbb{E}\left[\sum_{t=1}^{T}(f_t(w_t) - f_t(u))\right] \leq \sigma \log(d) + \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{d} z_t^{p_t}[i]^2 w_t[i]\right] / \sigma$$

To bound the r.h.s we have

$$\mathbb{E}\left[\sum_{i=1}^{d} z_t^{p_t}[i]^2 w_t[i] | z_1, ..., z_{t-1}\right] = \sum_{j} P[p_t = j] \sum_{i=1}^{d} z_t^{j}[i]^2 w_t[i]$$

$$\sum_{j} w_t[j] \left(\frac{y_t[j]}{w_t[j]}\right)^2 w_t[j] \leq d.$$

The bandit algorithm has regret therefore expected regret bound

$$\mathbb{E}\left[\sum_{t=1}^{T} y_t[p_t]\right] - \min_{i\in[d]} \sum_{t=1}^{T} y_t[i] \leq \sigma \log(d) + Td/\sigma$$

In particular - setting $\sigma = \sqrt{dT/\log(d)}$ we get $2\sqrt{d\log(d)T}$ bound.