# The Follow Perturbed Leader Algorithm Protected from Unbounded One-Step Losses

V. V. V'yugin

Institute for Information Transmission Problems
Russian Academy of Sciences

ALT 2009, Porto 3-5 October 2009

## Experts algorithms for online prediction

**Specific loss functions**

Littlestone and Warmuth's Weighted Majority Algorithm (1989).
Vovk's Aggregating Algorithm (1990).
Weighted Average Forecasters with convex loss functions (Cesa-Bianchi and Lugosi, 2006)

**General losses**

Freund and Schapire's online allocation algorithm Hedge (1997).
Hannan's Follow the Perturbed Leader FPL algorithm (1957) ; rediscovered by Kalai and Vempala (2005).

## Prediction with expert advice

$s_t^i$ - loss of expert $i = 1, \dots N$ at step $t$, $s_t^i \geq 0$

$s_t$ - Learner loss at step $t$, $s_t \geq 0$

$s_{1:T}^i = \sum_{t=1}^{T} s_t^i$ - cumulative loss of expert $i$ on steps $\leq T$

$s_{1:T} = \sum_{t=1}^{T} s_t^i$ - cumulative loss of Learner on steps $\leq T$

Learner's goal

$$s_{1:T} \leq \min_i s_{1:T}^i + \text{regret}$$

## Follow leader algorithms

At each step $t$, observing Experts losses Learner makes a decision to follow one of the experts $i \in \{1, 2, \dots N\}$.

Learner suffers loss $s_t = s_t^i$ at step $t$.

In the traditional framework, we suppose that one-step losses are bounded, for example, $0 \le s_t^i \le 1$ for all $i$ and $t$.

## Following the Perturbed Leader FPL algorithm:

Output prediction of an expert $i$ which minimizes

$$s^i_{1:t-1} - \frac{1}{\varepsilon}\xi^i,$$

where $\xi^i$, $i = 1, \ldots N$, $t = 1, 2, \ldots$, is a sequence of i.i.d random variables distributed according to the exponential distribution with the density $p(x) = \exp\{-x\}$, and $\varepsilon$ is *a learning rate*.

The method of following the perturbed leader was discovered by Hannan (1957) and rediscovered by Kalai and Vempala (2005).

### Why randomization? An example.

In the deterministic framework, Learner can perform much worse than each expert:

let the current losses of two experts on steps $t = 0, 1, \dots$ be $s^1_{0,1,2,3,4,5,6,\dots} = \frac{1}{2}, 0, 1, 0, 1, 0, 1, \dots$ and $s^2_{0,1,2,3,4,5,6\dots} = 0, 1, 0, 1, 0, 1, 0, \dots$.

"Follow Leader" algorithm always chooses the wrong prediction.

One-step losses of experts are bounded $0 \le s_t^i \le 1$.

Kalai and Vempala, Hutter and Poland show that the expected cumulative loss of the FPL algorithm with variable learning rate $\varepsilon_t = O(1/\sqrt{t})$ has the upper bound

$$E(s_{1:T}) \le \min_{i=1,\dots,N} s^i_{1:T} + O(\sqrt{T \log N}),$$

where $N$ is the number of experts.

## Asymptotic performance of probabilistic algorithms

A probabilistic algorithm is called *asymptotically consistent in the mean* if

$$\limsup_{T\to\infty}\frac{1}{T}E(s_{1:T}-\min_{i=1,\dots N}s^i_{1:T})\le 0.$$

A probabilistic learning algorithm is called *Hannan consistent* if

$$\limsup_{T\to\infty}\frac{1}{T}\left(s_{1:T}-\min_{i=1,\dots N}s^i_{1:T}\right)\le 0$$

almost surely, where $s_{1:T}$ is its random cumulative loss.

## Unbounded one-step losses

In that follows we allow losses at any step to be unbounded

$$s_t^i \in (0, +\infty).$$

We use only general losses - the notion of a loss function is not used

## Unbounded losses: relevant results

Poland and Hutter (2005): an asymptotically consistent learning algorithm for $s_t^i \le t^{1/16}$.

Allenberg et al. (2006): algorithm for one-step losses $(s_t^i)^2 \le t^a$ with the expected regret

$$2\sqrt{N \ln N}(T+1)^{\frac{1}{2}(1+a+\beta)},$$

where $a > 0$, and $\beta > 0$ is a parameter of the algorithm.

*Volume* of a game at step $t$

$$v_t = \sum_{j=1}^{t} \max_i s_j^i.$$

*Scaled fluctuation* of a game at step $t$.

$$\mathrm{fluc}(t) = \frac{\Delta v_t}{v_t} = \frac{\max_i s_t^i}{v_t},$$

where $\Delta v_t = v_t - v_{t-1}$.

By definition $v_{t-1} \le v_t$ for all $t$ and $0 \le \mathrm{fluc}(t) \le 1$ for all $t$.

## Theorem

*For any probabilistic algorithm of choosing an expert and for any $\varepsilon$ such that $0 < \varepsilon < 1$, two experts exist such that*

$$v_t \to \infty \text{ as } t \to \infty,$$
$$\mathrm{fluc}(t) \geq 1 - \varepsilon,$$
$$\frac{1}{v_t} E(s_{1:t} - \min_{i=1,2} s^i_{1:t}) \geq \frac{1}{2}(1 - \varepsilon)$$

*for all $t$.*

A probabilistic algorithm can be asymptotically consistent only in games where

$$\mathrm{fluc}(t) = \frac{\Delta v_t}{v_t} \to 0 \text{ as } t \to \infty.$$

What is a suitable sufficient condition on a game?

We consider games, such that $\mathrm{fluc}(t) \le \gamma(t)$ for all $t$, where $\gamma(t)$ is a computable non-increasing real function such that $0 \le \gamma(t) \le 1$ for all $t$ and $\gamma(t) \to 0$ as $t \to \infty$.
We consider non-degenerate games, i.e., such that $v_t \to \infty$ as $t \to \infty$.

## Learning rate of the FPL algorithm

We define a learning rate

$$\varepsilon_t = \frac{1}{\mu_t v_{t-1}}, \text{ where}$$

$$\mu_t = \sqrt{\frac{e^2-1}{1+\ln N}}(\gamma(t))^{1/2}.$$

It holds $\mu_t \le \mu_{t-1}$ and $v_t \ge v_{t-1}$ for all $t$.

**FPL algorithm PROT.**
FOR $t = 1, \dots T$
Define

$$I_t = \mathrm{argmin}_{i=1,2,\dots N} \left\{ s^i_{1:t-1} - \frac{1}{\epsilon_t} \xi^i \right\},$$

where $\epsilon_t = 1/(\mu_t v_{t-1})$ is the learning rate.

Receive one-step losses $s^i_t$ for experts $i = 1, \dots, N$, and receive one-step loss $s_t = s^{I_t}_t$ of the master algorithm.
ENDFOR
Here, $\xi^1, \dots \xi^N$ are i.i.d random variables, distributed according to the density $p(x) = \exp\{-x\}$.

## FPL algorithm PROT: regret

### Theorem

*Let a game satisfies*

$$\mathrm{fluc}(t) \le \gamma(t) \text{ for all } t,$$

*where $\gamma(t)$ is a computable non-increasing positive real function such that $0 \le \gamma(t) \le 1$ for all $t$.*
*Then the expected cumulated loss of the* FPL *algorithm* PROT *with variable learning rate is bounded by*

$$E(s_{1:T}) \le \min_i s^i_{1:T} + 2\sqrt{(e^2-1)(1+\ln N)}\sum_{t=1}^{T}(\gamma(t))^{1/2}\Delta v_t.$$

**Theorem**

*Let a game satisfies*

$$\mathrm{fluc}(t) \le \gamma(t) \text{ for all } t,$$

*where $\gamma(t)$ is a computable non-increasing positive real function such that*

- $0 \le \gamma(t) \le 1$ *for all* $t$,
- $\gamma(t) \to 0$, *and also,*
- $v_t \to \infty$ *as* $t \to \infty$.

*Then the algorithm* PROT *is asymptotically consistent in the mean*

$$\limsup_{T\to\infty} \frac{1}{v_T} E(s_{1:T} - \min_{i=1,\dots N} s^i_{1:T}) \le 0.$$

## Theorem

*Let a game satisfies*

$$\mathrm{fluc}(t) \le \gamma(t) \text{ for all } t,$$

*where $\gamma(t)$ is a computable non-increasing positive real function such that*

$$\sum_{t=1}^{\infty} (\gamma(t))^2 < \infty.$$

*Let also, $v_t \to \infty$ as $t \to \infty$.*
*Then the algorithm* PROT *is Hannan consistent, i.e.,*

$$\limsup_{T\to\infty} \frac{1}{v_T} \left( s_{1:T} - \min_{i=1,\dots N} s^i_{1:T} \right) \le 0$$

*almost surely.*

## Details of the proof

We use an auxiliary **IFPL** algorithm. General scheme:

$$\mathrm{Loss}_{1:T}(\mathbf{FPL}) \leq \mathrm{Loss}_{1:T}(\mathbf{IFPL}) + \mathrm{regret}_1,$$
$$\mathrm{Loss}_{1:T}(\mathbf{IFPL}) \leq \min_i s^i_{1:T} + \mathrm{regret}_2,$$

$$\mathrm{Loss}_{1:T}(\mathbf{FPL}) \leq \min_i s^i_{1:T} + \mathrm{regret}_1 + \mathrm{regret}_2$$

**IFPL algorithm.**
FOR $t = 1, \dots T$
Define the learning rate

$$\epsilon'_t = \frac{1}{\mu_t v_t}, \text{ where}$$
$$\mu_t = \sqrt{\frac{e^2-1}{1+\ln N}}(\gamma(t))^{1/2}$$

Choose an expert with the minimal perturbed cumulated loss on steps $\le t$

$$J_t = \mathrm{argmin}_{i=1,2,\dots N}\left\{s^i_{1:t} - \frac{1}{\epsilon'_t}\xi^i\right\}.$$

Receive the one step loss $\tilde{s}_t = s_t^{J_t}$ of the $\mathrm{IFPL}$ algorithm.
ENDFOR

## Lemma

*The cumulated expected losses of the* FPL *and* IFPL *algorithms satisfy the inequality*

$$E(s_{1:T}) \le E(\tilde{s}_{1:T}) + (e^2 - 1)\sum_{t=1}^{T} (\gamma(t))^{1-\alpha_t} \Delta v_t$$

*for all* $T$, *where* $\alpha_t < 1$.

The optimal choice

$$\alpha_t = \frac{1}{2}\left(1 - \frac{\ln \frac{1+\ln N}{e^2-1}}{\ln \gamma(t)}\right)$$

## Lemma 2

### Lemma

*The expected cumulative loss of the* IFPL *algorithm is bounded*

$$E(\tilde{s}_{1:T}) \leq \min_i s^i_{1:T} + (1 + \ln N) \sum_{t=1}^{T} (\gamma(t))^{\alpha_t} \Delta v_t$$

*for all* $T$.

By definition $\mu_t = (\gamma(t))^{\alpha_t}$.

### Lemma

*Let $g(x)$ be a positive nondecreasing real function such that $x/g(x)$, $g(x)/x^2$ are non-increasing for $x > 0$ and $g(x) = g(-x)$ for all $x$.*
*Let the assumptions of the previous theorems hold and*

$$\sum_{t=1}^{\infty} \frac{g(\Delta v_t)}{g(v_t)} < \infty.$$

*Then the* FPL *algorithm* PROT *is Hannan consistent.*

Put $g(x) = x^2$.

Let $g(x)$ be a positive nondecreasing real function such that $x/g(x)$, $g(x)/x^2$ are non-increasing for $x>0$ and $g(x)=g(-x)$ for all $x$. Let also, $a_t$ be a nondecreasing sequence of real numbers such that $a_t\to\infty$ as $t\to\infty$ and $X_t$ be a sequence of independent random variables such that $E(X_t)=0$, for $t=1,2,\dots$. Using Petrov's book we obtain

$$\sum_{t=1}^{\infty}\frac{E(g(X_t))}{g(a_t)}<\infty$$

implies

$$\frac{1}{a_T}\sum_{t=1}^{T}X_t\to 0$$

as $T\to\infty$ almost surely.

We get $X_t=s_t-E(s_t)$

## Corollary

*Assume that $s_t^i \le t^a$ for all $t$ and $i = 1, \dots N$, and*

$$\liminf_{t \to \infty} \frac{v_t}{t^{a+\delta}} > 0,$$

*where $a$ and $\delta$ are positive real numbers. Let also in the algorithm PROT, $\gamma(t) = t^{-\delta}$ and $\mu_t$ is defined above. Then*

- *the algorithm PROT is asymptotically consistent in the mean for any $a > 0$ and $\delta > 0$;*
- *this algorithm is Hannan consistent for any $a > 0$ and $\delta > \frac{1}{2}$;*
- *the expected loss of this algorithm is bounded*

$$E(s_{1:T}) \le \min_i s_{1:T}^i + 2\sqrt{(e^2 - 1)(1 + \ln N)}\, T^{1 - \frac{1}{2}\delta + a}$$

*as $T \to \infty$.*

## Corrolary

The main result is also valid for the standard time scaling, i.e., when $v_T = T$ for all $T$, and when losses of experts are bounded, i.e., $a = 0$. Then the expected regret has the upper bound

$$2\sqrt{(e^2-1)(1+\ln N)}\sum_{t=1}^{T}(\gamma(t))^{1/2} \leq 4\sqrt{(e^2-1)(1+\ln N)T}$$

## Conclusions

The requirement $\mathrm{fluc}(t) \to 0$ imply that the rate of growth of the losses has to be slower than exponential.

Previous results relate to the case where losses grow at most polynomial.

Therefore, these results are more general in the regime "faster than polynomial, but slower than exponential".

The paper also cover the case where $v_T \ll T$. In this case consistency results also hold.