A Second-order Bound with Excess Losses

Pierre Gaillard [1 2] Gilles Stoltz [2] Tim van Erven [3]

[1]EDF R&D, Clamart, France

[2]GREGHEC: HEC Paris – CNRS, Jouy-en-Josas, France

[3]Université Paris-Sud, Orsay, France

June 14, 2014

Setting of prediction with expert advice

In each round t

- the learner makes a prediction by choosing a vector $\widehat{\boldsymbol{p}}_t = (\widehat{p}_{1,t}, \ldots, \widehat{p}_{K,t})$ of non-negative weights that sum to one
- nature chooses $\boldsymbol{\ell}_t = (\ell_{1,t}, \ldots, \ell_{K,t}) \in [a, b]^K$
- every expert k incurs loss $\ell_{k,t} \in [a, b]$
- the learner's loss is $\widehat{\ell}_t = \widehat{\boldsymbol{p}}_t \cdot \boldsymbol{\ell}_t \qquad \left(= \sum_{k=1}^{K} \widehat{p}_{k,t} \ell_{k,t} \right)$

The goal of the learner is to control its cumulative loss, which he can do by controling its cumulative regret against any expert k:

$$R_{k,T} \triangleq \sum_{t=1}^{T} \left(\widehat{\ell}_t - \ell_{k,t} \right)$$

Setting of prediction with expert advice

In each round t

- the learner makes a prediction by choosing a vector $\widehat{\boldsymbol{p}}_t = (\widehat{p}_{1,t}, \ldots, \widehat{p}_{K,t})$ of non-negative weights that sum to one
- nature chooses $\boldsymbol{\ell}_t = (\ell_{1,t}, \ldots, \ell_{k,t}) \in [a,b]^K$
- every expert k incurs loss $\ell_{k,t} \in [a,b]$
- the learner's loss is $\widehat{\ell}_t = \widehat{\boldsymbol{p}}_t \cdot \boldsymbol{\ell}_t \qquad \left(= \sum_{k=1}^K \widehat{p}_{k,t} \ell_{k,t} \right)$

In the worst case, the best bound that can be guaranteed is of order

$$R_{k,T} \triangleq \sum_{t=1}^{T} \left(\widehat{\ell}_t - \ell_{k,t} \right) \leqslant \square \sqrt{T \ln K}$$

Setting of prediction with expert advice

In each round t

- the learner makes a prediction by choosing a vector $\widehat{\boldsymbol{p}}_t = (\widehat{p}_{1,t}, \dots, \widehat{p}_{K,t})$ of non-negative weights that sum to one
- nature chooses $\boldsymbol{\ell}_t = (\ell_{1,t}, \dots, \ell_{k,t}) \in [a, b]^K$
- every expert k incurs loss $\ell_{k,t} \in [a, b]$
- the learner's loss is $\widehat{\ell}_t = \widehat{\boldsymbol{p}}_t \cdot \boldsymbol{\ell}_t \qquad \left(= \sum_{k=1}^K \widehat{p}_{k,t} \ell_{k,t} \right)$

But this can be improved, for example when losses take values in [0, 1] one can get

$$R_{k,T} \triangleq \sum_{t=1}^{T} \left(\widehat{\ell}_t - \ell_{k,t} \right) \leqslant \square \sqrt{(\ln K) \sum_{t=1}^{T} \ell_{k,t}}$$

Setting of prediction with expert advice

In each round t

- the learner makes a prediction by choosing a vector $\widehat{\boldsymbol{p}}_t = (\widehat{p}_{1,t}, \ldots, \widehat{p}_{K,t})$ of non-negative weights that sum to one
- nature chooses $\boldsymbol{\ell}_t = (\ell_{1,t}, \ldots, \ell_{K,t}) \in [a, b]^K$
- every expert k incurs loss $\ell_{k,t} \in [a, b]$
- the learner's loss is $\widehat{\ell}_t = \widehat{\boldsymbol{p}}_t \cdot \boldsymbol{\ell}_t \qquad \left(= \sum_{k=1}^K \widehat{p}_{k,t} \ell_{k,t}\right)$

Cesa-Bianchi et al (2007) raised the question of improving further by proving second-order bounds.

Two bounds established by Cesa-Bianchi et al (2007)

The first bound is of the form

$$R_{k,t} \leqslant \frac{\ln K}{\eta} + \eta \sum_{t=1}^{T} \ell_{k,t}^2$$

but no method is known to optimize it in η.

The second bound is of the form

$$R_{k,T} = O\left(\sqrt{(\ln K)\ \sum_{t=1}^{T} v_t}\right), \text{ where } v_t = \sum_{k \leqslant K} p_{k,t} \left(\widehat{\ell}_t - \ell_{k,t}\right)^2$$

However, it does not reflect that it is harder to compete with some experts that with other ones.

Two bounds established by Cesa-Bianchi et al (2007)

The first bound is of the form

$$R_{k,t} \leqslant \frac{\ln K}{\eta} + \eta \sum_{t=1}^{T} \ell_{k,t}^2 \Rightarrow R_{k,T} = O\left(\sqrt{(\ln K) \sum_{t=1}^{T} \ell_{k,t}^2}\right),$$

but no method is known to optimize it in η $\Rightarrow$ Impossible tuning.

The second bound is of the form

$$R_{k,T} = O\left(\sqrt{(\ln K) \sum_{t=1}^{T} v_t}\right) \text{, where } v_t = \sum_{k \leqslant K} p_{k,t} \left(\widehat{\ell}_t - \ell_{k,t}\right)^2$$

However, it does not reflect that it is harder to compete with some experts that with other ones.

A Second-order Bound with Excess Losses

We provide a third form of second-order bound

$$R_{k,T} = O\left(\sqrt{(\ln K) \sum_{t=1}^{T} \left(\widehat{\ell}_t - \ell_{k,t}\right)^2}\right), \qquad (1)$$

which holds for all k experts simulateneously.

Key element in the analysis: consider multiple learning rates and develop tuning techniques that go with it.

Interests of the bound: bounds of form (1) entail

- optimal scalling in the setting of experts reporting confidence [Blum and Mansour, 2007].
- improvement for small excess losses.
- constant regret in the special case of i.i.d. losses.
- regret bounds on the cumulative predictive risk of the associated strategy without any assumption on the underlying stochastic process [Wintenberger, 2014].

Prod with multiple learning rates (ML-Prod)

Parameters: a vector $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_K)$ of positive learning rates

Initialization: a vector $\boldsymbol{w}_0 = (w_{1,0}, \ldots, w_{K,0})$ of nonnegative weights that sum to 1

For each round $t = 1, 2, \ldots$

- form the mixture $\widehat{\boldsymbol{p}}_t$ defined component-wise by $p_{k,t} = \eta_k w_{k,t-1} \big/ \boldsymbol{\eta} \cdot \boldsymbol{w}_{t-1}$
- observe the loss vector $\boldsymbol{\ell}_t$ and incur loss $\widehat{\ell}_t = \boldsymbol{p}_t \cdot \boldsymbol{\ell}_t$
- for each expert k perform the update $w_{k,t} = w_{k,t-1}\big(1 + \eta_k\big(\widehat{\ell}_t - \ell_{k,t}\big)\big)$

If $\eta_k \leqslant 1/2$ and $\boldsymbol{\ell}_t \in [0,1]^K$, the cumulative loss is bounded as

$$\underbrace{\sum_{t=1}^{T} \widehat{\ell}_t}_{\text{Our loss}} \leqslant \min_{k=1,\ldots,K} \left\{ \underbrace{\sum_{t=1}^{T} \ell_{k,t}}_{\text{Loss of expert } k} + \underbrace{\frac{1}{\eta_k} \ln \frac{1}{w_{k,0}} + \eta_k \sum_{t=1}^{T} \big(\widehat{\ell}_t - \ell_{k,t}\big)^2}_{\text{Regret}} \right\}$$

Prod with multiple learning rates (ML-Prod)

Parameters: a vector $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_K)$ of positive learning rates

Initialization: a vector $\boldsymbol{w}_0 = (w_{1,0}, \ldots, w_{K,0})$ of nonnegative weights that sum to 1

For each round $t = 1, 2, \ldots$

- form the mixture $\widehat{\boldsymbol{p}}_t$ defined component-wise by $p_{k,t} = \eta_k w_{k,t-1} \big/ \boldsymbol{\eta} \cdot \boldsymbol{w}_{t-1}$
- observe the loss vector $\boldsymbol{\ell}_t$ and incur loss $\widehat{\ell}_t = \boldsymbol{p}_t \cdot \boldsymbol{\ell}_t$
- for each expert k perform the update $w_{k,t} = w_{k,t-1}\big(1 + \eta_k\big(\widehat{\ell}_t - \ell_{k,t}\big)\big)$

If we could optimize $\eta_k = \sqrt{\ln(1/w_{k,0}) / \sum_t (\widehat{\ell}_t - \ell_{k,t})^2}$

$$\underbrace{\sum_{t=1}^{T} \widehat{\ell}_t}_{\text{Our loss}} \leqslant \min_{k=1,\ldots,K} \left\{ \underbrace{\sum_{t=1}^{T} \ell_{k,t}}_{\text{Loss of expert } k} + \underbrace{2\sqrt{\ln \frac{1}{w_{k,0}} \sum_{t=1}^{T} \big(\widehat{\ell}_t - \ell_{k,t}\big)^2}}_{\text{Regret}} \right\}$$

Adaptive version of ML-Prod

Parameters: a rule to sequentially pick positive learning rates

Initialization: a vector $\boldsymbol{w}_0 = (w_{1,0}, \ldots, w_{K,0})$ of nonnegative weights that sum to 1

For each round $t = 1, 2, \ldots$

- form the mixture $\widehat{\boldsymbol{p}}_t$ defined component-wise by $p_{k,t} \propto \eta_{k,t-1} w_{k,t-1}$
- observe the loss vector $\boldsymbol{\ell}_t$ and incur loss $\widehat{\ell}_t = \boldsymbol{p}_t \cdot \boldsymbol{\ell}_t$
- for each expert k perform the update

$$w_{k,t} = \left(w_{k,t-1} \left(1 + \eta_{k,t-1} \left(\widehat{\ell}_t - \ell_{k,t} \right) \right) \right)^{\frac{\eta_{k,t}}{\eta_{k,t-1}}}$$

If $\eta_{k,t} \leqslant 1/2$, $(\eta_{k,t})$ is non-increasing in t and $\boldsymbol{\ell}_t \in [0,1]^K$, the cumulative regret $R_{k,T} = \sum_{t \leqslant T} (\widehat{\ell}_t - \ell_{k,t})$ is bounded simultaneously for all expert k as

$$R_{k,T} \leqslant \frac{1}{\eta_{k,0}} \ln \frac{1}{w_{k,0}} + \sum_{t=1}^{T} \eta_{k,t-1} \left(\widehat{\ell}_t - \ell_{k,t} \right)^2 + \underbrace{\frac{1}{\eta_{k,T}} \ln \left(1 + \frac{1}{\varepsilon} \sum_{k'=1}^{K} \sum_{t=1}^{T} \left(\frac{\eta_{k',t-1}}{\eta_{k',t}} - 1 \right) \right)}_{\text{Cost of tuning multiple learning rates}}$$

Adaptive version of ML-Prod

Parameters: a rule to sequentially pick positive learning rates

Initialization: a vector $\boldsymbol{w}_0 = (w_{1,0}, \dots, w_{K,0})$ of nonnegative weights that sum to 1

For each round $t = 1, 2, \dots$

- form the mixture $\widehat{\boldsymbol{p}}_t$ defined component-wise by $p_{k,t} \propto \eta_{k,t-1} w_{k,t-1}$
- observe the loss vector $\boldsymbol{\ell}_t$ and incur loss $\widehat{\ell}_t = \boldsymbol{p}_t \cdot \boldsymbol{\ell}_t$
- for each expert k perform the update

$$w_{k,t} = \left(w_{k,t-1} \left(1 + \eta_{k,t-1} \left(\widehat{\ell}_t - \ell_{k,t} \right) \right) \right)^{\frac{\eta_{k,t}}{\eta_{k,t-1}}}$$

With uniform initial weights $\boldsymbol{w}_0 = (1/K, \dots, 1/K)$ and learning rates, for $t \geqslant 1$,

$$\eta_{k,t-1} = \min \left\{ \frac{1}{2}, \sqrt{\frac{\ln K}{1 + \sum_{s=1}^{t-1} \left(\widehat{\ell}_s - \ell_{k,s} \right)^2}} \right\},$$

the cumulative regret is bounded simultaneously for all expert k as

$$R_{k,T} = O\left(\ln K + \sqrt{(\ln K) \sum_{t=1}^{T} \left(\widehat{\ell}_t - \ell_{k,t} \right)^2 \ln \ln T} \right),$$

Small excess losses

If a strategy satisfies a bound of the form

$$\sum_{t=1}^{T} \widehat{\ell}_t \leqslant \min_{k=1,\dots,K} \left\{ \sum_{t=1}^{T} \ell_{k,t} + C_1 \sqrt{\ln K \sum_{t=1}^{T} \left(\widehat{\ell}_t - \ell_{k,t}\right)^2} + C_2 \right\}$$

Then, if $\boldsymbol{\ell}_t \in [0,1]^K$, is also satisfies

$$\sum_{t=1}^{T} \widehat{\ell}_t \leqslant \min_{k=1,\dots,K} \left\{ \sum_{t=1}^{T} \ell_{k,t} + 2C_1 \sqrt{\ln K \sum_{t:\ell_{k,t} \geqslant \widehat{\ell}_t} \left(\ell_{k,t} - \widehat{\ell}_t\right)} + O\Big(C_2 + C_1^2 \ln K\Big) \right\}$$

This bound is invariant by translation of the losses and implies the small losses bound $R_{k,T} \leqslant O\Big(\sqrt{(\ln K)\sqrt{\sum_t \ell_{k,t}}}\Big)$.

Expert that report their confidence [Blum and Mansour, 2007]

In each round $t = 1, \dots, T$

- each expert k expresses its confidence as a number $I_{k,t} \in [0, 1]$
- the learner makes a prediction by choosing a vector $\widehat{\boldsymbol{p}}_t = (\widehat{p}_{1,t}, \dots, \widehat{p}_{K,t})$ of non-negative weights that sum to one
- nature chooses $\boldsymbol{\ell}_t = (\ell_{1,t}, \dots, \ell_{K,t}) \in [a, b]^K$
- every expert k incurs loss $\ell_{k,t} \in [a, b]$
- the learner's loss is $\widehat{\ell}_t = \widehat{\boldsymbol{p}}_t \cdot \boldsymbol{\ell}_t \qquad \left(= \sum_{k=1}^K \widehat{p}_{k,t} \ell_{k,t} \right)$

The learner aims at minimizing its confidence regret simultaneously for all experts

$$R^c_{k,T} = \sum_{t=1}^{T} I_{k,t} \left(\widehat{\ell}_t - \ell_{k,t} \right)$$

The special case $I_{k,t} = 0$ expresses that expert k is inactive in round t.

Expert that report their confidence [Blum and Mansour, 2007]

In each round $t = 1, \dots, T$

- each expert k expresses its confidence as a number $I_{k,t} \in [0, 1]$
- the learner makes a prediction by choosing a vector $\widehat{\boldsymbol{p}}_t = (\widehat{p}_{1,t}, \dots, \widehat{p}_{K,t})$ of non-negative weights that sum to one
- nature chooses $\boldsymbol{\ell}_t = (\ell_{1,t}, \dots, \ell_{K,t}) \in [a, b]^K$
- every expert k incurs loss $\ell_{k,t} \in [a, b]$
- the learner's loss is $\widehat{\ell}_t = \widehat{\boldsymbol{p}}_t \cdot \boldsymbol{\ell}_t \qquad \left(= \sum_{k=1}^K \widehat{p}_{k,t} \ell_{k,t} \right)$

The learner aims at minimizing its confidence regret simultaneously for all experts

$$R^c_{k,T} = \sum_{t=1}^{T} I_{k,t} \left(\widehat{\ell}_t - \ell_{k,t} \right)$$

Because the confidence regret scales linearly with $I_{k,t}$, one would therefore like to obtain bounds on the confidence regret that scale linearly as well.

Expert that report their confidence [Blum and Mansour, 2007]

In each round $t = 1, \dots, T$

- each expert k expresses its confidence as a number $I_{k,t} \in [0,1]$
- the learner makes a prediction by choosing a vector $\widehat{\boldsymbol{p}}_t = (\widehat{p}_{1,t}, \dots, \widehat{p}_{K,t})$ of non-negative weights that sum to one
- nature chooses $\boldsymbol{\ell}_t = (\ell_{1,t}, \dots, \ell_{K,t}) \in [a,b]^K$
- every expert k incurs loss $\ell_{k,t} \in [a,b]$
- the learner's loss is $\widehat{\ell}_t = \widehat{\boldsymbol{p}}_t \cdot \boldsymbol{\ell}_t \qquad \left(= \sum_{k=1}^K \widehat{p}_{k,t} \ell_{k,t} \right)$

The best available stated bound [Blum and Mansour, 2007] is

$$R^c_{k,T} = \sum_{t=1}^T I_{k,t} \big(\widehat{\ell}_t - \ell_{k,t} \big) = O\left(\sqrt{(\ln K) \sum_{t=1}^T I_{k,t} \ell_{k,t}} \right).$$

Application to experts that report their confidence

If a strategy satisfies a standard regret bound of the form

$$\sum_{t=1}^{T} \widehat{\ell}_t \leqslant \min_{k=1,\dots,K} \left\{ \sum_{t=1}^{T} \ell_{k,t} + C_1 \sqrt{(\ln K) \sum_{t=1}^{T} \left(\widehat{\ell}_t - \ell_{k,t}\right)^2} + C_2 \right\}$$

Then, if $\ell_t \in [0,1]^K$, applying the strategy on the modified losses

$$\tilde{\ell}_{k,t} = I_{k,t}\ell_{k,t} + (1 - I_{k,t})\ell_{k,t} \,,$$

leads to an algorithm with a confidence regret bound of the form

$$R^c_{k,T} \leqslant C_1 \sqrt{(\ln K) \sum_{t=1}^{T} I^2_{k,t} \left(\widehat{\ell}_t - \ell_{k,t}\right)^2} + C_2 \leqslant C_1 \sqrt{(\ln K) \sum_{t=1}^{T} I^2_{k,t}} + C_2 \quad \text{for all } k.$$

Stochastic (i.i.d.) losses

We now turn to a stochastic setting considered by [Van Erven et al, 2011] where the loss vectors are identically distributed.

Assumption [Van Erven et al, 2011]

The loss vectors $\boldsymbol{\ell}_t \in [0,1]^K$ are independent random variables such that there exists an action $k^\star$ and some $\alpha \in (0,1]$ such that

$$\forall t \geqslant 1, \qquad \min_{k \neq k^\star} \mathbb{E}\big[\ell_{k,t} - \ell_{k^\star,t}\big] \geqslant \alpha\,.$$

If some strategy satisfies

$$\sum_{t=1}^T \widehat{\ell}_t \leqslant \sum_{t=1}^T \ell_{k^\star,t} + 2\sqrt{\ln K \sum_{t=1}^T \left(\widehat{\ell}_t - \ell_{k^\star,t}\right)^2}$$

Then

- $$\mathbb{E}\left[\sum_{t=1}^T \widehat{\ell}_t\right] \leqslant \mathbb{E}\left[\sum_{t=1}^T \ell_{k^\star,t}\right] + \frac{4\ln K}{\alpha}$$
- For any $\delta \in (0,1)$, with probability at least $1-\delta$
$$\sum_{t=1}^T \widehat{\ell}_t \leqslant \sum_{t=1}^T \ell_{k^\star,t} + \frac{4\ln K}{\alpha} + \frac{12}{\alpha}\sqrt{\ln \frac{\square}{\delta} \ln K}$$

Application to cumulative risk [Wintenberger, 2014]

We turn to a stochastic setting considered by [Wintenberger, 2014] where (X_t, Y_t) are random element observed recursively.

In each round $t = 1, \dots, T$

- the learner makes a prediction by choosing $\widehat{f}_t : \mathcal{X} \to \mathbb{R}$ with knowledge of the past $\mathcal{F}_{t-1} = \big\{(X_1, Y_1), \dots, (X_{t-1}, Y_{t-1})\big\}$
- nature reveals $(X_t, Y_t) \in \mathcal{X} \times \mathbb{R}$
- the learner's loss is $\mathbb{E}\Big[\ell\big(Y_t, \widehat{f}_t(X_t)\big)\big|\mathcal{F}_{t-1}\Big]$ (not observed)

The accuracy of the learner $\widehat{f} = (\widehat{f}_1, \dots, \widehat{f}_T)$ is quantified by its cumulative risk

$$\mathcal{R}_T\big(\widehat{f}\big) = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\Big[\ell\big(Y_t, \widehat{f}_t(X_t)\big)\big|\mathcal{F}_{t-1}\Big]$$

Application to cumulative risk [Wintenberger, 2014]

We turn to a stochastic setting considered by [Wintenberger, 2014] where (X_t, Y_t) are random element observed recursively.

In each round $t = 1, \dots, T$

- each expert $k = 1, \dots, K$ suggests a prediction $f_{k,t} : \mathcal{X} \to \mathbb{R}$
- the learner assigns weight to each expert and predict

$$\widehat{f}_t = \sum_{k=1}^{K} \widehat{p}_{k,t} f_{k,t}$$

- nature reveals $(X_t, Y_t) \in \mathcal{X} \times \mathbb{R}$

The accuracy of the learner $\widehat{f} = (\widehat{f}_1, \dots, \widehat{f}_T)$ is quantified by its cumulative risk

$$\mathcal{R}_T(\widehat{f}) = \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[\ell\left(Y_t, \widehat{f}_t(X_t)\right) \middle| \mathcal{F}_{t-1}\right]$$

Application to cumulative risk [Wintenberger, 2014]

We turn to a stochastic setting considered by [Wintenberger, 2014] where (X_t, Y_t) are random element observed recursively.

In each round $t = 1, \dots, T$

- each expert $k = 1, \dots, K$ suggests a prediction $f_{k,t} : \mathcal{X} \to \mathbb{R}$
- the learner assigns weight to each expert and predict

$$\widehat{f}_t = \sum_{k=1}^{K} \widehat{p}_{k,t} f_{k,t}$$

- nature reveals $(X_t, Y_t) \in \mathcal{X} \times \mathbb{R}$

Wintenberger (2014) proved that deterministic second-order bounds in excess losses applied with $\ell_{k,t} = \ell(Y_t, f_{k,t}(X_t))$ can be converted to bound in cumulative risk with probability at least $1 - \delta$,

$$\mathcal{R}_T(\widehat{f}) \leqslant \min_k \left\{ \mathcal{R}_T(\widehat{f}_k) + \square \sqrt{(\ln K) \sum_{t=1}^{T} (\ell_{k,t} - \widehat{\ell}_t)^2} \left(1 + \log \frac{1}{\delta} (\ln K)^{-1} \right) \right\} + \square \ln \frac{K}{\delta}$$