The Pareto Regret Frontier

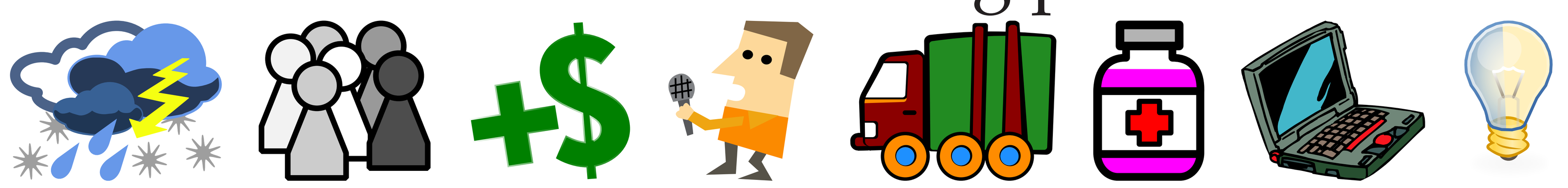
Wouter M. Koolen

Queensland University of Technology · CWI Centrum Wiskunde & Informatica

Poster Sun12
December 8th 2013

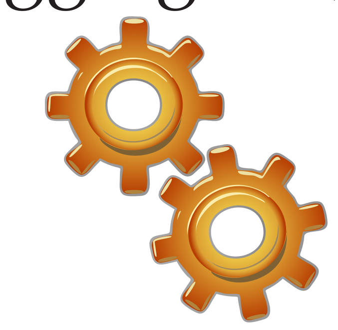
The online learning philosophy

Want to solve a learning problem?



Step 1: Get hold of some experts

Step 2: Aggregate (in production)

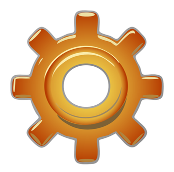


Step 3:

Problem solved?

HEDGE aggregates K experts s.t. after T rounds

$$\underbrace{L_T - L_T^k}_{\text{regret}} \leq \sqrt{\tfrac{1}{2} T \ln K} \quad \text{for all } k$$



With tight lower bound

No! Plea for favouritism

What if

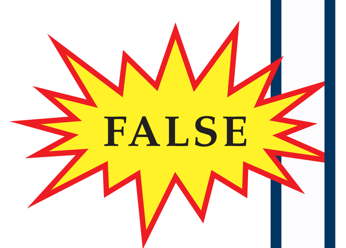
- Lots of experts?
- Special experts?

We need biased aggregation

Simply tune a little differently?

Natural guess: for every distribution $\mathbb{P}$ on experts, we can ensure

$$L_T - L_T^k \leq \sqrt{\tfrac{1}{2} T \left(-\ln \mathbb{P}(k)\right)} \quad \text{for all } k$$



So what **can** we guarantee? And **how**?

The setup

On each round t the learner plays a probability distribution $\boldsymbol{p}_t$ on K experts. Then the vector of expert losses $\boldsymbol{\ell}_t \in [0,1]^K$ is revealed, and the learner suffers

$$\textbf{dot loss} \qquad \boldsymbol{p}_t^\intercal \boldsymbol{\ell}_t.$$

After T rounds, the **regret** w.r.t. expert k is

$$\text{Regret}_T^k := \sum_{t=1}^T \left(\boldsymbol{p}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k\right)$$

A candidate trade-off $\boldsymbol{r} \in \mathbb{R}^K$ is called T**-realisable** if there is a strategy that keeps the regret w.r.t. each k below r^k, i.e. if

$$\exists \boldsymbol{p}_1 \forall \boldsymbol{\ell}_1 \cdots \exists \boldsymbol{p}_T \forall \boldsymbol{\ell}_T \forall k : \text{Regret}_T^k \leq r^k.$$

Combinatorial characterisation ($K = 2$ experts)

The T-realisable Pareto front is piece-wise linear with $T+1$ vertices:

$$\langle f_T(i), f_T(T-i)\rangle \quad 0 \leq i \leq T \quad \text{where} \quad f_T(i) := \sum_{j=0}^{i} j 2^{j-T} \binom{T-j-1}{T-i-1}$$

The optimal strategy at vertex i assigns to expert 1 probability

$$p_T(0) = 0, \quad p_T(T) = 1, \quad p_T(i) = \frac{f_{T-1}(i) - f_{T-1}(i-1)}{2} \quad 0 < i < T$$

and it interpolates linearly between neighbouring vertices.

Normalised large horizon behaviour

$$\textbf{limit frontier} := \lim_{T\to\infty} \frac{\text{frontier}_T}{\sqrt{T}}$$

Asymptotic characterisation ($K = 2$ experts)

The limit frontier is the smooth curve

$$\langle f(-u), f(+u)\rangle \quad u \in \mathbb{R}, \quad \text{where} \quad f(u) := u\Phi(u) + \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}}$$

and $\Phi(u) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{x^2}{2}} \, \mathrm{d}x$. The optimal strategy converges to

$$p(u) = \Phi(u)$$

Conclusion

Unfairness is a key resource!

Regret $\sqrt{T(-\ln \mathbb{P}(k))}$ is realisable for any prior $\mathbb{P}$ on $K = 2$ experts …but **fundamentally suboptimal**

$\sqrt{2.6T}(-\ln \mathbb{P}(k))$ is realisable for any prior $\mathbb{P}$ on $K > 2$ experts (by biased one-vs-all chain)

Thank you

Interesting follow-ups:

- $K > 2$ experts exact frontier
- Other losses
- Luckiness stratifications, e.g. $\sqrt{L_T^k}$, $\sqrt{\frac{L_T^k(T-L_T^k)}{T}}$, …
- Horizon-free biased rates ρ: $\text{Regret}_t^k \leq \rho^k \sqrt{t}$ for all t
- Use of limit algorithm

Pareto frontier for $K = 2$ experts. Finite horizons (left). Asymptotic (middle & right)