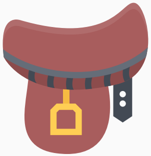
Supermartingales in Online Learning

Wouter M. Koolen
CWI and University of Twente

SAVI 2 Workshop Oberwolfach, May 8, 2024

Before we begin: Postdoc Position



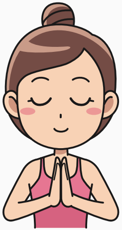
On efficient computation of saddle points arising in GRO/KLInf/GLRT/...

Co-hosted by CWI & INRIA team 4TUNE

- Pierre Gaillard (Grenoble)
- Rémy Degenne (Lille)
- Adrien Taylor (Paris)

Vacancy

Let's begin: Warm Thanks



Tim van Erven

Zakaria Mhammedi

Muriel Pérez-Ortiz

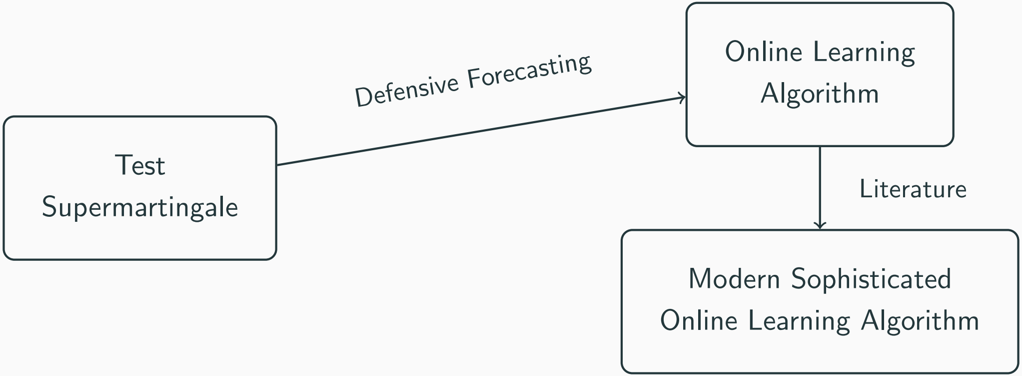
Online Learning: Is it E-relevant?

Online Learning: Is it E-relevant?

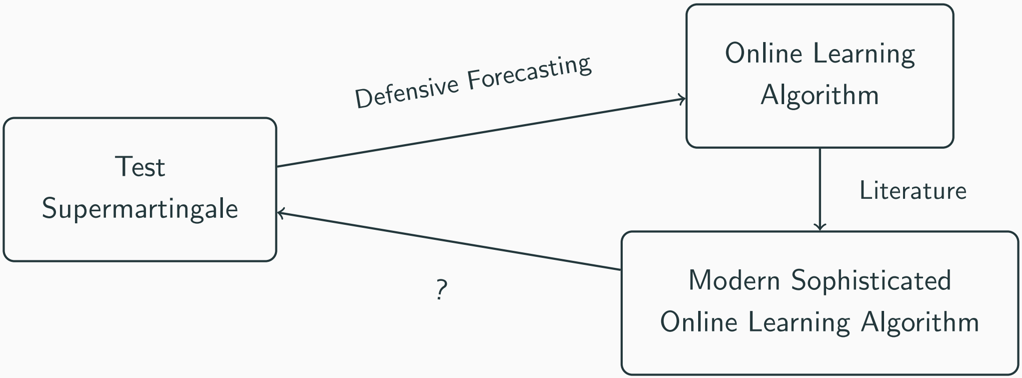
Test
Supermartingale

Online Learning: Is it E-relevant?

Online Learning: Is it E-relevant?



Online Learning: Is it E-relevant?



Starting Point: Hedge

Online Learning Setup

Protocol (Hedge setting with K experts)

For $t = 1, 2, \ldots$

- Learner plays $\boldsymbol{w}_t \in \triangle_K$
- Adversary picks $\boldsymbol{\ell}_t \in [0, 1]^K$

Online Learning Setup

Protocol (Hedge setting with K experts)

For $t = 1, 2, \ldots$

- Learner plays $\boldsymbol{w}_t \in \triangle_K$
- Adversary picks $\boldsymbol{\ell}_t \in [0, 1]^K$

Definition (Regret)

The regret after T rounds w.r.t expert k is defined as

$$R_T^k := \sum_{t=1}^{T} \left(\boldsymbol{w}_t^\top \boldsymbol{\ell}_t - \ell_t^k \right)$$

Online Learning Setup

Protocol (Hedge setting with K experts)

For $t = 1, 2, \ldots$

- Learner plays $\boldsymbol{w}_t \in \triangle_K$
- Adversary picks $\boldsymbol{\ell}_t \in [0, 1]^K$

Definition (Regret)

The regret after T rounds w.r.t expert k is defined as

$$R_T^k := \sum_{t=1}^{T} \left(\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k \right)$$

Goal: strategy for Learner keeping all R_T^k small.

Skeptic Perspective

Someone claims that their strategy for setting $\boldsymbol{w}_t$ ensures sublinear regret $R_T^k = o(T)$.

Skeptic Perspective

Someone claims that their strategy for setting $\boldsymbol{w}_t$ ensures sublinear regret $R_T^k = o(T)$.

We're at a workshop about sequential testing by betting. Let's!

Skeptic Perspective

Someone claims that their strategy for setting $\boldsymbol{w}_t$ ensures sublinear regret $R_T^k = o(T)$.

We're at a workshop about sequential testing by betting. Let's!

Forecasts $\boldsymbol{w}_t$ are claimed to agree with Outcomes $\boldsymbol{\ell}_t$ in the sense that

$$\forall k : \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \left(\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k \right) \leq 0$$

Skeptic Perspective

Someone claims that their strategy for setting $\boldsymbol{w}_t$ ensures sublinear regret $R_T^k = o(T)$.

We're at a workshop about sequential testing by betting. Let's!

Forecasts $\boldsymbol{w}_t$ are claimed to agree with Outcomes $\boldsymbol{\ell}_t$ in the sense that

$$\forall k : \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \left(\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k \right) \leq 0$$

Let's read that as the claim that $\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \ldots$ come from any joint $\mathbb{P}$ satisfying

$$\forall k, t : \mathop{\mathbb{E}}_{\boldsymbol{\ell}_t} \left[\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k \middle| \mathcal{F}_{t-1} \right] \leq 0$$

Betting

If we assume the null is

$$\mathcal{H}_0 = \left\{ \mathbb{P} \text{ on } \boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \ldots \middle| \forall k, t : \mathop{\mathbb{E}}_{\boldsymbol{\ell}_t}\left[\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k \middle| \mathcal{F}_{t-1} \right] \leq 0 \right\}$$

Then what are the available e-values to bet on?

Betting

If we assume the null is

$$\mathcal{H}_0 \;=\; \left\{ \mathbb{P} \text{ on } \boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \ldots \middle| \forall k, t : \mathop{\mathbb{E}}_{\boldsymbol{\ell}_t}\left[\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k \middle| \mathcal{F}_{t-1} \right] \;\le\; 0 \right\}$$

Then what are the available e-values to bet on?

Theorem (Larsson, ISI WSC 2021 Virtual Talk)

Given $\mathcal{F}_{t-1}$, all admissible e-values for $\mathcal{H}_0$ are of the form

$$E_{\boldsymbol{\eta}_t}(\boldsymbol{\ell}_t) \;:=\; 1 + \sum_{k=1}^{K} \eta_t^k \left(\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k \right)$$

for $\eta_t^k \ge 0$ (since they correspond to inequality constraints) and $\boldsymbol{\eta}_t$ ensuring non-negativity $\min_{\boldsymbol{\ell}_t \in [0,1]^K} E_{\boldsymbol{\eta}_t}(\boldsymbol{\ell}_t) \ge 0$, i.e. $\sum_{k=1}^K \left(\eta_t^k - w_t^k \mathbf{1}^\intercal \boldsymbol{\eta}_t \right)_+ \le 1$.

Actual betting strategy

Simple and effective: mix over experts k, repeating fixed bet $\eta > 0$ and

$$1 + \eta\left(\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k\right) \;\geq\; e^{\eta\left(\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k\right) - \eta^2/2}$$

Definition (Hedge supermartingale; Chernov and Vovk 2009)

$$\Phi_T \;:=\; \sum_{k=1}^K \frac{1}{K} \prod_{t=1}^T e^{\eta\left(\boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k\right) - \eta^2/2} \;=\; \sum_{k=1}^K \frac{1}{K} e^{\eta R_T^k - T\eta^2/2}$$

Say we reject $\mathcal{H}_0$ when $\Phi_T \geq 1/\alpha$. This occurs for $\eta = \sqrt{\frac{2 \ln \frac{K}{\alpha}}{T}}$ if

$$\exists k : R_T^k \;\geq\; \frac{1}{\eta} \ln \frac{K}{\alpha} + T \frac{\eta}{2} \;=\; \sqrt{2T \ln \frac{K}{\alpha}}$$

Defensive Forecasting

We have a test supermartingale Φ_T against $\mathcal{H}_0$: the hypothesis that the Learner guarantees sub-linear regret.

Defensive Forecasting

We have a test supermartingale Φ_T against $\mathcal{H}_0$: the hypothesis that the Learner guarantees sub-linear regret.

So?

Defensive Forecasting

We have a test supermartingale Φ_T against $\mathcal{H}_0$: the hypothesis that the Learner guarantees sub-linear regret.

So?

We can turn this test into a strategy for Learner.

Idea: Always pick $\boldsymbol{w}_t$ to ensure Φ_t does not get big.

Defensive Forecasting

Let's see. We can guarantee

$$\Phi_{T+1} \;\leq\; \sum_{k=1}^{K} \frac{1}{K} e^{\eta R_T^k - T\eta^2/2} \left(1 + \eta \left(\boldsymbol{w}_{T+1}^\intercal \boldsymbol{\ell}_{T+1} - \ell_{T+1}^k\right)\right) \;\overset{(\star)}{=}\; \Phi_T$$

Affine in $\boldsymbol{\ell}_{T+1}$. So set $\boldsymbol{w}_{T+1}$ to cancel coefficient:

$$\mathbf{0} \;=\; \sum_{k=1}^{K} \frac{1}{K} e^{\eta R_T^k - T\eta^2/2} \eta \left(\boldsymbol{w}_{T+1} - \boldsymbol{e}_k\right) \qquad \Rightarrow \qquad w_{T+1}^k \;=\; \frac{\frac{1}{K} e^{\eta R_T^k - T\eta^2/2} \eta}{\sum_{j=1}^{K} \frac{1}{K} e^{\eta R_T^j - T\eta^2/2} \eta}$$

Online Learning Consequence

Theorem (Freund and Schapire, 1997)

The algorithm

$$w_{T+1}^{k} = \frac{e^{\eta R_T^k}}{\sum_{j=1}^{K} e^{\eta R_T^j}}$$

with $\eta = \sqrt{\frac{2 \ln K}{T}}$ *guarantees regret bounded by*

$$\forall k : R_T^k \leq \sqrt{2 T \ln K}$$

Postmortem

- We started with an online learning protocol.
- We desired small regret compared to any expert k
- We put forward a test supermartingale against that goal
- And then we synthesised an algorithm, from that test, achieving the goal.

Postmortem

- We started with an online learning protocol.
- We desired small regret compared to any expert k
- We put forward a test supermartingale against that goal
- And then we synthesised an algorithm, from that test, achieving the goal.

E-Lessons

- (Desired) regret guarantees are generating the null
- Several goals $\Rightarrow$ mixture martingale

Squint

Squint

Squint tightens all the screws on the Hedge supermartingale.

What?

- No tuning parameter (η)
- Anytime
- Stochastic Luckiness
- Comparator adaptivity; quantile bounds; countably many experts $K = \infty$
- Same computational cost

How?

- Refined bets
- Prior on experts
- Prior on η (improper!)

Squint Supermartingale

Let us define the instantaneous regret in round t w.r.t. expert k by

$$r_t^k \;:=\; \boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k$$

We know that critical r_t^k are small. So let's use ever-so-slightly-tighter e-value

$$1 + \eta r_t^k \;\geq\; e^{\eta r_t^k - \eta^2 (r_t^k)^2}$$

Definition (Squint supermartingale; Koolen and van Erven 2015)

Fix prior $\boldsymbol{\pi} \in \triangle_K$. Define

$$\Phi_T \;:=\; \sum_{k=1}^K \pi_k \int_0^{\frac{1}{2}} \frac{e^{\eta R_T^k - \eta^2 V_T^k} - 1}{\eta} \, \mathrm{d}\eta \qquad \text{where} \qquad V_T^k \;=\; \sum_{t=1}^T (r_t^k)^2$$

Ehh, did we need non-negativity?

$$\Phi_T \;:=\; \sum_{k=1}^{K} \pi_k \int_0^{\frac{1}{2}} \frac{e^{\eta R_T^k - \eta^2 V_T^k} - 1}{\eta} \,\mathrm{d}\eta$$

Mixture of centred supermartingales $e^{\sum_t \cdots} - 1$ under improper prior $\frac{1}{\eta}\,\mathrm{d}\eta$ possibly negative

Ehh, did we need non-negativity?

$$\Phi_T \;:=\; \sum_{k=1}^K \pi_k \int_0^{\frac{1}{2}} \frac{e^{\eta R_T^k - \eta^2 V_T^k} - 1}{\eta} \, \mathrm{d}\eta$$

Mixture of centred supermartingales $e^{\sum_t \cdots} - 1$ under improper prior $\frac{1}{\eta} \, \mathrm{d}\eta$ possibly negative

Still

$$\Phi_T \geq - \ln T.$$

Defensive Forecasting for Squint

We have

$$\Phi_{T+1} \;\leq\; \sum_{k=1}^{K} \pi_k \int_0^{\frac{1}{2}} \frac{e^{\eta R_T^k - \eta^2 V_T^k} \left(1 + \eta \left(\boldsymbol{w}_{T+1}^\intercal \boldsymbol{\ell}_{T+1} - \ell_{T+1}^k\right)\right) - 1}{\eta} \,\mathrm{d}\eta \;\overset{(\star)}{=}\; \Phi_T$$

for the unique equaliser choice

$$\mathbf{0} = \sum_{k=1}^{K} \pi_k \int_0^{\frac{1}{2}} e^{\eta R_T^k - \eta^2 V_T^k} \left(\boldsymbol{w}_{T+1} - \boldsymbol{e}_k\right) \mathrm{d}\eta \qquad \text{i.e.} \qquad w_{T+1}^k = \frac{\pi_k \int_0^{\frac{1}{2}} e^{\eta R_T^k - \eta^2 V_T^k} \,\mathrm{d}\eta}{\sum_{j=1}^{K} \pi_j \int_0^{\frac{1}{2}} e^{\eta R_T^j - \eta^2 V_T^j} \,\mathrm{d}\eta}$$

Cool feature: $\boldsymbol{w}_T$ has a closed form expression (Gaussian CDFs) though Φ_T does not.

Small is Beautiful for Squint

Theorem

$$\forall T : \Phi_T \leq \Phi_0 = 0 \qquad \textit{implies} \qquad \forall k, T : R_T^k \ \leq \ 2\sqrt{V_T^k \ln \frac{\ln T}{\pi_k}}$$

Small is Beautiful for Squint

Theorem

$$\forall T : \Phi_T \le \Phi_0 = 0 \qquad \textit{implies} \qquad \forall k, T : R_T^k \;\le\; 2\sqrt{V_T^k \ln \frac{\ln T}{\pi_k}}$$

Why? Thinking about $\eta = \sqrt{\frac{\ln \frac{\ln T}{\pi_k}}{V_T^k}}$ gives

$$\Phi_T \;=\; \sum_{k=1}^K \pi_k \int_0^{\frac{1}{2}} \frac{e^{\eta R_T^k - \eta^2 V_T^k} - 1}{\eta} \mathrm{d}\eta \;\approx\; \sum_{k=1}^K \pi_k e^{\sqrt{\frac{\ln \frac{\ln T}{\pi_k}}{V_T^k}} R_T^k - \ln \frac{\ln T}{\pi_k}} - \ln T$$

Small is Beautiful for Squint

Theorem

$$\forall T : \Phi_T \le \Phi_0 = 0 \qquad \textit{implies} \qquad \forall k, T : R_T^k \;\le\; 2\sqrt{V_T^k \ln \frac{\ln T}{\pi_k}}$$

Why? Thinking about $\eta = \sqrt{\frac{\ln \frac{\ln T}{\pi_k}}{V_T^k}}$ gives

$$\Phi_T \;=\; \sum_{k=1}^K \pi_k \int_0^{\frac{1}{2}} \frac{e^{\eta R_T^k - \eta^2 V_T^k} - 1}{\eta} \mathrm{d}\eta \;\approx\; \sum_{k=1}^K \pi_k e^{\sqrt{\frac{\ln \frac{\ln T}{\pi_k}}{V_T^k}} R_T^k - \ln \frac{\ln T}{\pi_k}} - \ln T$$

In fact the quantile upgrade is also true:

$$\forall \boldsymbol{q} \in \triangle_K, T : \mathop{\mathbb{E}}_{k \sim \boldsymbol{q}} \left[R_T^k\right] \;\le\; 2\sqrt{\mathop{\mathbb{E}}_{k \sim \boldsymbol{q}} \left[V_T^k\right] \left(\mathrm{KL}(\boldsymbol{q} \| \boldsymbol{\pi}) + \ln \ln T\right)}$$

E-Lessons

We should explore

- deeply improper priors
- supermartingales possibly negative yet bounded below
- Mixtures and duality of KL (Donsker-Varadhan)

E-Lessons

We should explore

- deeply improper priors
- supermartingales possibly negative yet bounded below
- Mixtures and duality of KL (Donsker-Varadhan)

Muriel's talk

Intermezzo

Many cool ideas/extensions/techniques

Upgrade to online convex optimisation (continuously many actions).

MetaGrad (van Erven and Koolen, 2016)

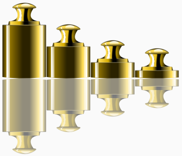
Black-Box reductions (Cutkosky and Orabona, 2018)

FreeGrad (Mhammedi and Koolen, 2020), this blog post E-laborates

Coin Betting (Orabona and Pal, 2016)

Muscada

Multi-Scale Update to the Protocol



Fix a vector $\boldsymbol{\sigma} \in (0, \infty)^K$ of positive loss ranges.

Now let's say the losses $\boldsymbol{\ell}_t$ are such that $\ell_t^k \in [\pm\sigma_k]$.

We want regret bounded by

$$\forall k : R_T^k \;\leq\; \sigma_k \sqrt{T \ln K}$$

Connection to chaining.

Failure

Let's try something akin to

$$\Phi_T = \sum_k \frac{1}{K} e^{\eta_k R_T^k - T\eta_k^2/2}$$

Recall that

$$e^{\eta r_t^k - \eta^2/2} \qquad \text{where} \qquad r_t^k = \boldsymbol{w}_t^\intercal \boldsymbol{\ell}_t - \ell_t^k$$

is an e-value for $r_t^k \in [\pm 1]$, which follows from $\boldsymbol{\ell}_t \in [0,1]^K$. But now $\ell_t^k \in [\pm\sigma_k]$.

Problem For any k, even with σ_k small, $|r_t^k|$ can be as high as $\max_j \sigma_j$.

Muscada Supermartingale

Inspiration:

Fact (Duality for KL)

For any $\boldsymbol{\pi} \in \triangle$ *and* $\boldsymbol{X} \in \mathbb{R}^K$, $\ln \sum_k \pi_k e^{X_k} = \max_{\boldsymbol{w} \in \triangle_K} \langle \boldsymbol{w}, \boldsymbol{X} \rangle - \mathrm{KL}(\boldsymbol{w} \| \boldsymbol{\pi})$.

Muscada Supermartingale

Inspiration:

Fact (Duality for KL)

For any $\boldsymbol{\pi} \in \triangle$ *and* $\boldsymbol{X} \in \mathbb{R}^K$, $\ln \sum_k \pi_k e^{X_k} = \max_{\boldsymbol{w} \in \triangle_K} \langle \boldsymbol{w}, \boldsymbol{X} \rangle - \mathrm{KL}(\boldsymbol{w} \| \boldsymbol{\pi})$.

Define $\boldsymbol{\mu}_T$ by $\mu_T^k := \sigma_k \sqrt{T \ln K}$. Recall that we want $R_T^k \leq \mu_T^k$.

Definition (Muscada supermartingale)

$$\Phi_T \;:=\; \Phi(\boldsymbol{R}_T - \boldsymbol{\mu}_T, \boldsymbol{\eta}_T) \;:=\; \max_{\boldsymbol{w} \in \triangle(K)} \langle \boldsymbol{w}, \boldsymbol{R}_T - \boldsymbol{\mu}_T \rangle - D_{\boldsymbol{\eta}_T}(\boldsymbol{w}, \boldsymbol{u}).$$

where for $\boldsymbol{w}, \boldsymbol{u} \in \triangle_K$ the relative entropy at multi-scale $\boldsymbol{\eta}$ is

$$D_{\boldsymbol{\eta}}(\boldsymbol{w}, \boldsymbol{u}) \;=\; \sum_{k=1}^{K} \frac{w_k \ln(w_k / u_k) - w_k + u_k}{\eta_k}$$

Muscada Analysis

Recall $\mu_T^k := \sigma_k \sqrt{T \ln K}$, and let us fix $\eta_T^k \approx \frac{1}{\sigma_k}\sqrt{\frac{\ln K}{T}}$

$$
\begin{aligned}
\Phi_t &\leq \Phi(\boldsymbol{R}_t - \boldsymbol{\mu}_t, \boldsymbol{\eta}_{t-1}) && \boldsymbol{\eta} \mapsto D_{\boldsymbol{\eta}} \text{ decr.} \\
&= \Phi(\boldsymbol{R}_t - \boldsymbol{\mu}_{t-1} - 4\boldsymbol{\eta}_{t-1}\boldsymbol{\sigma}^2, \boldsymbol{\eta}_{t-1}) && \text{by def. of } \boldsymbol{\mu}_t \\
&\leq \Phi(\boldsymbol{R}_{t-1} - \boldsymbol{\mu}_{t-1}, \boldsymbol{\eta}_{t-1}) && \text{by range control} \\
&= \max_{\boldsymbol{w} \in \triangle(K)} \langle \boldsymbol{w}, \boldsymbol{R}_{t-1} - \boldsymbol{\mu}_{t-1} \rangle - D_{\boldsymbol{\eta}_{t-1}}(\boldsymbol{w}, \boldsymbol{u}) && \text{by def. of } \Phi \\
&= \langle \boldsymbol{w}_t, \boldsymbol{R}_{t-1} - \boldsymbol{\mu}_{t-1} \rangle - D_{\boldsymbol{\eta}_{t-1}}(\boldsymbol{w}_t, \boldsymbol{u}) && \text{by def. of } \boldsymbol{w}_t \\
&\leq \max_{\boldsymbol{w} \in \triangle(K)} \langle \boldsymbol{w}, \boldsymbol{R}_{t-1} - \boldsymbol{\mu}_{t-1} \rangle - D_{\boldsymbol{\eta}_{t-1}}(\boldsymbol{w}, \boldsymbol{u}) && \text{since } \boldsymbol{w}_t \in \triangle(K) \\
&= \Phi(\boldsymbol{R}_{t-1} - \boldsymbol{\mu}_{t-1}, \boldsymbol{\eta}_{t-1}) \;=\; \Phi_{t-1} && \text{by def. of } \Phi, \Phi_t.
\end{aligned}
$$

Hence, $\Phi_t \leq \Phi_{t-1}$, as we were to show.

Postmortem

Indeed

$$R_T^k \;\leq\; \sigma_k \sqrt{T \ln K}$$

E-lessons

- Should investigate how to combine test supermartingales with subtle dependence

Conclusion

Conclusion

- Many cool relations between testing and learning
- Let's talk more!

Thanks!

References i

Chernov, A. and V. Vovk (2009). **"Prediction with expert evaluators' advice".** In: *CoRR* abs/0902.4127.

Cutkosky, A. and F. Orabona (2018). **"Black-Box Reductions for Parameter-free Online Learning in Banach Spaces".** In: *Proceedings of the 31st Annual Conference on Learning Theory (COLT)*. Vol. 75.

Freund, Y. and R. E. Schapire (1997). **"A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting".** In: *Journal of Computer and System Sciences* 55.

Koolen, W. M. and T. van Erven (June 2015). **"Second-order Quantile Methods for Experts and Combinatorial Games".** In: *Proceedings of the 28th Annual Conference on Learning Theory (COLT)*.

References ii

Mhammedi, Z. and W. M. Koolen (July 2020). **"Lipschitz and Comparator-Norm Adaptivity in Online Learning".** In: *Proceedings of the 33rd Annual Conference on Learning Theory (COLT)*.

Orabona, F. and D. Pal (2016). **"Coin Betting and Parameter-Free Online Learning".** In: *Advances in Neural Information Processing Systems 29*.

Van Erven, T. and W. M. Koolen (Dec. 2016). **"MetaGrad: Multiple Learning Rates in Online Learning".** In: *Advances in Neural Information Processing Systems (NeurIPS) 29*.