Efficient Minimax Strategies for Square Loss Games

Wouter M. Koolen Alan Malek Peter L. Bartlett

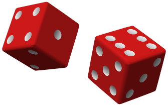
Queensland University of Technology

Berkeley UNIVERSITY OF CALIFORNIA

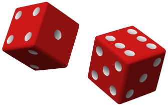
Large-scale Online Learning and Decision Making Workshop
September 11, 2014

Motivation

- Interested in foundations of OLDM.
- Learning formulated as sequential game of regret minimisation.
- Minimax optimal strategy
  - known in a few cases (NML, GDE, L^*-experts, ...)
  - tractable in even fewer



Motivation



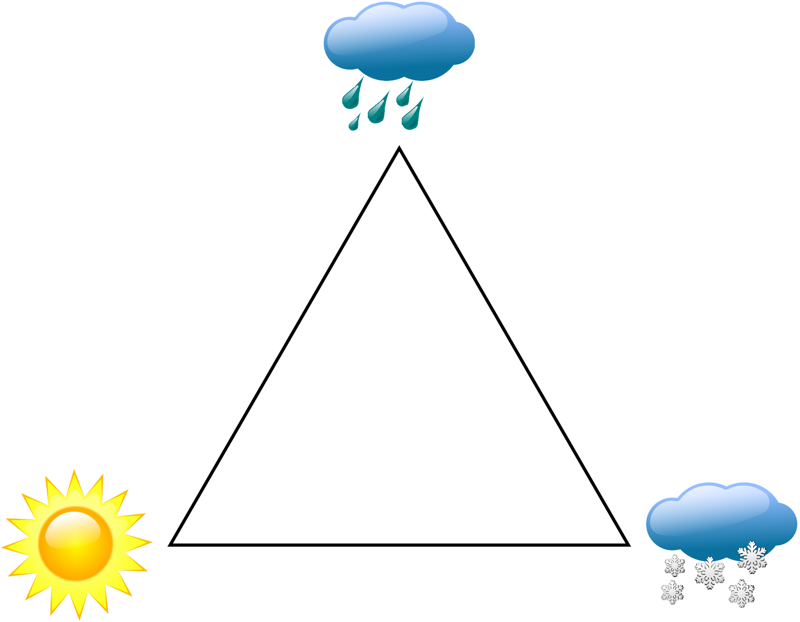
- Interested in foundations of OLDM.
- Learning formulated as sequential game of regret minimisation.
- Minimax optimal strategy
  - known in a few cases (NML, GDE, L^*-experts, ...)
  - tractable in even fewer
- We stumbled across two natural games with efficient minimax solutions.
- So efficient that we dared to submit to **LS**OLDM.

Outline

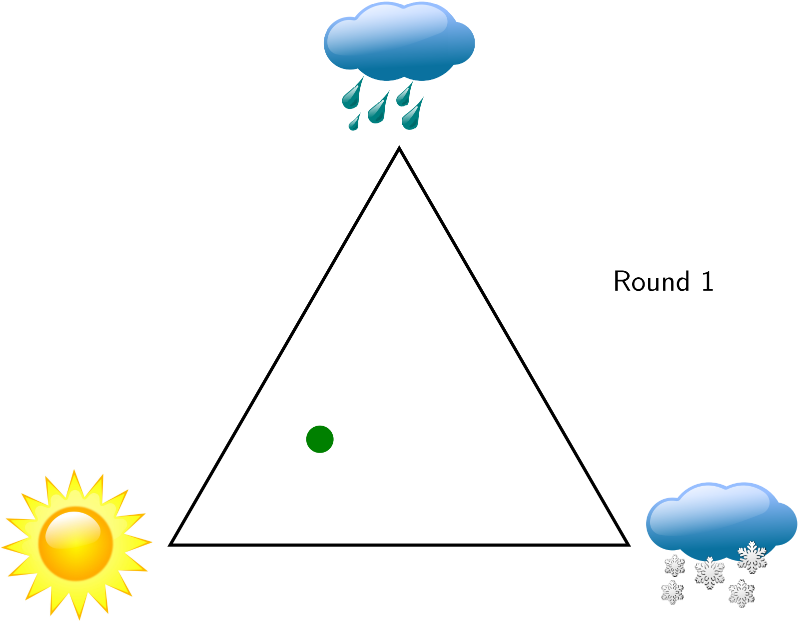
Section 1

Brier game

Brier game: weather prediction example

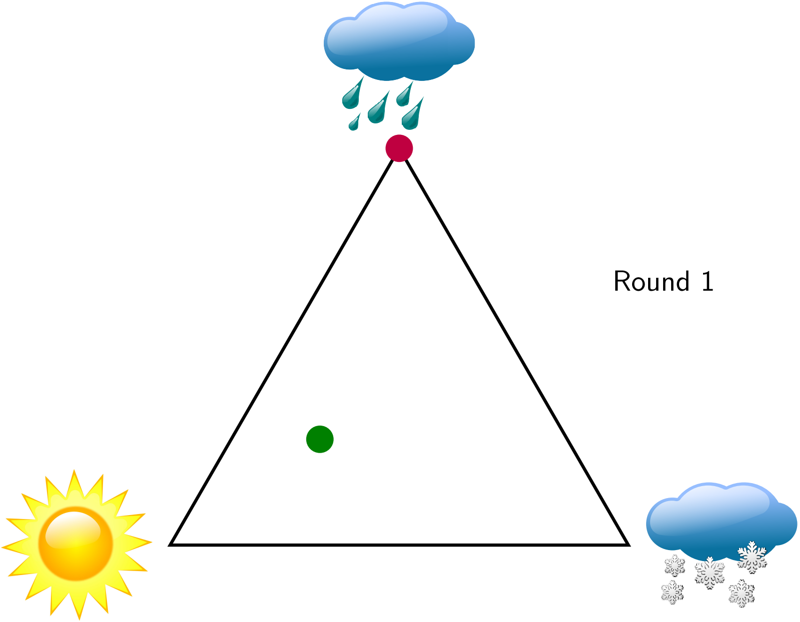


Brier game: weather prediction example

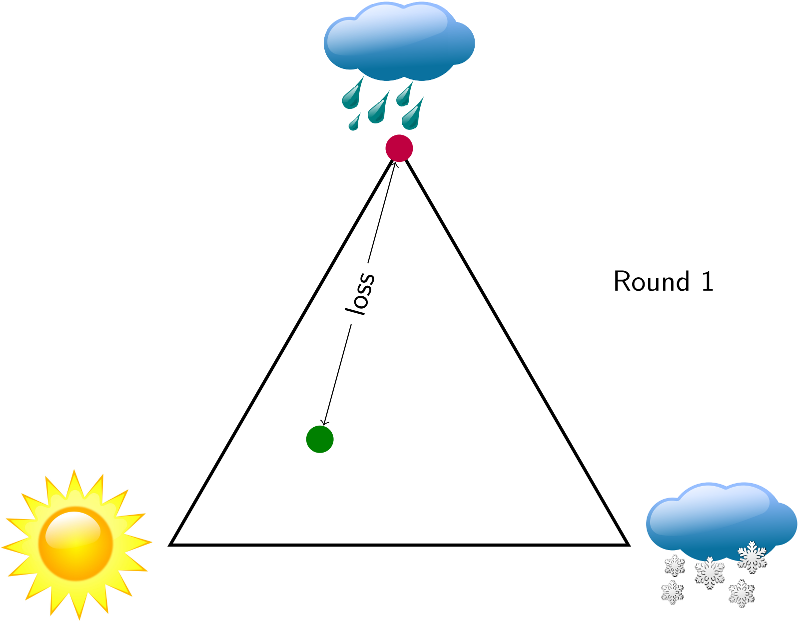


Round 1

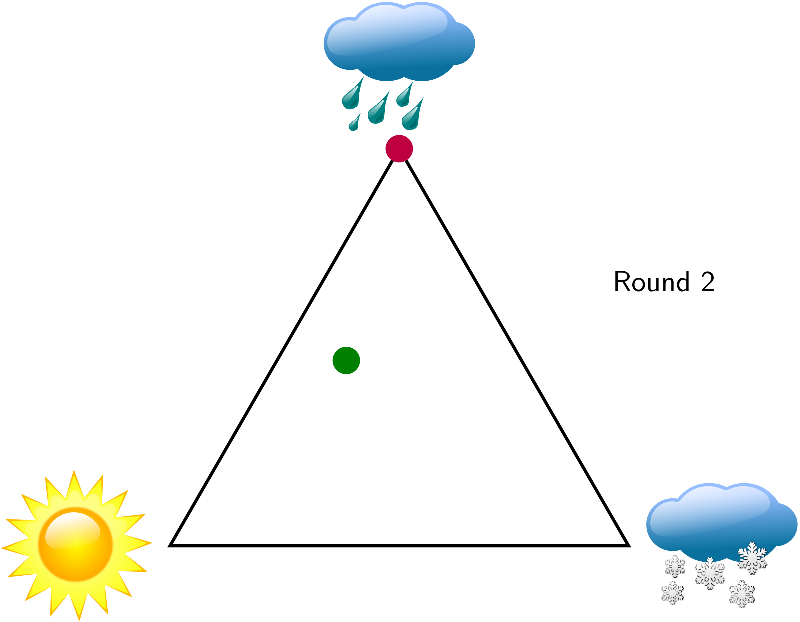
Brier game: weather prediction example



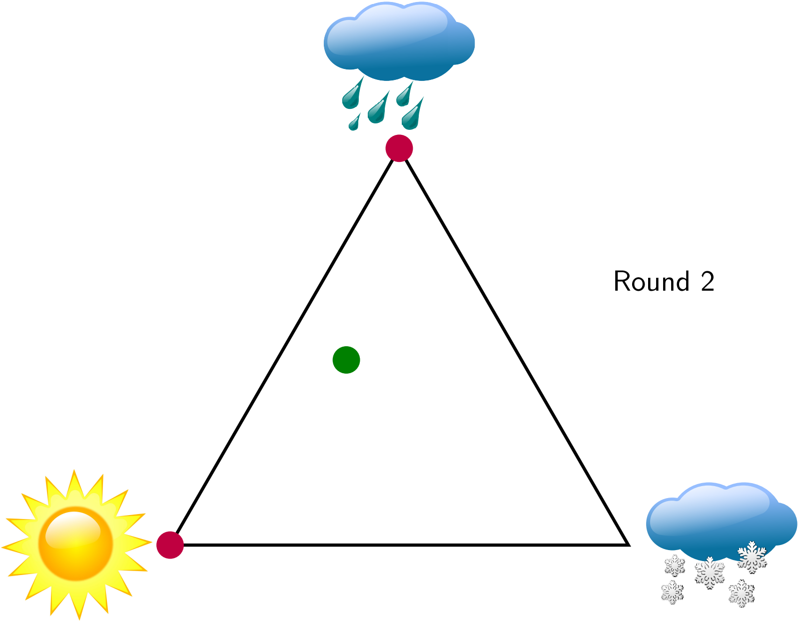
Brier game: weather prediction example



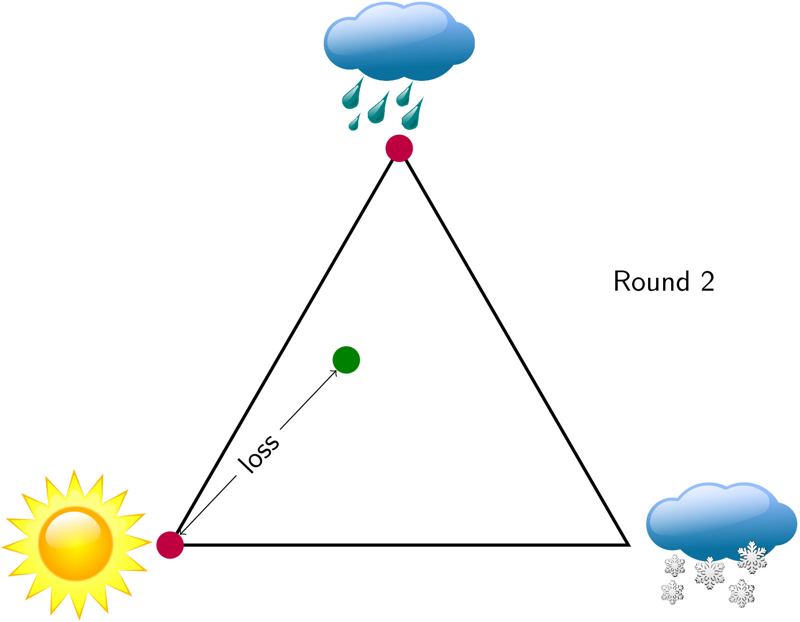
Brier game: weather prediction example



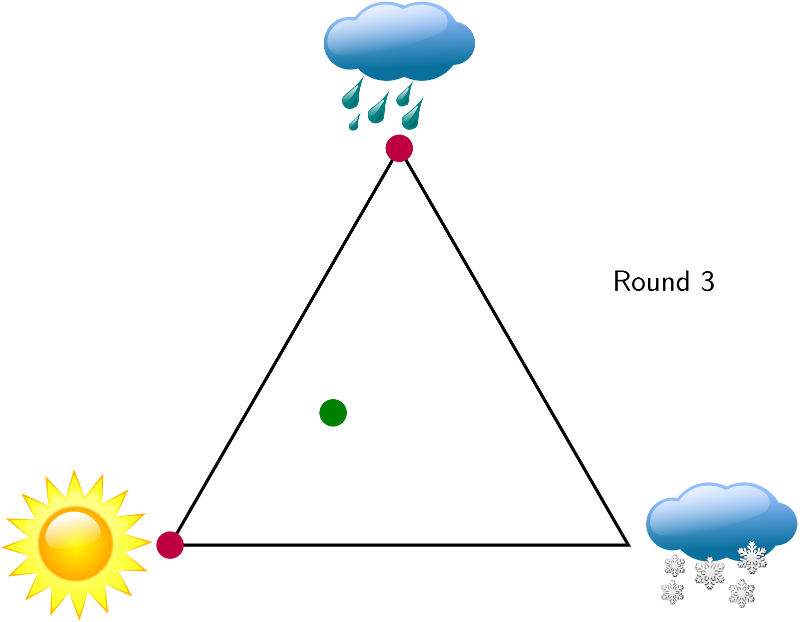
Brier game: weather prediction example



Brier game: weather prediction example

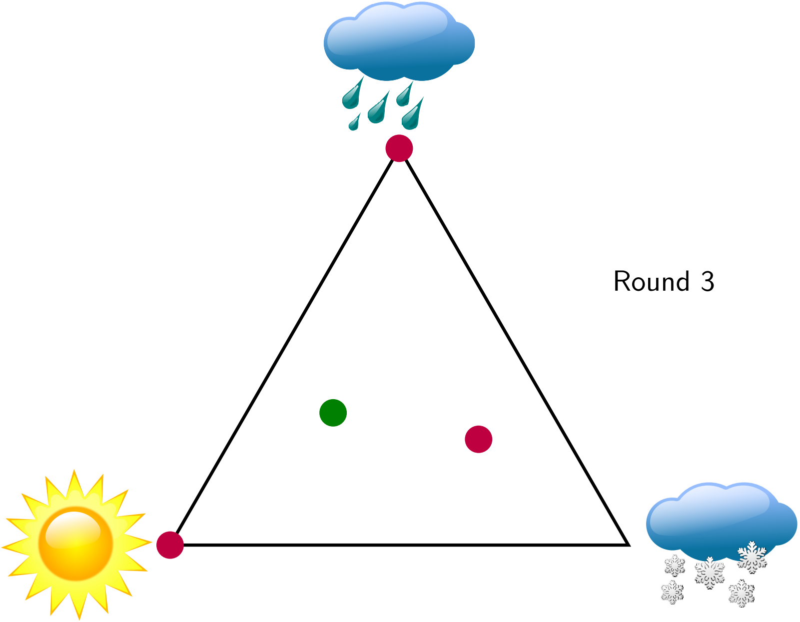


Brier game: weather prediction example

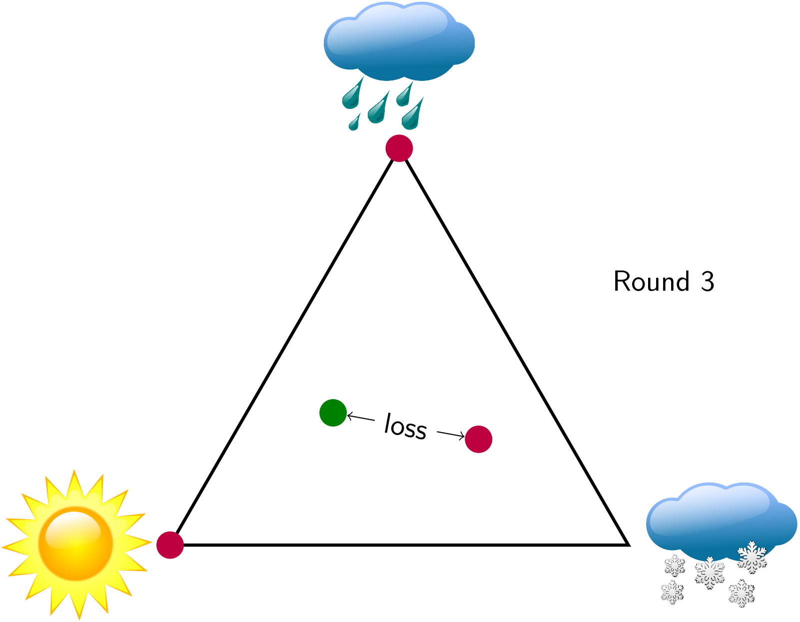


Round 3

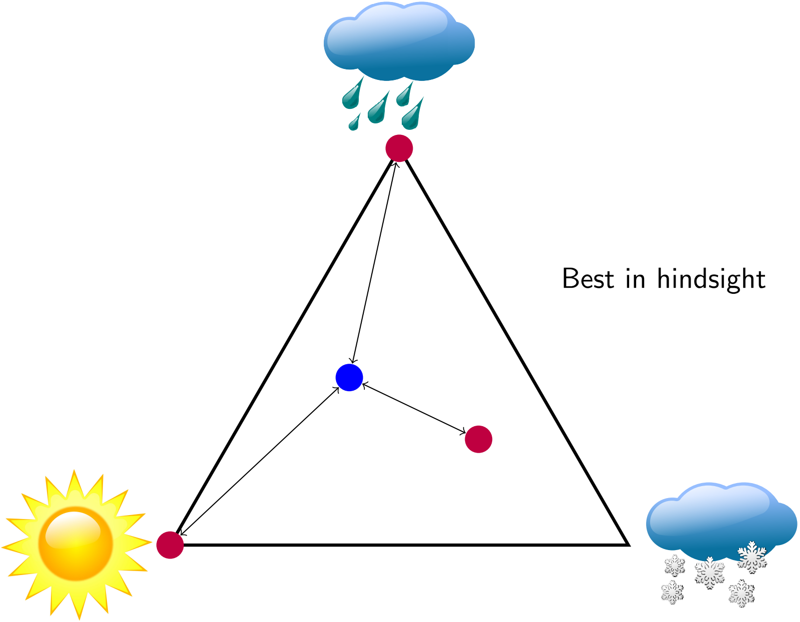
Brier game: weather prediction example



Brier game: weather prediction example



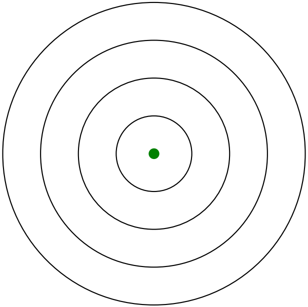
Brier game: weather prediction example



Brier loss

Brier loss equals squared Euclidean distance between $\boldsymbol{a}, \boldsymbol{x} \in \triangle$:

$$\|\boldsymbol{a} - \boldsymbol{x}\|^2 \; = \; (\boldsymbol{a} - \boldsymbol{x})^\intercal (\boldsymbol{a} - \boldsymbol{x})$$



Square loss is proper, convex, and bounded.

Objective: close to best prediction

Learner	$\boldsymbol{a}_1$	$\boldsymbol{a}_2$	$\ldots$	$\boldsymbol{a}_T$
Nature	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\ldots$	$\boldsymbol{x}_T$

$$\text{Regret} \; := \; \sum_{t=1}^{T} \|\boldsymbol{a}_t - \boldsymbol{x}_t\|^2 - \min_{\boldsymbol{a}} \sum_{t=1}^{T} \|\boldsymbol{a} - \boldsymbol{x}_t\|^2$$

Minimax regret

Problem:

$$\min_{\boldsymbol{a}_1} \max_{\boldsymbol{x}_1} \ldots \min_{\boldsymbol{a}_T} \max_{\boldsymbol{x}_T} \left(\sum_{t=1}^{T} \|\boldsymbol{a}_t - \boldsymbol{x}_t\|^2 - \min_{\boldsymbol{a}} \sum_{t=1}^{T} \|\boldsymbol{a} - \boldsymbol{x}_t\|^2 \right)$$

Minimax regret

Problem:

$$\min_{\boldsymbol{a}_1} \max_{\boldsymbol{x}_1} \ldots \min_{\boldsymbol{a}_T} \max_{\boldsymbol{x}_T} \left(\sum_{t=1}^{T} \|\boldsymbol{a}_t - \boldsymbol{x}_t\|^2 - \min_{\boldsymbol{a}} \sum_{t=1}^{T} \|\boldsymbol{a} - \boldsymbol{x}_t\|^2 \right)$$

Game-theoretic analysis gives us:

- Minimax strategy $\boldsymbol{a}_t$
- Maximin strategy $\boldsymbol{x}_t$
- Value of the game (minimax regret).

Recurrence

Value-to-go:

$$V(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T) \;:=\; -\min_{\boldsymbol{a}} \sum_{t=1}^{T} \|\boldsymbol{a} - \boldsymbol{x}_t\|^2$$

$$V(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{t-1}) \;:=\; \min_{\boldsymbol{a}_t} \max_{\boldsymbol{x}_t} \Big(\|\boldsymbol{a}_t - \boldsymbol{x}_t\|^2 + V(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t) \Big)$$

The minimax regret equals value-to-go $V(\epsilon)$ from empty history.

Our approach: manual backwards induction ...

Crux

For each $0 \leq t \leq T$ the value-to-go

$$V(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t)$$

is *quadratic* function of simple statistics

$$\sum_{s=1}^{t} \boldsymbol{x}_s \quad \text{and} \quad \sum_{s=1}^{t} \boldsymbol{x}_s^\intercal \boldsymbol{x}_s.$$

Idea: proof by induction. Base case $t = T$ is easy. Induction step hinges on single-round min-max solution.

Consequences I

In the state $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ with statistics $\boldsymbol{s} = \sum_{t=1}^{n} \boldsymbol{x}_t$ and $\sigma^2 = \sum_{t=1}^{n} \boldsymbol{x}_t^\intercal \boldsymbol{x}_t$ the Brier game on $\triangle_d$ has value-to-go

$$V(\boldsymbol{s}, \sigma^2) \;=\; \alpha_n \boldsymbol{s}^\intercal \boldsymbol{s} - \sigma^2 + \mathrm{const}_n$$

and minimax and maximin strategies given by

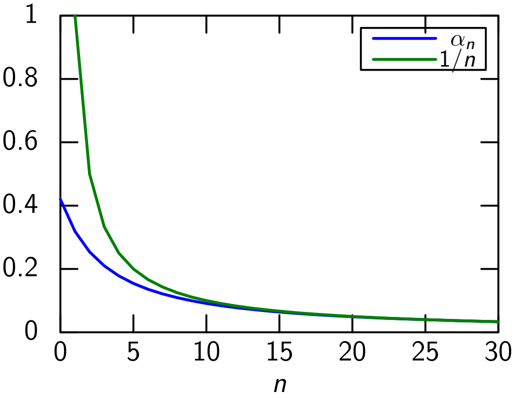
$$\boldsymbol{a}^*(\boldsymbol{s}, \sigma^2) \;=\; \boldsymbol{p}^*(\boldsymbol{s}, \sigma^2) \;=\; \frac{\mathbf{1}}{d} + \alpha_{n+1} \left(\boldsymbol{s} - n \frac{\mathbf{1}}{d} \right)$$

with coefficients defined recursively by

$$\alpha_T \;=\; \frac{1}{T} \qquad\qquad \alpha_{n-1} \;=\; \alpha_n^2 + \alpha_n.$$

Consequences II

- Minimax shrinks Follow-the-Leader towards uniform:



- Computation: $O(T)$ pre-processing, then $O(d)$ per round.
- The regret is at most

$$1 + \ln(T).$$

- Mixed data points are friendly.

Extension: Mahalanobis loss

Gravity of errors differs among dimensions.

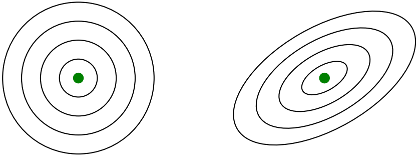
Technical tool: generalise squared Euclidean distance

$$\|\boldsymbol{a} - \boldsymbol{x}\|^2 = (\boldsymbol{a} - \boldsymbol{x})^\intercal (\boldsymbol{a} - \boldsymbol{x})$$

to squared Mahalanobis distance (proper!)

$$\|\boldsymbol{a} - \boldsymbol{x}\|_{\boldsymbol{W}}^2 = (\boldsymbol{a} - \boldsymbol{x})^\intercal \boldsymbol{W}^{-1} (\boldsymbol{a} - \boldsymbol{x})$$

for some fixed coefficient matrix $\boldsymbol{W} \succ \mathbf{0}$.

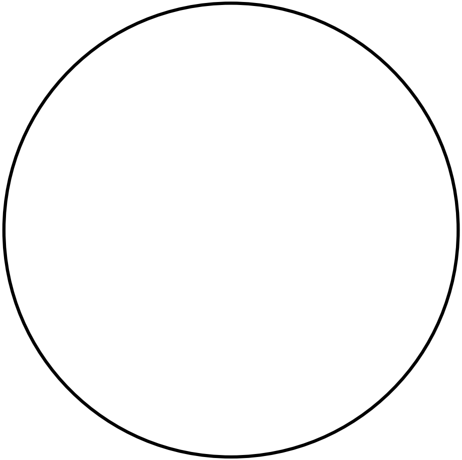


All results scale up (under simplex alignment condition on $\boldsymbol{W}$).

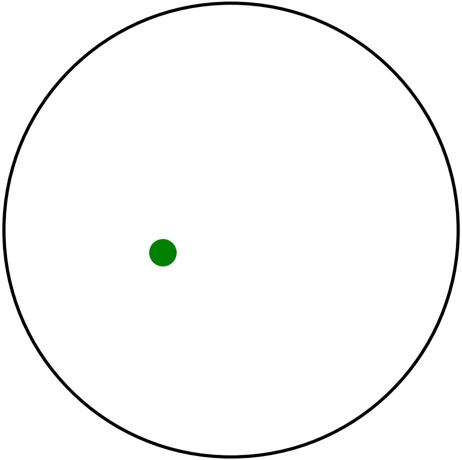
Section 2

Ball game

Ball game

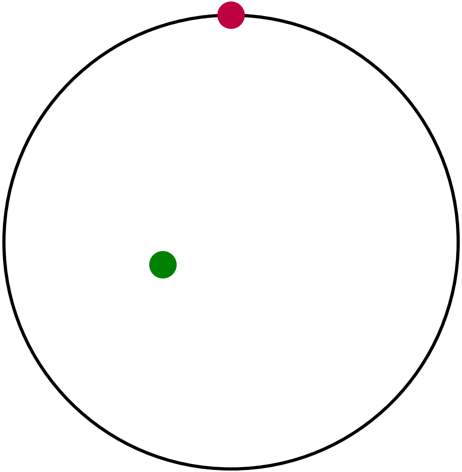


Ball game



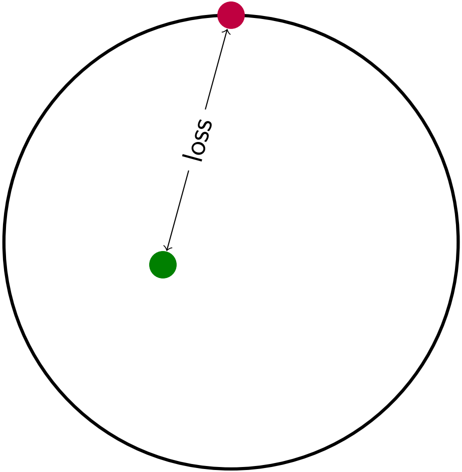
Round 1

Ball game



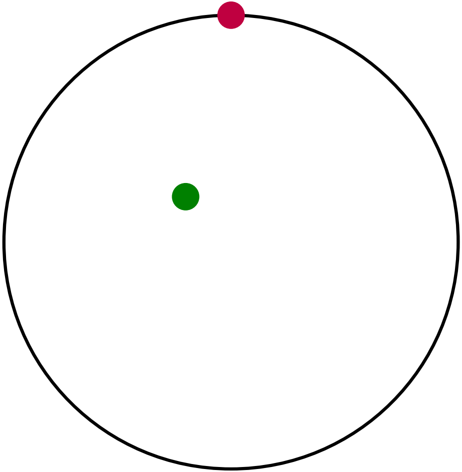
Round 1

Ball game



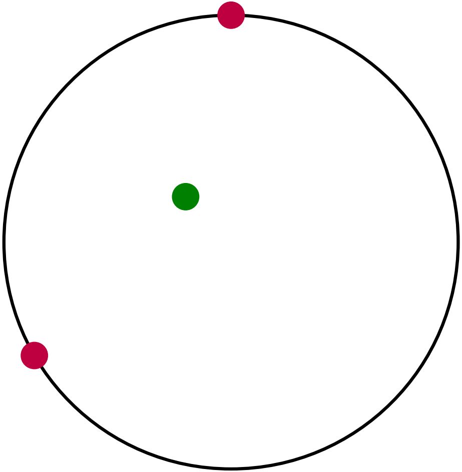
Round 1

Ball game



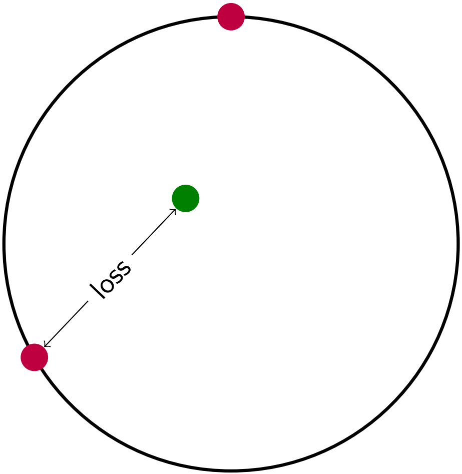
Round 2

Ball game



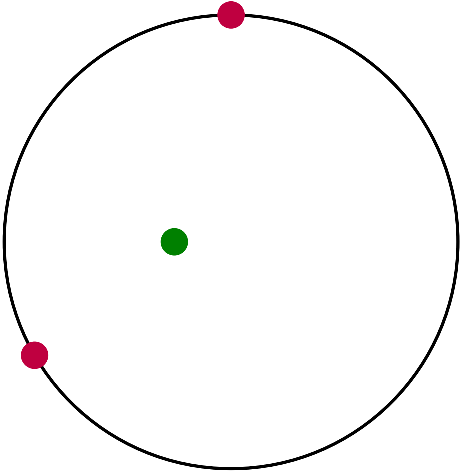
Round 2

Ball game



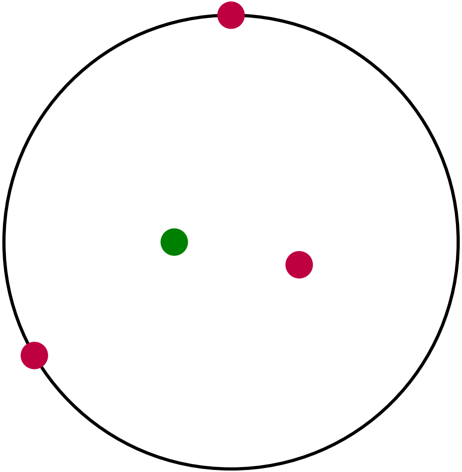
Round 2

Ball game



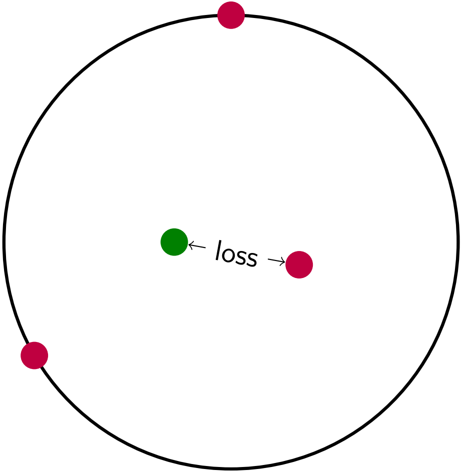
Round 3

Ball game



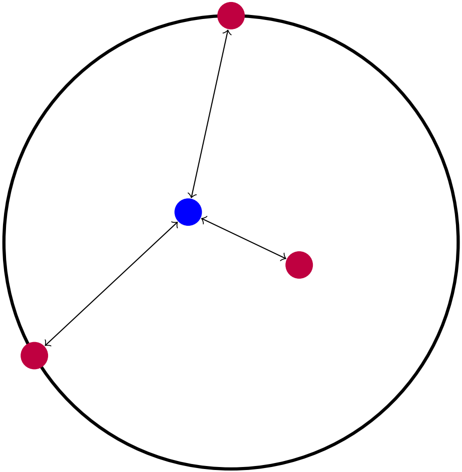
Round 3

Ball game



Round 3

Ball game



Best in hindsight

Minimax regret

Problem:

$$\min_{\boldsymbol{a}_1} \max_{\boldsymbol{x}_1} \dots \min_{\boldsymbol{a}_T} \max_{\boldsymbol{x}_T} \left(\sum_{t=1}^{T} \|\boldsymbol{a}_t - \boldsymbol{x}_t\|_{\boldsymbol{W}}^2 - \min_{\boldsymbol{a}} \sum_{t=1}^{T} \|\boldsymbol{a} - \boldsymbol{x}_t\|_{\boldsymbol{W}}^2 \right)$$

Note: Brier and Ball game only differ in domain of $\boldsymbol{a}_t$ and $\boldsymbol{x}_t$

Minimax analysis

Consider the ball game with loss $\|\boldsymbol{a} - \boldsymbol{x}\|_{\boldsymbol{W}}^2$. The value-to-go for state $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ with statistics $\boldsymbol{s} = \sum_{t=1}^n \boldsymbol{x}_t$ and $\sigma^2 = \sum_{t=1}^n \boldsymbol{x}_t^\intercal \boldsymbol{W}^{-1} \boldsymbol{x}_t$ is

$$V(\boldsymbol{s}, \sigma^2) \;=\; \boldsymbol{s}^\intercal \boldsymbol{A}_n \boldsymbol{s} - \sigma^2 + \mathrm{const}_n.$$

The minimax strategy plays

$$\boldsymbol{a}^*(\boldsymbol{s}, \sigma^2) \;=\; \left(\boldsymbol{W}^{-1} + \lambda_{\max} \boldsymbol{I} - \boldsymbol{A}_{n+1}\right)^{-1} \boldsymbol{A}_{n+1} \boldsymbol{s}$$

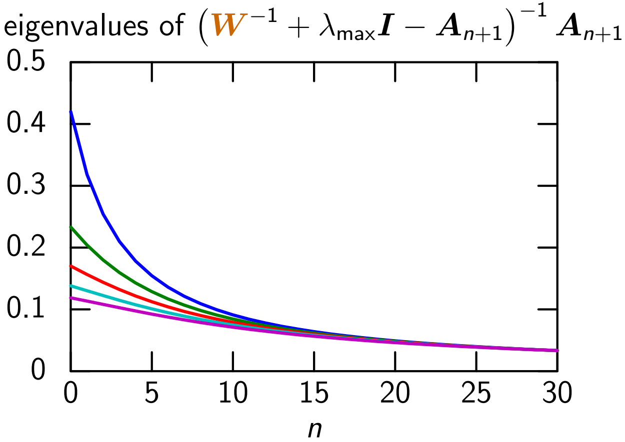
and the maximin strategy plays two unit length vectors with

$$\Pr\left(\boldsymbol{x} = \boldsymbol{a}_\perp \pm \sqrt{1 - \boldsymbol{a}_\perp^\intercal \boldsymbol{a}_\perp}\, \boldsymbol{v}_{\max}\right) = \frac{1}{2} \pm \frac{1}{2}\sqrt{\frac{\boldsymbol{a}_\parallel^\intercal \boldsymbol{a}_\parallel}{1 - \boldsymbol{a}_\perp^\intercal \boldsymbol{a}_\perp}},$$

where $\lambda_{\max}$ and $\boldsymbol{v}_{\max}$ correspond to the largest eigenvalue of $\boldsymbol{A}_{n+1}$ and $\boldsymbol{a}_\perp$ and $\boldsymbol{a}_\parallel$ are the components of $\boldsymbol{a}^*$ perpendicular and parallel to $\boldsymbol{v}_{\max}$. The coefficients $\boldsymbol{A}_n$ are determined recursively by base case $\boldsymbol{A}_T \;=\; \frac{1}{T} \boldsymbol{W}^{-1}$ and recursion

$$\boldsymbol{A}_{n-1} \;=\; \boldsymbol{A}_n \left(\boldsymbol{W}^{-1} + \lambda_{\max} \boldsymbol{I} - \boldsymbol{A}_n\right)^{-1} \boldsymbol{A}_n + \boldsymbol{A}_n.$$

The eigenvalue warp



Brier game had uniform shrinkage.

For ball game shrinkage rate depends on dimension.

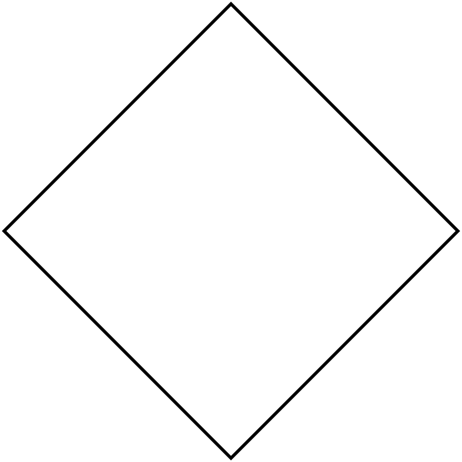
Ball game consequences

- Regret bounded by $\lambda_{\max}(\boldsymbol{W}^{-1})\big(1 + \ln(T)\big)$.
- Computation: $O(Td + d^3)$ pre-processing, $O(d^2)$ per round.
- Outcomes in ball interior are friendly.

Section 3

Diamond game

Counterexample: diamond game



$\boldsymbol{W} = \boldsymbol{I}$ case (accidentally?) works out; value-to-go is quadratic.
$\boldsymbol{W} \neq \boldsymbol{I}$ case fails. Complexity of value-to-go function explodes.

Section 4

Conclusion

Conclusion

- Two games where the minimax strategy can be followed in amortised constant computation per round.
- Value-to-go quadratic function of statistic
- Minimax strategy linear in statistic
  (Follow-the-Leader with subtle shrinkage)

What next

- Characterise interplay of action/outcome sets and loss that results in simple value-to-go function (conjugacy)
- Reduce other similar losses to square loss
- Consider other notions of "squared distance". (Bregman)
- Add covariates (regression)
- Consider non-stationarity
- Other losses (PCA, @#$! hard)
- Horizon-free/anytime algorithms

Thank you!