

# Efficient Minimax Strategies for Square Loss Games



**Wouter M. Koolen**   Alan Malek   Peter L. Bartlett



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

# Motivation

- ▶ Interested in foundations of OLDLM.
- ▶ 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 OLDLM.
- ▶ 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 **LSOLDLM**.



# Outline

Brier game

Ball game

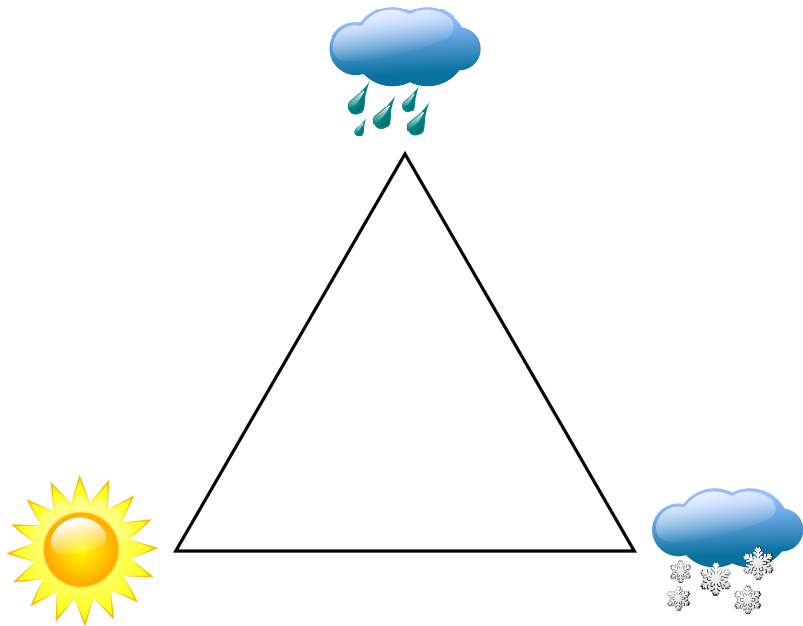
Diamond game

Conclusion

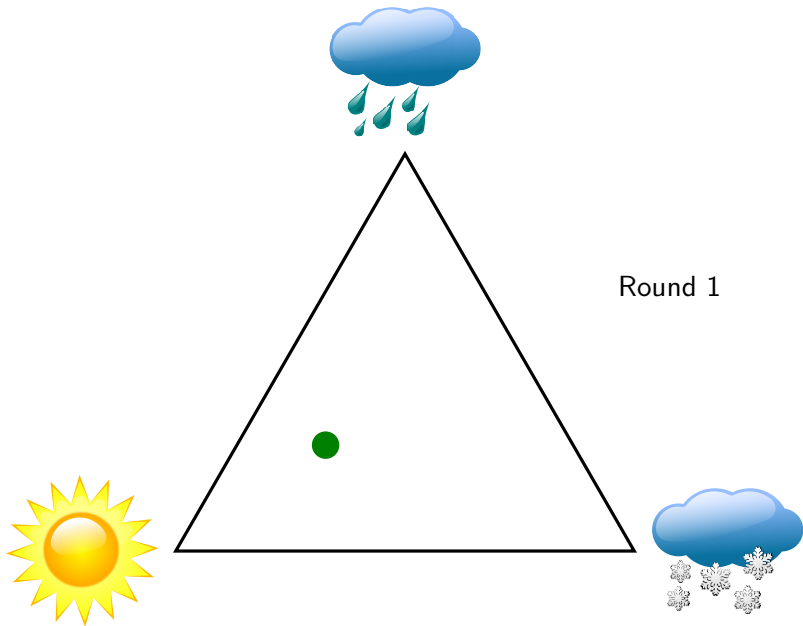
## Section 1

### Brier game

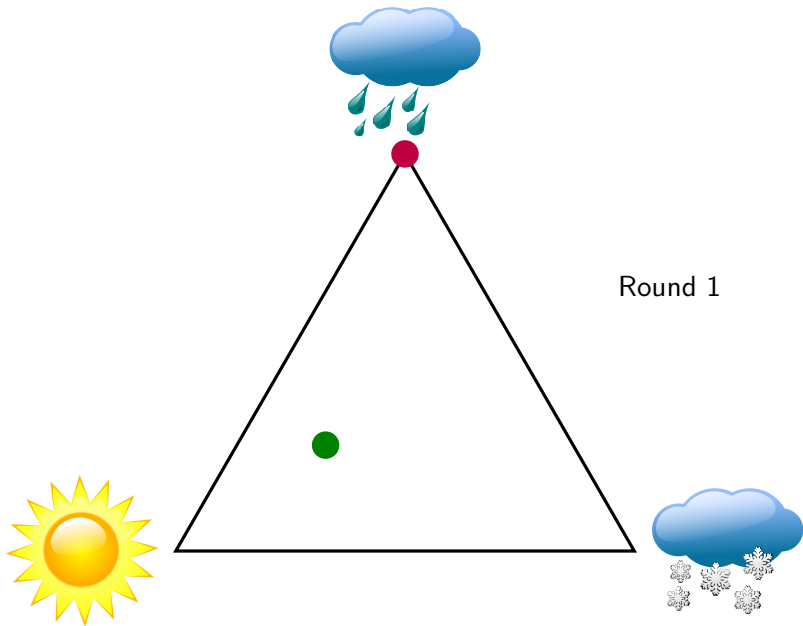
## Brier game: weather prediction example



## Brier game: weather prediction example

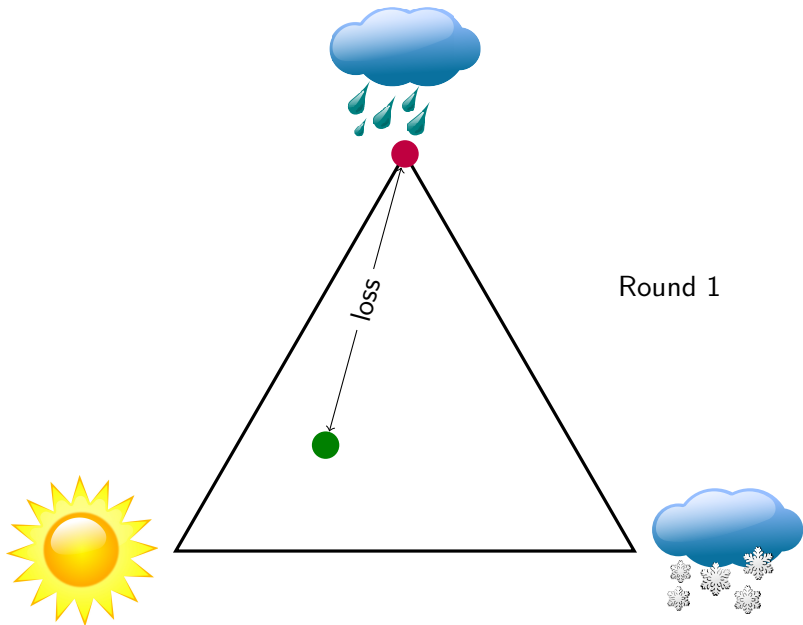


# Brier game: weather prediction example

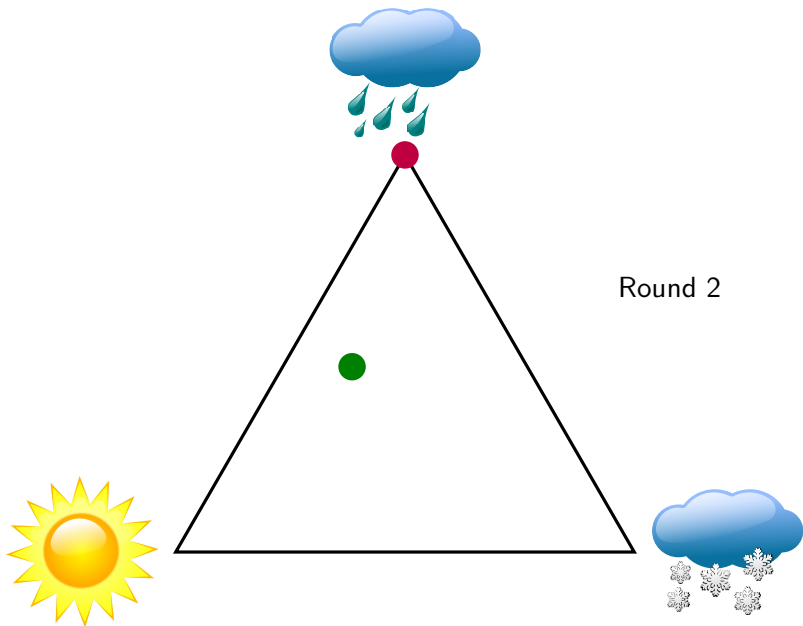




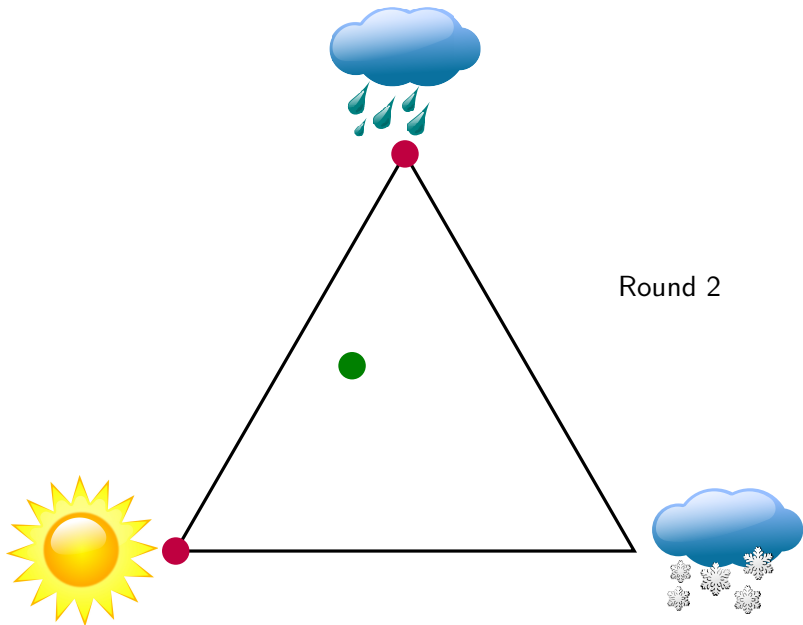
# Brier game: weather prediction example



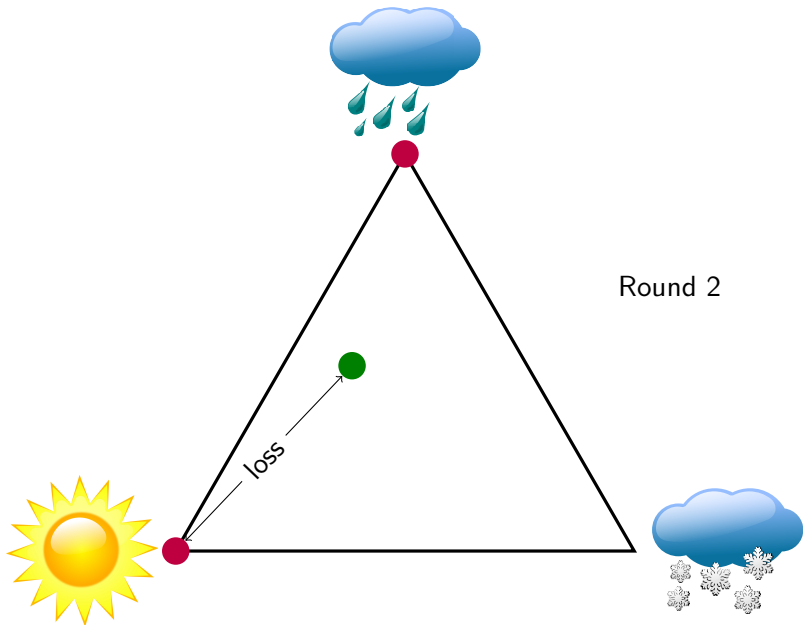
## Brier game: weather prediction example



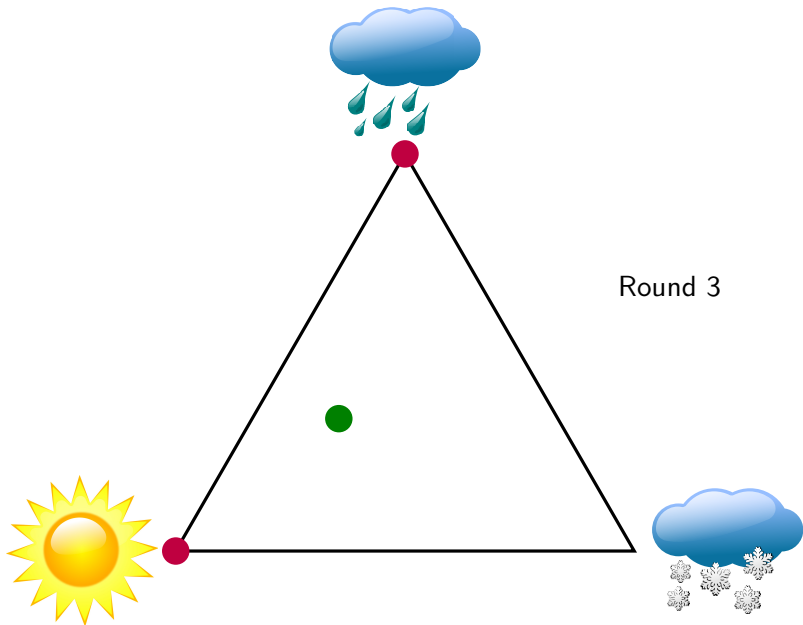
## Brier game: weather prediction example



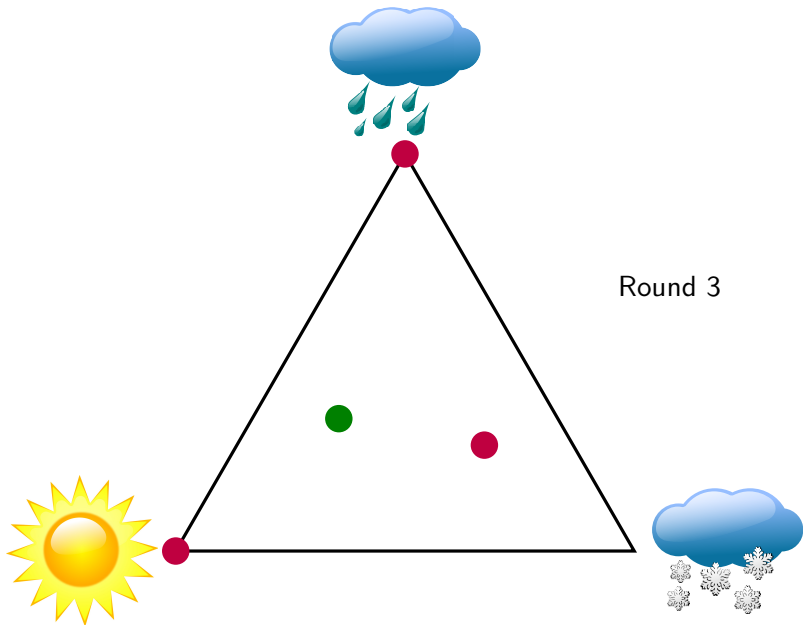
# Brier game: weather prediction example



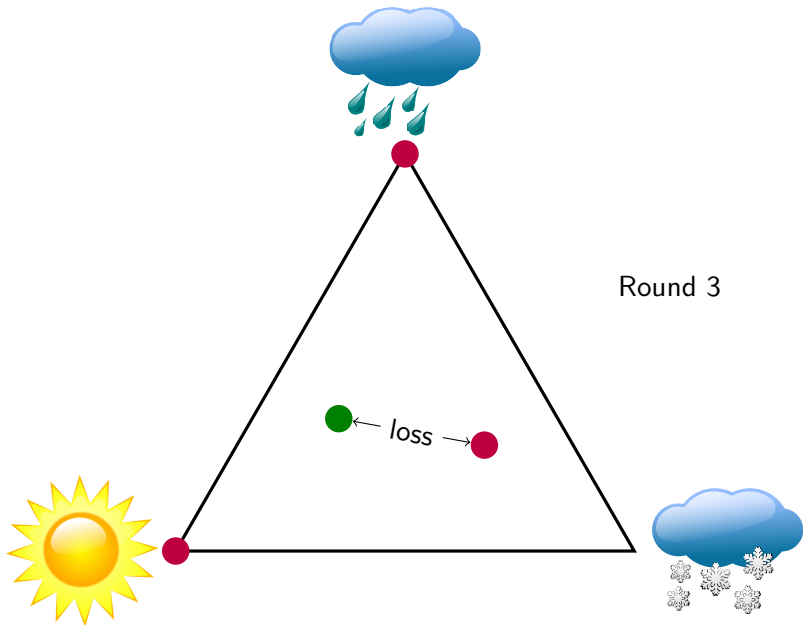
# Brier game: weather prediction example



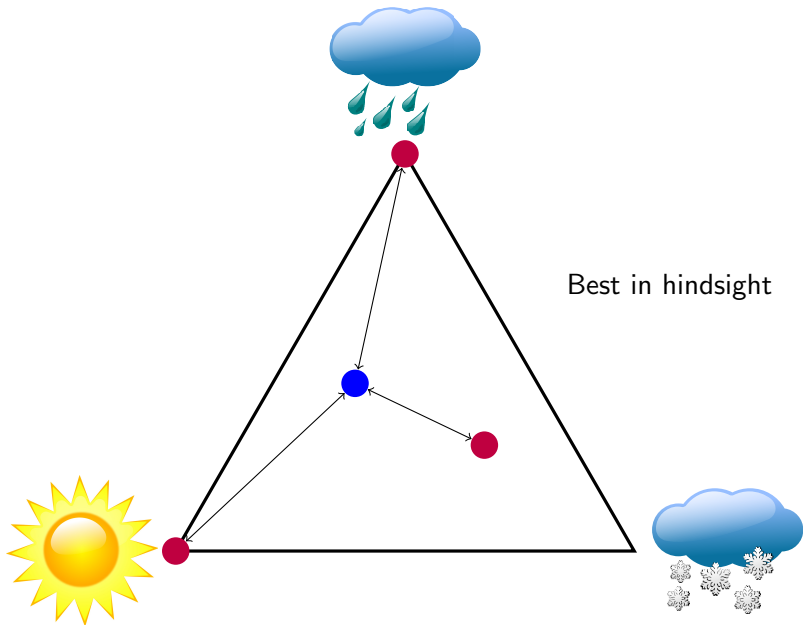
# Brier game: weather prediction example



# Brier game: weather prediction example



# Brier game: weather prediction example

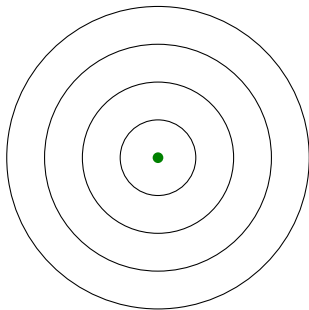




## Brier loss

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

$$\|\mathbf{a} - \mathbf{x}\|^2 = (\mathbf{a} - \mathbf{x})^\top (\mathbf{a} - \mathbf{x})$$



Square loss is proper, convex, and bounded.

Objective: close to best prediction

Learner	$\mathbf{a}_1$	$\mathbf{a}_2$	$\dots$	$\mathbf{a}_T$
Nature	$\mathbf{x}_1$	$\mathbf{x}_2$	$\dots$	$\mathbf{x}_T$

$$\text{Regret} := \sum_{t=1}^T \|\mathbf{a}_t - \mathbf{x}_t\|^2 - \min_a \sum_{t=1}^T \|\mathbf{a} - \mathbf{x}_t\|^2$$

## Minimax regret

Problem:

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

# Minimax regret

Problem:

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

Game-theoretic analysis gives us:

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

# Recurrence

Value-to-go:

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

$$V(\mathbf{x}_1, \dots, \mathbf{x}_{t-1}) := \min_{\mathbf{a}_t} \max_{\mathbf{x}_t} \left( \|\mathbf{a}_t - \mathbf{x}_t\|^2 + V(\mathbf{x}_1, \dots, \mathbf{x}_t) \right)$$

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

$$V(\mathbf{x}_1, \dots, \mathbf{x}_t)$$

is *quadratic* function of simple statistics

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

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

## Consequences I

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

$$V(\mathbf{s}, \sigma^2) = \alpha_n \mathbf{s}^\top \mathbf{s} - \sigma^2 + \text{const}_n$$

and minimax and maximin strategies given by

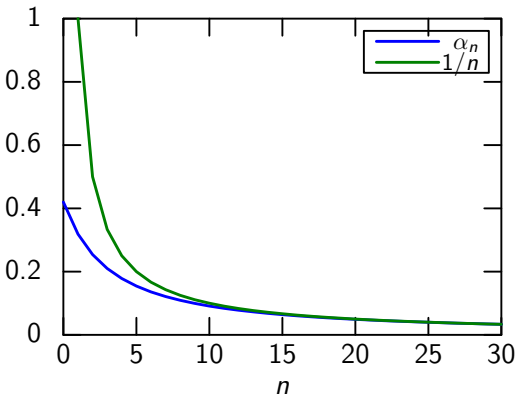
$$\mathbf{a}^*(\mathbf{s}, \sigma^2) = \mathbf{p}^*(\mathbf{s}, \sigma^2) = \frac{\mathbf{1}}{d} + \alpha_{n+1} \left( \mathbf{s} - n \frac{\mathbf{1}}{d} \right)$$

with coefficients defined recursively by

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

## 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.

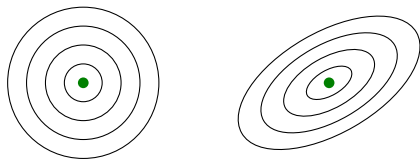
Technical tool: generalise squared Euclidean distance

$$\|\mathbf{a} - \mathbf{x}\|^2 = (\mathbf{a} - \mathbf{x})^\top (\mathbf{a} - \mathbf{x})$$

to squared Mahalanobis distance (proper!)

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

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

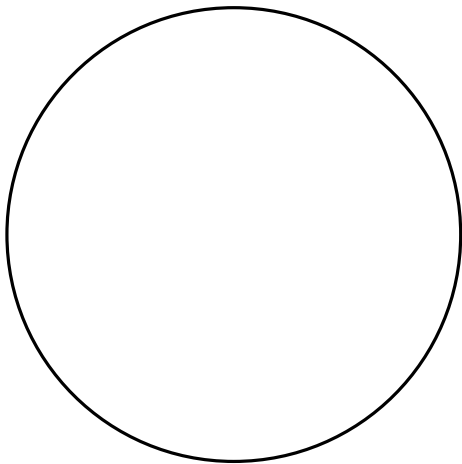


All results scale up (under simplex alignment condition on  $\mathbf{W}$ ).

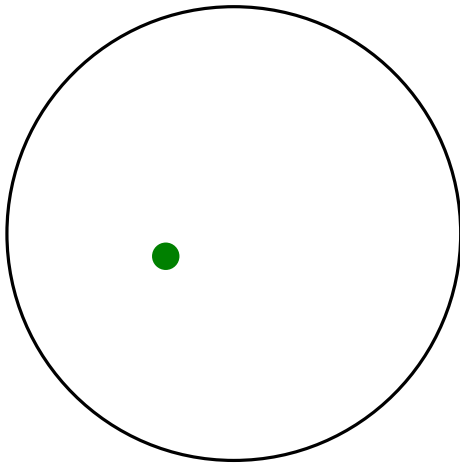
## Section 2

### Ball game

Ball game

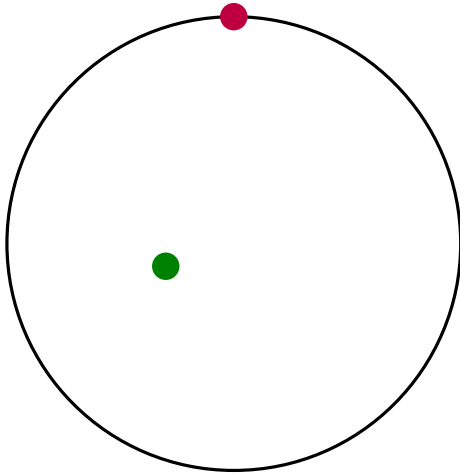


# Ball game



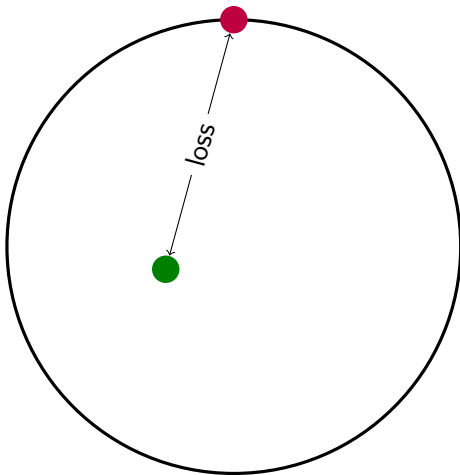
Round 1

# Ball game



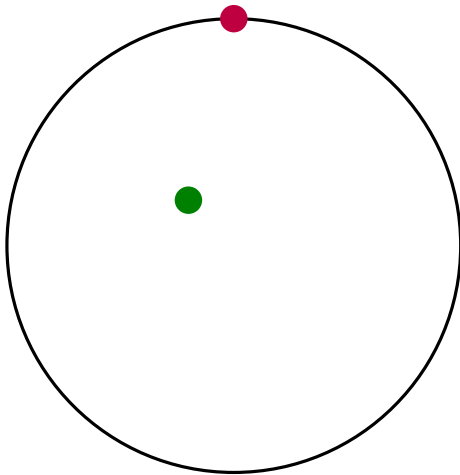
Round 1

# Ball game



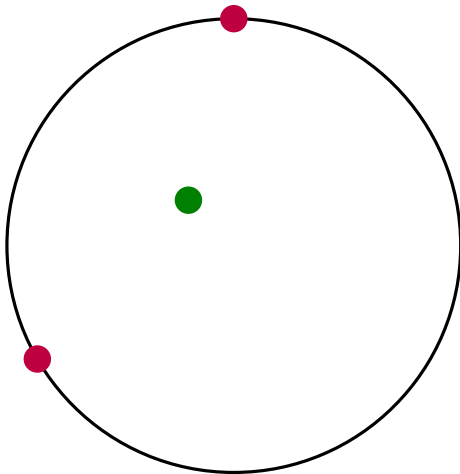
Round 1

# Ball game



Round 2

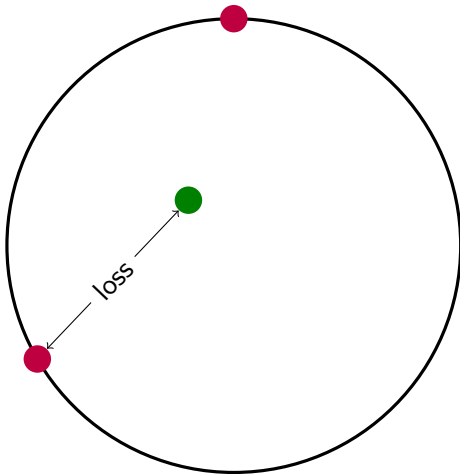
# Ball game



Round 2

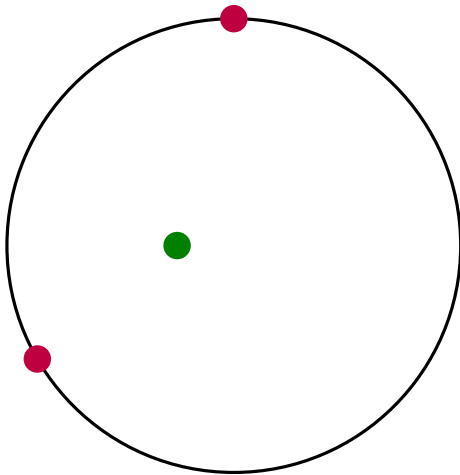


# Ball game



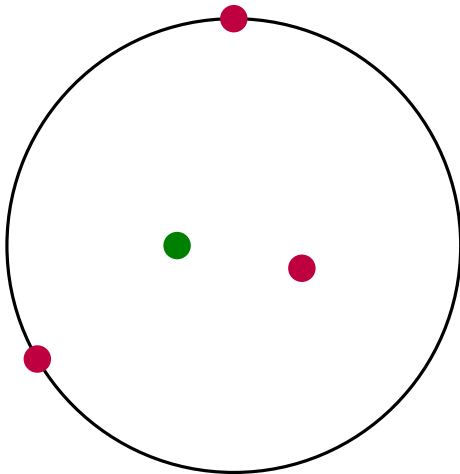
Round 2

# Ball game



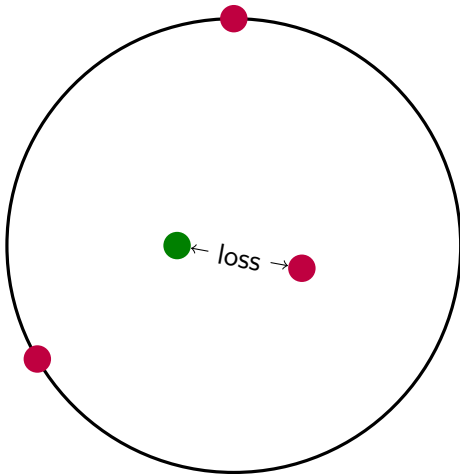
Round 3

# Ball game



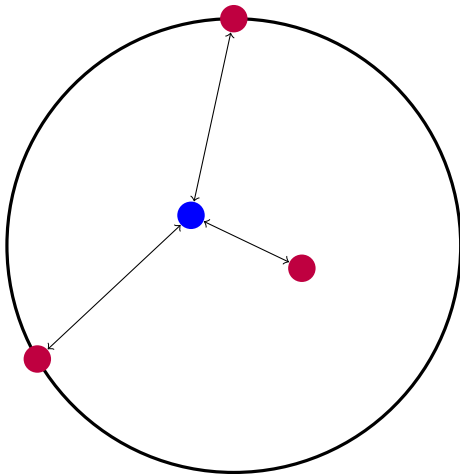
Round 3

# Ball game



Round 3

# Ball game



Best in hindsight

## Minimax regret

Problem:

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

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

## Minimax analysis

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

$$V(\mathbf{s}, \sigma^2) = \mathbf{s}^\top \mathbf{A}_n \mathbf{s} - \sigma^2 + \text{const}_n.$$

The minimax strategy plays

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

and the maximin strategy plays two unit length vectors with

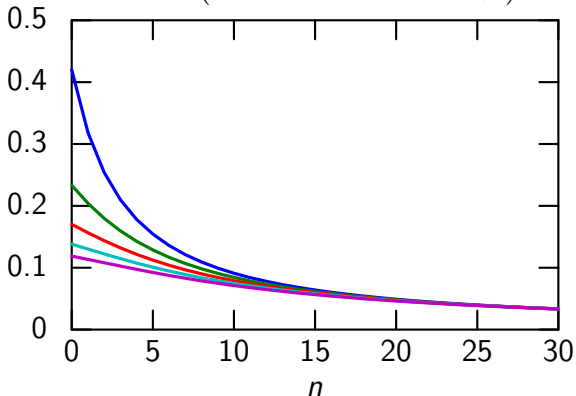
$$\Pr \left( \mathbf{x} = \mathbf{a}_\perp \pm \sqrt{1 - \mathbf{a}_\perp^\top \mathbf{a}_\perp} \mathbf{v}_{\max} \right) = \frac{1}{2} \pm \frac{1}{2} \sqrt{\frac{\mathbf{a}_\parallel^\top \mathbf{a}_\parallel}{1 - \mathbf{a}_\perp^\top \mathbf{a}_\perp}},$$

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

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

## The eigenvalue warp

eigenvalues of  $(\mathbf{W}^{-1} + \lambda_{\max} \mathbf{I} - \mathbf{A}_{n+1})^{-1} \mathbf{A}_{n+1}$



Brier game had uniform shrinkage.

For ball game shrinkage rate depends on dimension.



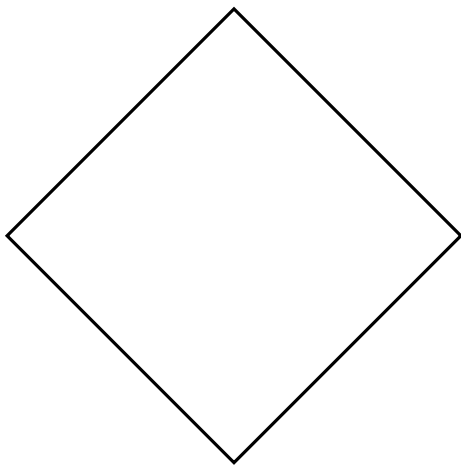
## Ball game consequences

- ▶ Regret bounded by  $\lambda_{\max}(\mathbf{W}^{-1})(1 + \ln(T))$ .
- ▶ 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



$W = I$  case (accidentally?) works out; value-to-go is quadratic.

$W \neq I$  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!