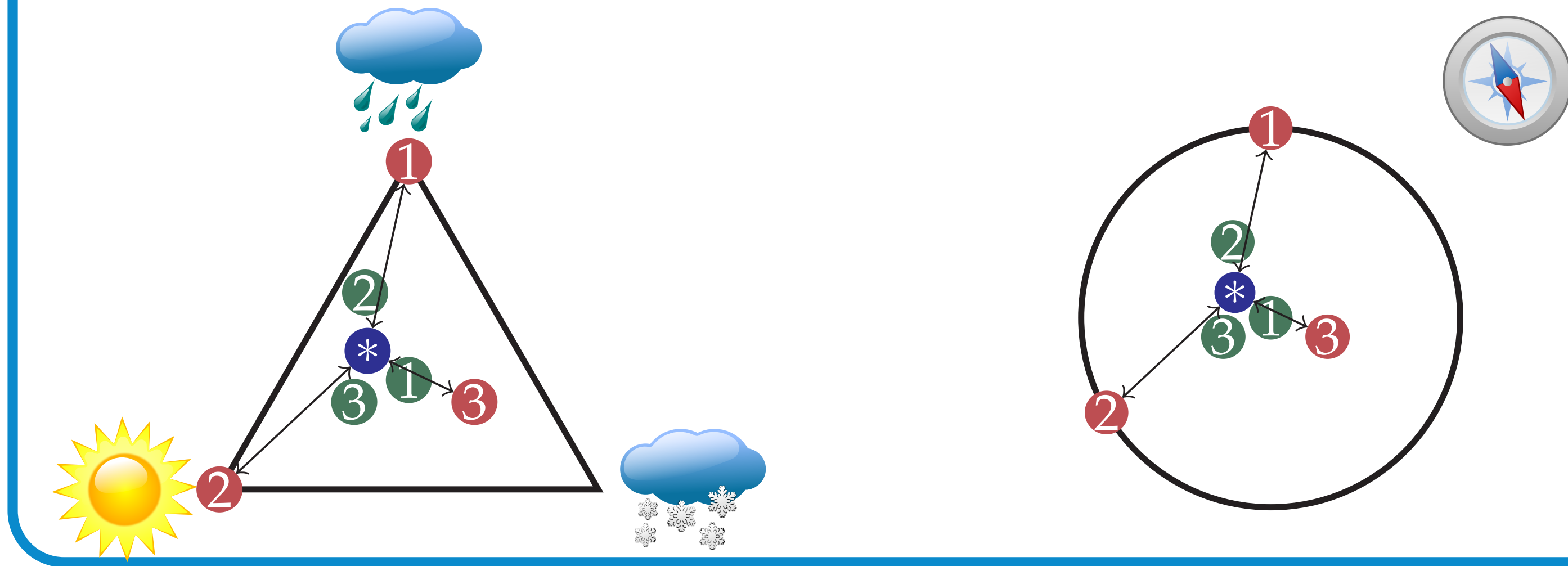


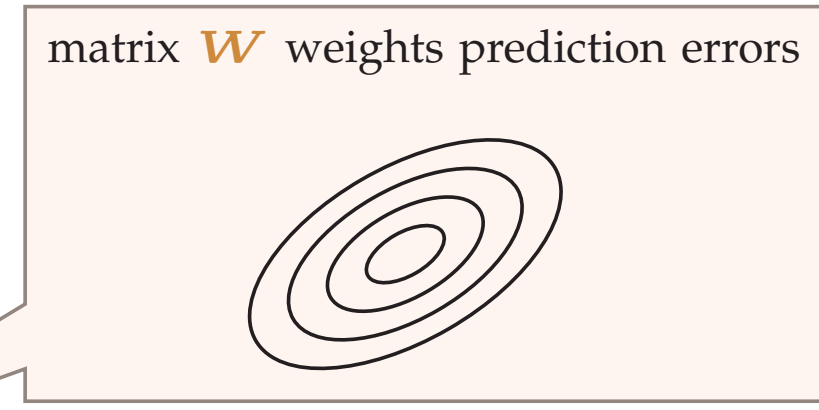
ONLINE PREDICTION



SQUARE LOSS GAMES

Fix a convex set \mathcal{C} , positive definite matrix \mathbf{W} , length T .
For each round $t = 1, \dots, T$,

- We play $\mathbf{a}_t \in \mathcal{C}$
- Nature reveals $\mathbf{x}_t \in \mathcal{C}$
- We incur loss



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

Our goal is to minimize regret w.r.t. best fixed action \mathbf{a} in hindsight

$$\text{Regret} := \sum_{t=1}^T \ell(\mathbf{a}_t, \mathbf{x}_t) - \min_{\mathbf{a}} \sum_{t=1}^T \ell(\mathbf{a}, \mathbf{x}_t)$$

MINIMAX REGRET

If we assume a perfect adversary, how well can we do?

$$V := \min_{\mathbf{a}_1} \max_{\mathbf{x}_1} \dots \min_{\mathbf{a}_T} \max_{\mathbf{x}_T} \text{Regret}$$

We play to minimize the worst-case regret. We can solve for the **value-to-go** V in any history using the recurrence:

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

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

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

To play the minimax strategy: after seeing $\mathbf{x}_1, \dots, \mathbf{x}_{t-1}$,

- Compute $V(\mathbf{x}_1, \dots, \mathbf{x}_t)$
- Choose \mathbf{a}_t as the minimizer of Equation (2)

Problem: this is expensive. Are there examples where V is a simple function of some statistics of $\mathbf{x}_1, \dots, \mathbf{x}_t$ that is simple to *precompute*?

SOLVING THE GAMES

Using sufficient statistics $\mathbf{s} = \sum_{\tau=1}^t \mathbf{x}_\tau$ and $\sigma^2 = \sum_{\tau=1}^t \mathbf{x}_\tau^\top \mathbf{W}^{-1} \mathbf{x}_\tau$,

Theorem 1 (Brier Game) Let $\mathcal{C} = \Delta$. For \mathbf{W} satisfying an alignment condition, the value-to-go is

Theorem 2 (Ball Game) Let $\mathcal{C} = \mathcal{O}$. For any positive definite \mathbf{W} the value-to-go is

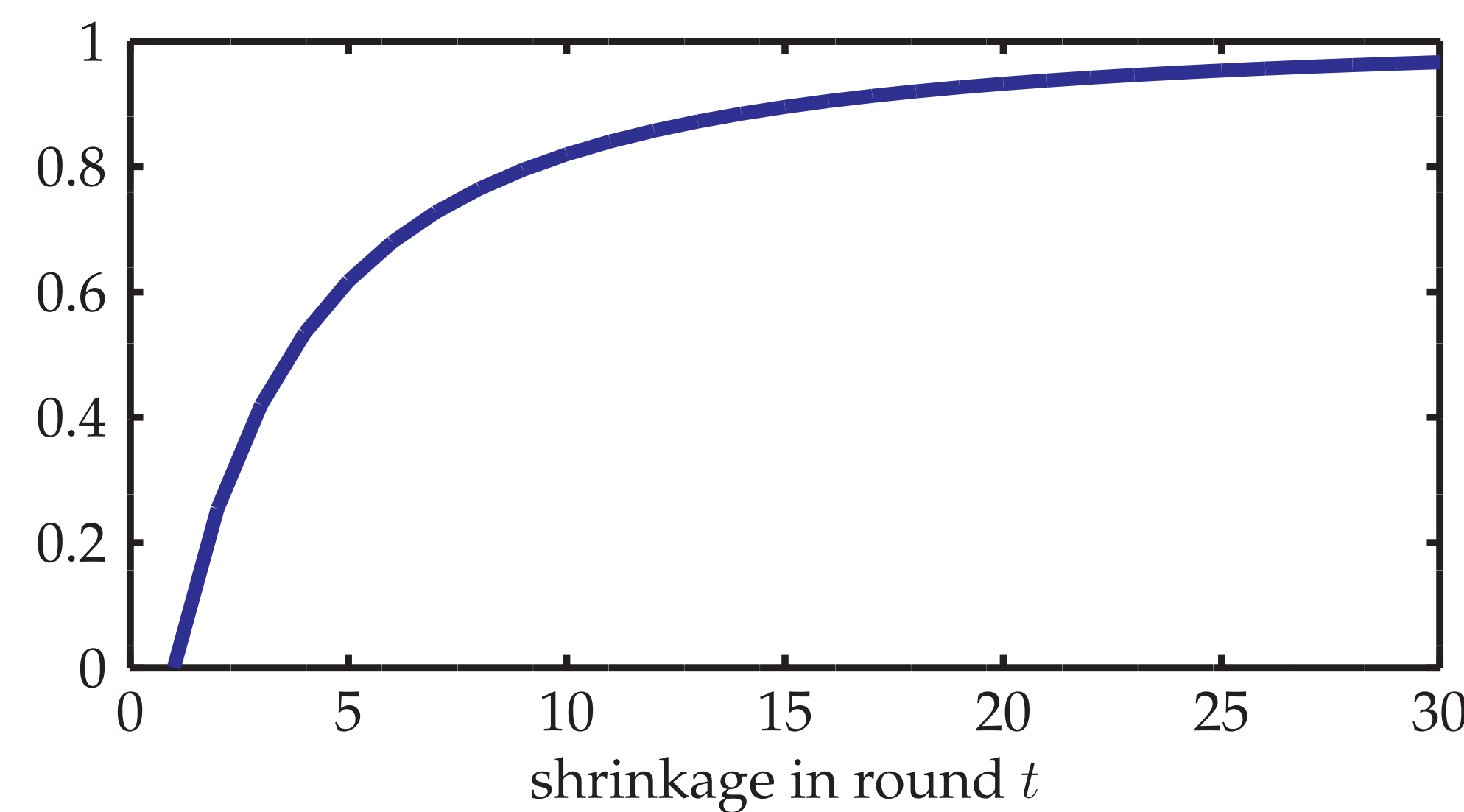
$$\alpha_t \mathbf{s}^\top \mathbf{W}^{-1} \mathbf{s} - \sigma^2 + (1 - t\alpha_t) \text{diag}(\mathbf{W}^{-1})^\top \mathbf{s} + \text{const}, \quad \leftarrow \text{QUADRATIC} \rightarrow \quad \mathbf{s}^\top \mathbf{A}_t \mathbf{s} - \sigma^2 + \text{const}.$$

and the minimax and maximin strategies for round $t + 1$ are given by

For round $t + 1$, the minimax strategy plays

$$\mathbf{a}^* = \mathbf{p}^* = \frac{\mathbf{s}}{t} t\alpha_{t+1} + \mathbf{c}(1 - t\alpha_{t+1}) \quad \leftarrow \text{LINEAR} \rightarrow \quad \mathbf{a}^* = (\lambda_{\max} \mathbf{I} - (\mathbf{A}_{t+1} - \mathbf{W}^{-1}))^{-1} \mathbf{A}_{t+1} \mathbf{s}$$

which is data mean $\frac{\mathbf{s}}{t}$ shrunk towards center \mathbf{c} .



and the maximin strategy plays two unit length vectors with

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

where λ_{\max} and \mathbf{v}_{\max} correspond to the largest eigenvalue of \mathbf{A}_{t+1} and \mathbf{a}_\perp and \mathbf{a}_\parallel are the components of \mathbf{a}^* perpendicular and parallel to \mathbf{v}_{\max} .

All coefficients (α_t, \mathbf{A}_t) are efficiently precomputable!

REGRET BOUNDS

- $\text{Regret}_{\text{Brier}} \propto \sum_{t=1}^T \alpha_t$.
- $\text{Regret}_{\text{Ball}} = \lambda_{\max}(\mathbf{W}^{-1}) \sum_{t=1}^T \alpha_t$.
- [1] show that $\sum_{t=1}^T \alpha_t = O(\log(T) - \log \log(T))$.
- Compare with $O(\log(T))$ of Follow the Leader.

[1] E. Takimoto, M. Warmuth The minimax strategy for Gaussian density estimation In COLT '00

COEFFICIENT RECURSIONS

Brier Game: $\alpha_T = \frac{1}{T}$ and

$$\alpha_t = \alpha_{t+1}^2 + \alpha_{t+1}.$$

Ball Game: $\mathbf{A}_T = \frac{1}{T} \mathbf{W}^{-1}$ and

$$\mathbf{A}_t = \mathbf{A}_{t+1} (\mathbf{W}^{-1} + \lambda_{\max} \mathbf{I} - \mathbf{A}_{t+1})^{-1} \mathbf{A}_{t+1} + \mathbf{A}_{t+1},$$

which maintains the eigenvectors of \mathbf{W} and updates the eigenvalues.

CONCLUSION

- Both games have tractable value functions and optimal strategies.
- All coefficients can be *precomputed* (need to know horizon T).
- Our strategies are subgame perfect: we are minimax optimal *from every history*.
- **FOLLOW-UP** We now know the minimax strategy for arbitrary outcome space \mathcal{C} .