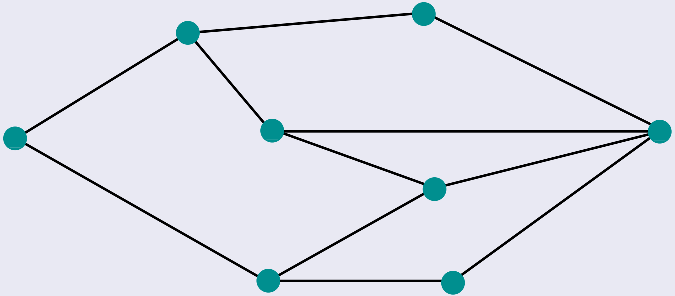
Combinatorial Prediction Games

Nicolò Cesa-Bianchi

Università degli Studi di Milano

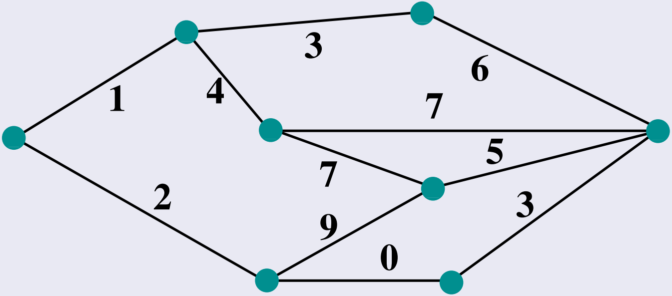
Contributors: Sébastien Bubeck, Sham Kakade, Gábor Lugosi

Example: Keeping a low-cost spanning tree



For every time step $t = 1, 2, \ldots$

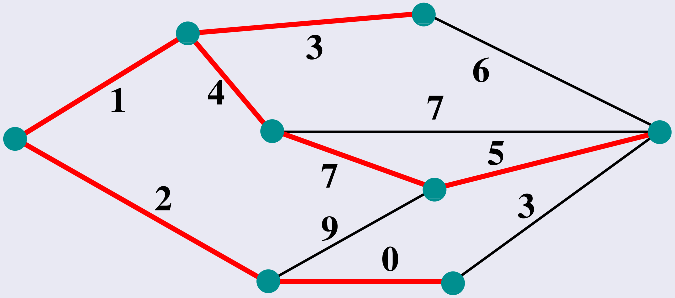
Example: Keeping a low-cost spanning tree



For every time step $t = 1, 2, \ldots$

1. Adversary chooses edge costs (hidden from the player)

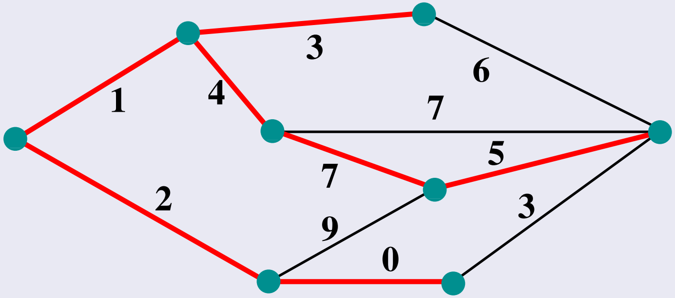
Example: Keeping a low-cost spanning tree



For every time step $t = 1, 2, \ldots$

1. Adversary chooses edge costs (hidden from the player)
2. Player chooses a tree

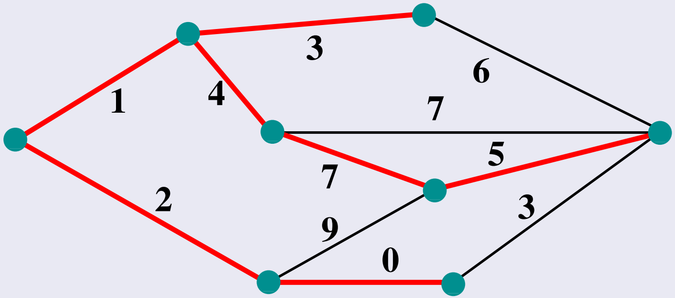
Example: Keeping a low-cost spanning tree



For every time step $t = 1, 2, \ldots$

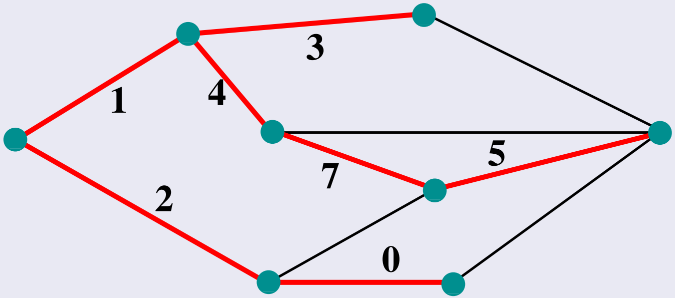
1. Adversary chooses edge costs (hidden from the player)
2. Player chooses a tree
3. Player's loss is total cost of tree edges $22 = 1 + 3 + 4 + 7 + 5 + 2 + 0$

Feedback models



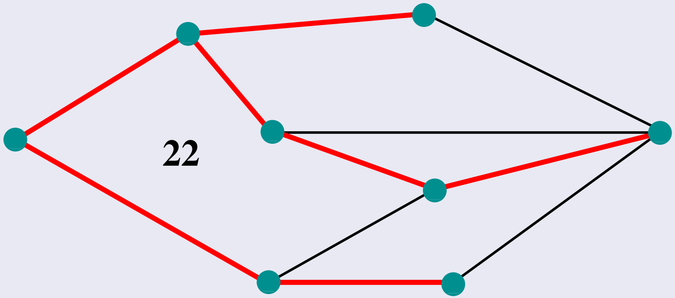
1. Full Information: player sees all edge costs

Feedback models



1. Full Information: player sees all edge costs
2. Semi-bandit: player sees cost of tree edges

Feedback models



1. Full Information: player sees all edge costs
2. Semi-bandit: player sees cost of tree edges
3. Bandit: player only sees total cost of tree

Combinatorial prediction games

For every time step $t = 1, 2, \ldots$

1. Adversary chooses loss vector $\ell_t \in \mathbb{R}^d$ (hidden from the player)
2. Player chooses an action $V_t \in \mathcal{S}$ and observes
   - loss vector ℓ_t (full info)
   - partial loss vector $\left(\ell_{i,t}\right)_{i \,:\, V_{i,t}=1}$ (semi-bandit)
   - loss $\ell_t^\top V_t$ (bandit)

- Actions are combinatorial objects represented by their incidence vectors $v \in \mathcal{S} \subseteq \{0,1\}^d$
- The loss of each action v at time t is linear over the incidence vector $\ell_t^\top v$ (the loss of a tree is the sum of its edge costs)
- $\left|\ell_t^\top v\right| \leqslant 1$ for all $v \in \mathcal{S}$ and $t = 1, 2, \ldots$

Regret

Adversary chooses $\ell_1, \ell_2, \dots \in \mathbb{R}^d$
Player chooses $\mathbf{V}_1, \mathbf{V}_2, \dots \in \mathcal{S}$

$$R_T = \sum_{t=1}^{T} \ell_t^\top \mathbf{V}_t - \min_{v \in \mathcal{S}} \sum_{t=1}^{T} \ell_t^\top v$$

Player's strategy

Template for randomized players

Player maintains $\boldsymbol{w}_t \in \mathbb{R}^d$

1. Obtain distribution p_t over $\mathcal{S}$ from $\boldsymbol{w}_t$ (decomposition)
2. Draw $\boldsymbol{V}_t$ from distribution p_t over $\mathcal{S}$ (sampling)
3. Estimate $\boldsymbol{\ell}_t$ using the unbiased estimate $\widehat{\boldsymbol{\ell}}_t$
4. Update $\boldsymbol{w}_t \to \boldsymbol{w}_{t+1}$ using $\widehat{\boldsymbol{\ell}}_t$

Unbiased loss estimates

- Full information: $\widehat{\boldsymbol{\ell}}_t = \boldsymbol{\ell}_t$
- Semi-bandit: $\widehat{\ell}_{i,t} = \dfrac{\ell_{i,t} V_{i,t}}{w_{i,t}}$
- Bandit: $\widehat{\boldsymbol{\ell}}_t = P_t^+ \boldsymbol{V}_t \boldsymbol{V}_t^\top \boldsymbol{\ell}_t$ where $P_t = \mathbb{E}_{p_t}\left[\boldsymbol{V}\boldsymbol{V}^\top\right]$

Algorithm ExpandedHedge

Representation of probabilities

- d weights over coordinates $\quad w_{j,t} = \exp\left(-\eta \widehat{L}_{j,t-1}\right)$
- $|\mathcal{S}|$ weights over actions $\quad w_t(\boldsymbol{v}) = \exp\left(-\eta \widehat{\mathbf{L}}_{t-1}^\top \boldsymbol{v}\right)$

Bandit case: GeometricHedge [Dani, Hayes, and Kakade, 2008]

- Distribution over actions:

$$p_t(\boldsymbol{v}) = (1-\gamma) \underbrace{\frac{w_t(\boldsymbol{v})}{Z_t}}_{\texttt{ExpandedHedge}} + \gamma\,\mu(\boldsymbol{v}) \qquad (0 \leqslant \gamma \leqslant 1)$$

- μ is an exploration distribution over the actions
- This controls the variance of the loss estimates by ensuring all coordinates are sampled often enough

Legendre potentials and Bregman divergences

Definition

A Legendre potential on a convex subset $\mathcal{D} \subseteq \mathbb{R}^d$ is any function $F : \mathcal{D} \to \mathbb{R}$ such that

1. F is strictly convex with continuous first partial derivatives on $\text{int}(\mathcal{D})$
2. $\nabla F(\boldsymbol{v})$ grows unboundedly as $\boldsymbol{v}$ approaches the boundary of $\mathcal{D}$

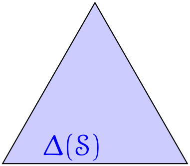
Definition

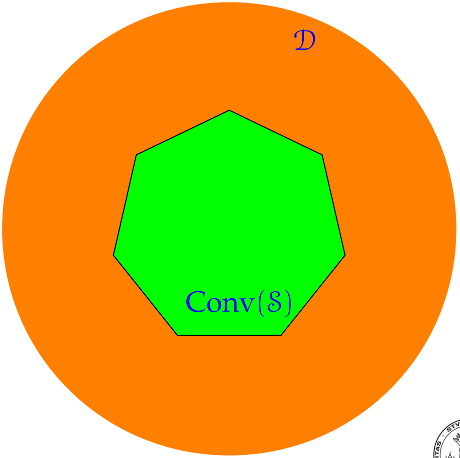
The Bregman divergence $D_F : \mathcal{D} \times \text{int}(\mathcal{D}) \to \mathbb{R}$ associated with a Legendre function F is

$$D_F(\boldsymbol{u}, \boldsymbol{v}) = F(\boldsymbol{u}) - F(\boldsymbol{v}) - (\boldsymbol{u} - \boldsymbol{v})^\top \nabla F(\boldsymbol{v})$$

Online Stochastic Mirror Descent (OSMD)

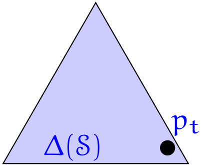
$\boldsymbol{w}_t \in \text{Conv}(\mathcal{S})$
F Legendre on $\mathcal{D} \supset \text{Conv}(\mathcal{S})$

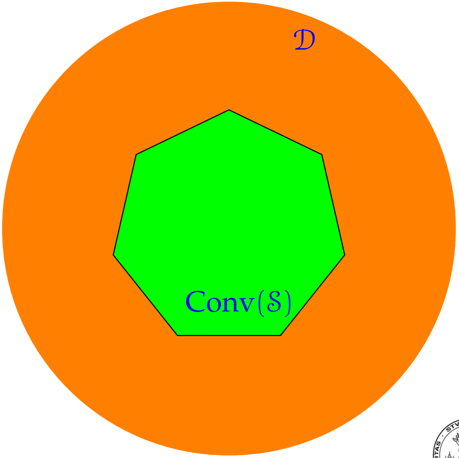




Online Stochastic Mirror Descent (OSMD)

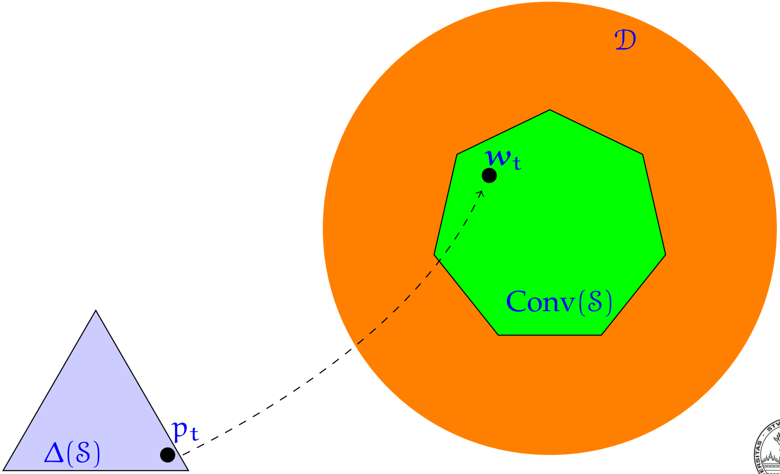
$\boldsymbol{w}_t \in \mathrm{Conv}(\mathcal{S})$
F Legendre on $\mathcal{D} \supset \mathrm{Conv}(\mathcal{S})$





Online Stochastic Mirror Descent (OSMD)

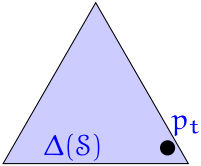
$\boldsymbol{w}_t \in \text{Conv}(\mathcal{S})$
F Legendre on $\mathcal{D} \supset \text{Conv}(\mathcal{S})$

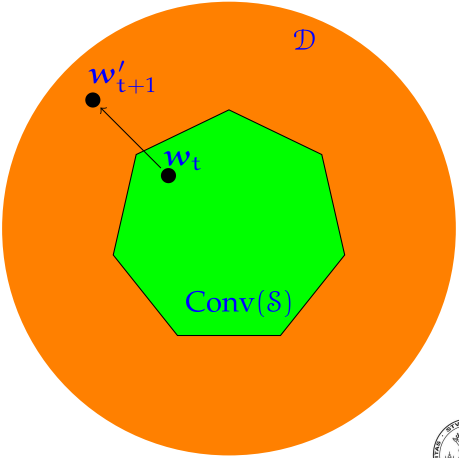


Online Stochastic Mirror Descent (OSMD)

$\boldsymbol{w}_t \in \mathrm{Conv}(\mathcal{S})$
F Legendre on $\mathcal{D} \supset \mathrm{Conv}(\mathcal{S})$

(1) $\nabla F(\boldsymbol{w}'_{t+1}) = \nabla F(\boldsymbol{w}_t) - \widehat{\boldsymbol{\ell}}_t$



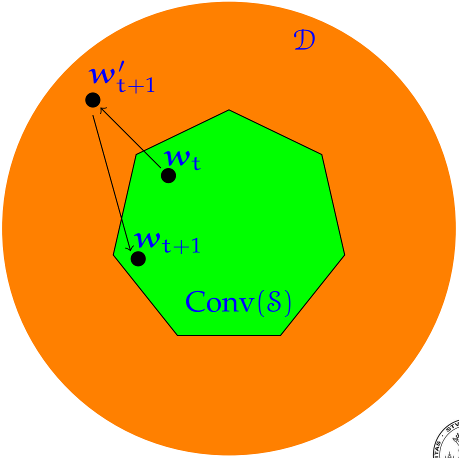


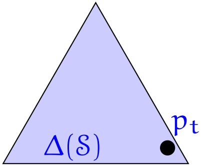
Online Stochastic Mirror Descent (OSMD)

$w_t \in \mathrm{Conv}(\mathcal{S})$
F Legendre on $\mathcal{D} \supset \mathrm{Conv}(\mathcal{S})$

(1) $\nabla F(w'_{t+1}) = \nabla F(w_t) - \widehat{\ell}_t$

(2) $w_{t+1} \in \underset{w \in \mathrm{Conv}(\mathcal{S})}{\mathrm{argmin}} \; D_F(w, w'_{t+1})$





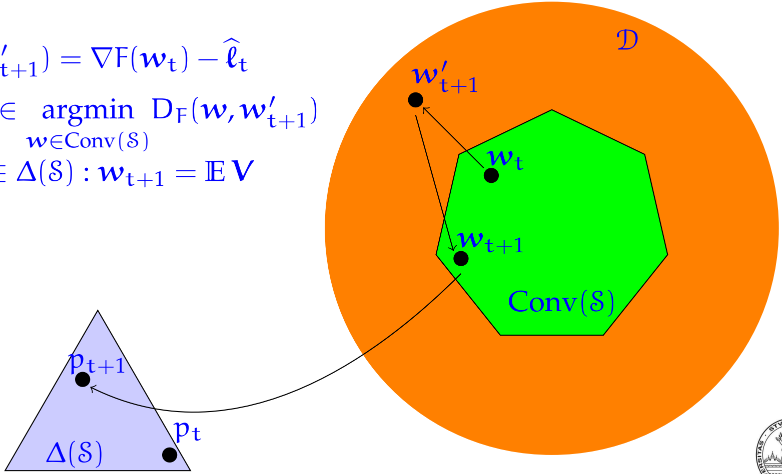
Online Stochastic Mirror Descent (OSMD)

$\boldsymbol{w}_t \in \text{Conv}(\mathcal{S})$
F Legendre on $\mathcal{D} \supset \text{Conv}(\mathcal{S})$

(1) $\nabla F(\boldsymbol{w}'_{t+1}) = \nabla F(\boldsymbol{w}_t) - \widehat{\boldsymbol{\ell}}_t$

(2) $\boldsymbol{w}_{t+1} \in \underset{\boldsymbol{w} \in \text{Conv}(\mathcal{S})}{\text{argmin}} \; D_F(\boldsymbol{w}, \boldsymbol{w}'_{t+1})$

(3) $p_{t+1} \in \Delta(\mathcal{S}) : \boldsymbol{w}_{t+1} = \mathbb{E}\,\boldsymbol{V}$



Efficient decomposition and sampling for OSMD

Main problems

1. Decomposition: Express $\boldsymbol{w}_t \in \mathrm{Conv}(\mathcal{S})$ as a distribution p_t over $\mathcal{S}$
2. Sampling: Draw an action from $\mathcal{S}$ according to p_t

- Carathéodory theorem ensures one can always find p_t with support of size at most $d + 1$
- This would also solve the sampling problem because the support is small enough
- Computing this set is generally intractable
- However, if $\mathrm{Conv}(\mathcal{S})$ has a polynomial number of faces, then p_t with small support can be found in polytime

Various instances

Classical case

- Full Info: Hedge [Freund, and Schapire, 1997]
- Semi-bandit/Bandit:
  - Exp3 [Auer, C-B, Freund, and Schapire, 2002]
  - INF [Audibert and Bubeck, 2009]

Combinatorial case

- Full Info: Component Hedge [Koolen, Warmuth, and Kivinen, 2010]
- Semi-bandit: MW [Kale, Reyzin, and Schapire, 2010]
- Bandit: Self-concordant F [Abernethy, Hazan, and Rakhlin, 2008]

Known regret bounds

$\lvert\mathcal{S}\rvert = 2^{\mathcal{O}(d)}$	Full	Semi-Bandit	Bandit
`ExpandedHedge`	$\sqrt{dT}$	$d\sqrt{T}$	$d^{3/2}\sqrt{T}$
`OMSD` (various F)	$\sqrt{dT}$	$\sqrt{dT}$	$d\sqrt{\theta T \ln T}$
Lower bound	$\sqrt{dT}$	$\sqrt{dT}$	$d\sqrt{T}$

- In general, $\theta = \mathcal{O}(d)$
- Hence `OMSD` and `ExpandedHedge` are comparable for bandits

Algorithm GeometricHedge [Dani, Hayes, and Kakade, 2008]

- Loss estimate $\widehat{\ell}_t = P_t^+ \mathbf{V}_t \mathbf{V}_t^\top \ell_t$ where $P_t = \mathbb{E}_{p_t}\left[\mathbf{V}\mathbf{V}^\top\right]$
- d weights over coordinates $w_{j,t} = \exp\left(-\eta \widehat{L}_{j,t-1}\right)$
- $|\mathcal{S}|$ weights over actions $w_t(\boldsymbol{v}) = \exp\left(-\eta \widehat{\mathbf{L}}_{t-1}^\top \boldsymbol{v}\right)$
- Distribution over actions:

$$p_t(\boldsymbol{v}) = (1-\gamma) \underbrace{\frac{w_t(\boldsymbol{v})}{Z_t}}_{\texttt{ExpandedHedge}} + \gamma\, \mu(\boldsymbol{v}) \qquad (0 \leqslant \gamma \leqslant 1)$$

Regret bound

$$\mathbb{E}\, R_T \leqslant d \sqrt{\left(\frac{B^2}{d\,\lambda_{\min}} + 1\right) T}$$

$\lambda_{\min}$ = smallest eigenvalue of $\mathbb{E}_\mu\left[\boldsymbol{V}\boldsymbol{V}^\top\right]$
$B = \max_{\boldsymbol{v}\in\mathcal{S}} \|\boldsymbol{v}\|$

- $B^2/\lambda_{\min}$ is proportional to the variance of loss estimates
- When $\lambda_{\min} \geqslant \dfrac{B^2}{d}$ we hit the optimal bound $d\sqrt{T}$
- If μ is uniform over all actions, the above happens when action space is approximately isotropic

Some cases is which uniform exploration is optimal

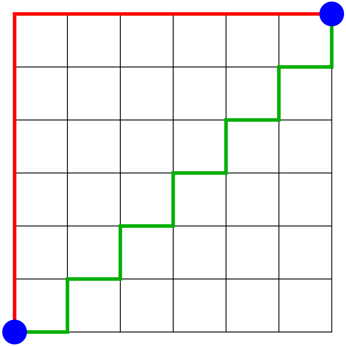
K bandits each with R actions	$\lambda_{\min} = \frac{1}{R}$
K-sized subsets of d elements	$\lambda_{\min} = \frac{K(d-K)}{d(d-1)}$
Permutations of K objects	$\lambda_{\min} = \frac{1}{K-1}$
Spanning trees of a K-clique	$\lambda_{\min} \geqslant \frac{1}{K} - \frac{17}{4K^2}$
Balanced cuts of a $2K$-clique	$\lambda_{\min} = \frac{1}{4} + \frac{8K-7}{4(2K-1)(2K-3)}$
Hamiltonian cycles in a K-clique	$\lambda_{\min} = \frac{2}{K-1}$

Sampling problem

K bandit problems	☺	independent draws from each bandit
K-sized subsets	☺	compute conditionals using dynamic programming
Permutations of K objects	☺	random sampling of perfect matchings
Spanning trees of a K-clique	☹	OK for uniform, unknown for weighted
Balanced cuts of a 2K-clique	☺	sampling ferromagnetic Ising model
Hamiltonian cycles in a K-clique	☹	notoriously hard

Routing: when uniform exploration fails

- Action set is a collection of paths between two vertices in a graph
- In this case $\lambda_{\min}$ with uniform exploration can get too small because there are "marginal" paths whose edges are seldom sampled



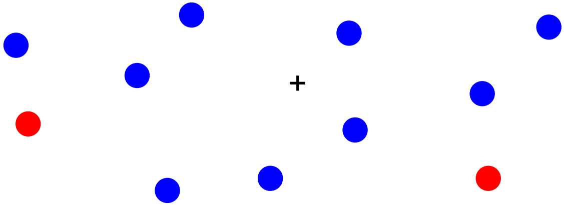
Change of representation

- Recall: actions are vectors $\boldsymbol{v} \in \{0,1\}^d$
- Problem when action set is not isotropic
  (marginal actions, $\lambda_{\min}$ small)

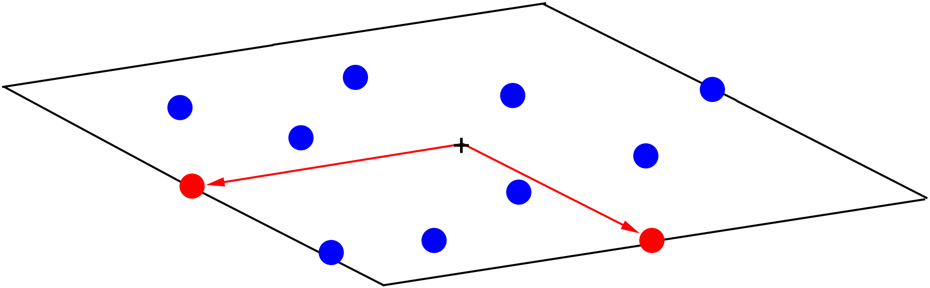
Basic steps

1. Choose a basis of $\mathbb{R}^d$ under which the action set looks isotropic
2. Run GeometricHedge with loss estimates computed in the new basis

Barycentric spanners [Dani, Hayes, and Kakade, 2008]



Barycentric spanners [Dani, Hayes, and Kakade, 2008]



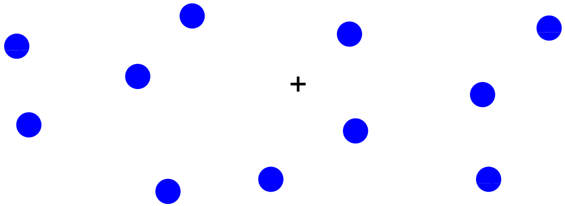
Remarks on barycentric spanners

- Exploration distribution μ is uniform over the d spanners
- All actions $\boldsymbol{v} \in \mathcal{S}$ have $[-1,+1]$ coordinates, $B^2 = d$
- $\mathbb{E}_\mu\left[\boldsymbol{V}\boldsymbol{V}^\top\right]$ is isotropic ($\lambda_{\min} = \frac{1}{d}$) in the spanner basis
- Computation of barycentric spanners efficient in some cases

Regret bound

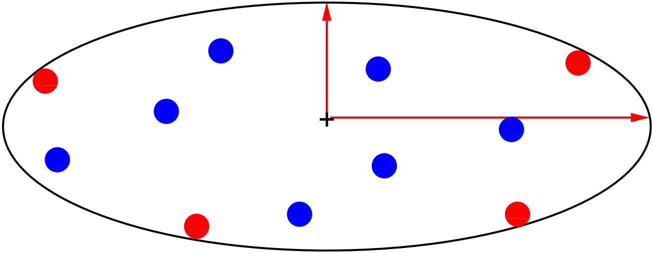
$$\mathbb{E}\, R_T \leqslant d\sqrt{\left(\frac{B^2}{d\,\lambda_{\min}} + 1\right) T} \simeq d^{3/2}\sqrt{T}$$

Löwner ellipsoid [C-B and Kakade, 2011]



Löwner ellipsoid

[C-B and Kakade, 2011]



Remarks on Löwner ellipsoid

- Exploration distribution μ over the $\mathcal{O}(d^2)$ contact points
- John's theorem ensures that $\mathbb{E}_\mu\left[\mathbf{V}\mathbf{V}^\top\right]$ is isotropic ($\lambda_{\min} = \frac{1}{d}$) in the basis where the Löwner ellipsoid is rescaled to the unit ball
- Since $\mathcal{S}$ is in the unit ball, $B^2 = 1$
- Computation of Löwner ellipsoid efficient in some cases

Regret bound

$$\mathbb{E}\, R_T \leqslant d \sqrt{\left(\frac{B^2}{d\,\lambda_{\min}} + 1\right) T} \simeq d\sqrt{T}$$

Regret bounds

$\lvert\mathcal{S}\rvert = 2^{\mathcal{O}(d)}$	Full	Semi-Bandit	Bandit
`ExpandedHedge`	$\sqrt{dT}$	$d\sqrt{T}$	$d\sqrt{T}$
`OMSD` (various F)	$\sqrt{dT}$	$\sqrt{dT}$	$d\sqrt{\theta T \ln T}$
Lower bound	$\sqrt{dT}$	$\sqrt{dT}$	$d\sqrt{T}$

Conclusions

- Combinatorial sequential prediction problems with linear loss in three feedback models: full info, semi-bandit, bandit
- Two main algorithms: `ExpandedHedge` and `OSMD`
- Key problem with bandit feedback is estimating the loss in each coordinate while keeping variance under control
- `ExpandedHedge` achieves optimal bandit rate by de-skewing the action set using the basis provided by Löwner ellipsoid
- Open: show how to design potential functions based on the topology of the action set and prove optimal regret bounds for `OSMD`