Approachability, Fast and Slow

Shie Mannor & Vianney Perchet

Electrical Engineering
Technion

L. Probabilités Modèles Aléatoires
Université Paris – Diderot

Conference On Learning Theory '13
Online learning (I)

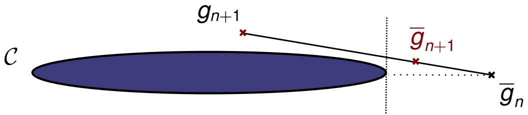
COLT 2013

Motivations

- **Approachability. Motivations**
  - generalizes regret to vectorial (multi-criteria) losses $g_n \in \mathbb{R}^d$
  - **Generic tool:** construct online learning & game theory algo.
- **Sequential decision pb against Nature (adversarial)**
  - Goal: $\overline{g}_n = \sum_{m=1}^{n} g_m/n$ converges to convex target set $\mathcal{C} \subset \mathbb{R}^d$
  - If possible (against any strategy of Nature), $\mathcal{C}$ is **approachable**

Motivations

- **Approachability. Motivations**
  - generalizes regret to vectorial (multi-criteria) losses $g_n \in \mathbb{R}^d$
  - **Generic tool:** construct online learning & game theory algo.
- **Sequential decision pb against Nature (adversarial)**
  - Goal: $\overline{g}_n = \sum_{m=1}^{n} g_m/n$ converges to convex target set $\mathcal{C} \subset \mathbb{R}^d$
  - If possible (against any strategy of Nature), $\mathcal{C}$ is **approachable**
- **At which rate ? :**
  - Regret: **either** $\overline{R}_n = 0$ **or** $\Omega(1/\sqrt{n})$ (... $1/n^{1/3}$ w. partial monit.)
  - If $\mathcal{C}$ is approach: $d_{\mathcal{C}}(\overline{g}_n) \leq O(1\sqrt{n})$ and



What about **possible rates** of approachability ? **only** 0 and $1/\sqrt{n}$??

Counter Examples - Fast Approachability

- Regret minimization: instance of approachability

First possible rate $\Omega(1/\sqrt{n})$ ... **slow approachability**

- Easy calibration: Each day, meteorologist predicts rain ($q_n = 1$) the next day. Accurate predictions if $\|\overline{p}_n - \overline{q}_n\| \to 0$

Predict the weather of yesterday... **fast rate** $\|\overline{p}_n - \overline{q}_n\| \leq 1/n$

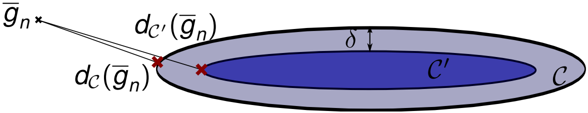
- **Convergence of empirical distribution:** Given $p \in \Delta(\{1, ..., d\})$, choose i_1, ..., i_n so that $\bar{\imath}_n = \sum_{m=1}^{n} \delta_{i_m}/n \to p$.

Bad idea: $i_m \sim p$ i.i.d. as $\mathbb{E}\|\bar{\imath}_n - p\| \simeq \sqrt{d/n}$

Good idea: **Fast approachability!** $\|\bar{\imath}_n - p\| \leq d/n$

Insightful Easy Cases

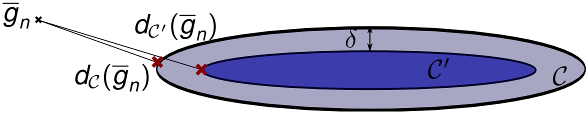
- **Approachable δ-shrinkage**



$$d_{\mathcal{C}}(\overline{g}_n) \leq \frac{d^2_{\mathcal{C}'}(\overline{g}_n)}{4\delta} \leq O\left(\frac{1}{\delta}\left(\frac{1}{\sqrt{n}}\right)^2\right) = O\left(\frac{1}{\delta n}\right)$$

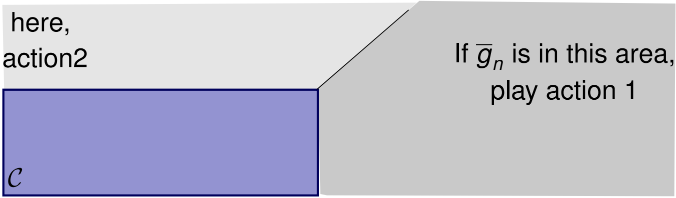
Insightful Easy Cases

- **Approachable δ-shrinkage**



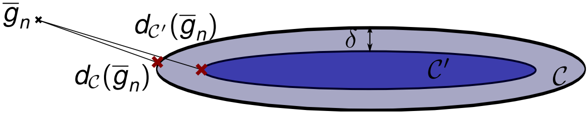
$$d_{\mathcal{C}}(\overline{g}_n) \leq \frac{d^2_{\mathcal{C}'}(\overline{g}_n)}{4\delta} \leq O\left(\frac{1}{\delta}\left(\frac{1}{\sqrt{n}}\right)^2\right) = O\left(\frac{1}{\delta n}\right)$$

- **Deterministically approachable polytope**



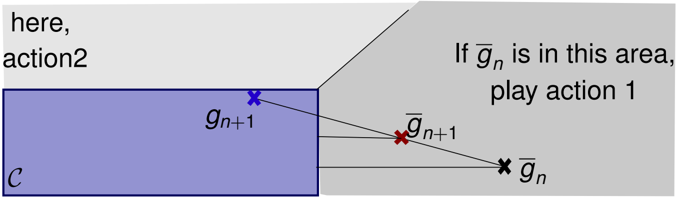
Insightful Easy Cases

- **Approachable δ-shrinkage**



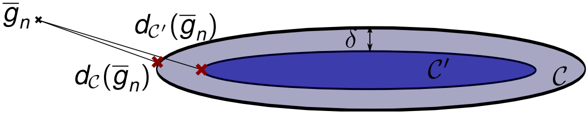
$$d_{\mathcal{C}}(\overline{g}_n) \leq \frac{d^2_{\mathcal{C}'}(\overline{g}_n)}{4\delta} \leq O\left(\frac{1}{\delta}\left(\frac{1}{\sqrt{n}}\right)^2\right) = O\left(\frac{1}{\delta n}\right)$$

- **Deterministically approachable polytope**



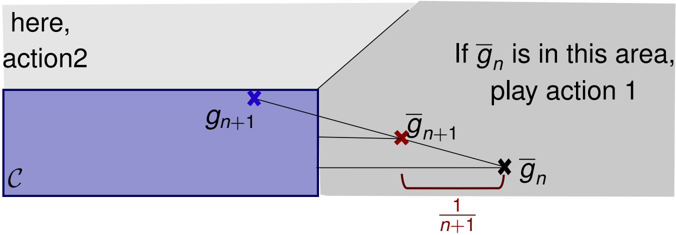
Insightful Easy Cases

- **Approachable δ-shrinkage**



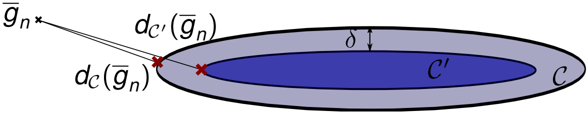
$$d_{\mathcal{C}}(\overline{g}_n) \leq \frac{d^2_{\mathcal{C}'}(\overline{g}_n)}{4\delta} \leq O\left(\frac{1}{\delta}\left(\frac{1}{\sqrt{n}}\right)^2\right) = O\left(\frac{1}{\delta n}\right)$$

- **Deterministically approachable polytope**



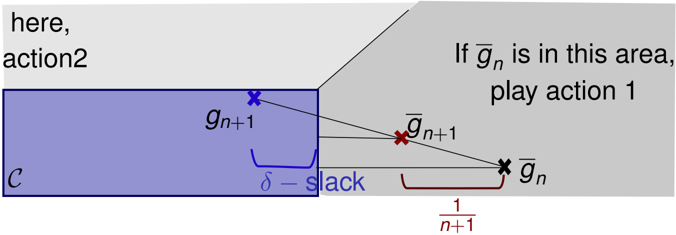
Insightful Easy Cases

- **Approachable δ-shrinkage**



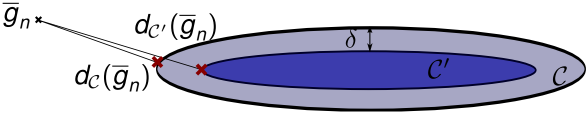
$$d_{\mathcal{C}}(\overline{g}_n) \leq \frac{d^2_{\mathcal{C}'}(\overline{g}_n)}{4\delta} \leq O\left(\frac{1}{\delta}\left(\frac{1}{\sqrt{n}}\right)^2\right) = O\left(\frac{1}{\delta n}\right)$$

- **Deterministically approachable polytope**



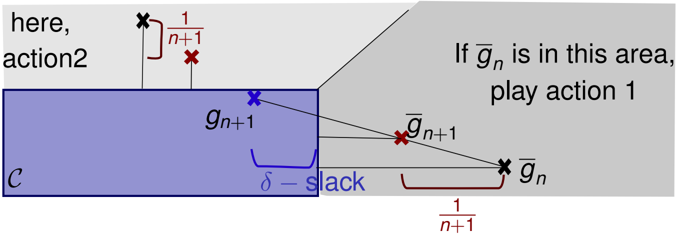
Insightful Easy Cases

- **Approachable δ-shrinkage**



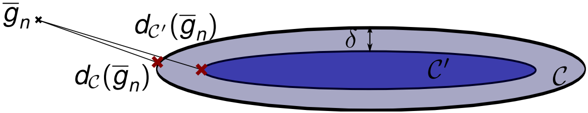
$$d_{\mathcal{C}}(\overline{g}_n) \leq \frac{d^2_{\mathcal{C}'}(\overline{g}_n)}{4\delta} \leq O\left(\frac{1}{\delta}\left(\frac{1}{\sqrt{n}}\right)^2\right) = O\left(\frac{1}{\delta n}\right)$$

- **Deterministically approachable polytope**



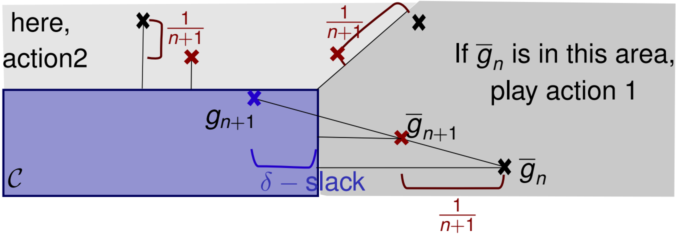
Insightful Easy Cases

- **Approachable δ-shrinkage**



$$d_{\mathcal{C}}(\overline{g}_n) \leq \frac{d^2_{\mathcal{C}'}(\overline{g}_n)}{4\delta} \leq O\left(\frac{1}{\delta}\left(\frac{1}{\sqrt{n}}\right)^2\right) = O\left(\frac{1}{\delta n}\right)$$

- **Deterministically approachable polytope**



Main Results

A closed and convex set $\mathcal{C} \subset \mathbb{R}^d$ is fast approachable

- Either if it is **deterministically approachable**
- Or if there exists $\delta > 0$ such that, whenever a **random action** is required, there is a **δ-slack**

$\exists$ strategy of DM such that $\forall$ strategy of Nature, $\mathbb{E}[d_{\mathcal{C}}(\overline{g}_n)] \leq O(1/n)$.

Converse statement

If approachability requires a random action without slack then

(under additional geometric condition) $\mathcal{C}$ is **slow-approachable**

$\exists$ strategy of Nature such that $\forall$ strategy of DM, $\mathbb{E}[d_{\mathcal{C}}(\overline{g}_n)] \geq \Omega(1/\sqrt{n})$.