EECS598: Prediction and Learning: It's Only a Game Fall 2013

Lecture 22: Blackwell's Approachability Theorem

Prof. Jacob Abernethy *Scribe: Petter Nilsson*

Announcements

- Prof. Jacob Abernethy will be away Tuesday-Friday next week. It is possible to schedule a meeting this week to discuss homework and/or project.
- Start to think about presenting projects in class. Participation credit will be given.
- Reminder: No class on November 27th.

22.1 Blackwell's Approachability Theorem

When the payoff function is multivariate, the minmax theory needs to be modified. Optimizing a vector is not well defined, nor is the notion of equilibria in a vector-valued game.

Given is a payoff function $r : \Delta_n \times \Delta_m \to \mathbb{R}^d$ which is biaffine. The bi-affinity is required to preserve the expectation of mixed strategies, i.e. that

$$r(p,q) = \mathbb{E}_{i\sim p,\, j\sim q} r(i,j). \tag{22.1}$$

Rather than maximizing, the goal in this setting is to direct the payoff vector to some convex set $S \subset \mathbb{R}^d$. We are interested in some duality result like the following

$$\forall q\, \exists p \;\text{ s.t }\; r(p,q) \in S \implies \exists p \;\text{ s.t }\; \forall q \;\; r(p,q) \in S, \tag{22.2}$$

which is a direct translation of the classical minimax result for scalar-valued payoff functions. This statement is however **not true** in the multivariate setting, as the following counter-example shows:

Counter-example to (22.2)**:** If $S = \{(p,q) : p = q\}$ and $r(p,q) = (p,q)$, one can trivially, for all q, choose $p(q) = q$ to guarantee that $r(p(q),q) \in S$. It is however not possible to find a p which works for all q, indeed the only p which works for a given q is $p = q$.

However, the duality statement (22.2) holds when S is a half-space $\{x \mid v \cdot x \geq c\}$. To see this, define a zero-sum game with scalar payoff given by the projection onto the normal direction v of the half-space:

$$M_{i,j} = v \cdot r(e_i, e_j). \tag{22.3}$$

Then the condition $r(p,q) \in S$ is equivalent to $p^T M q \geq c$ (because of bi-affinity), so the von Neumann minimax theorem applies. More succinctly, the following two conditions are equivalent (here S is a general convex set).

$$\forall q\, \exists p \;\text{ s.t }\; r(p,q) \in S, \tag{22.4}$$

$$\text{For all half-spaces } H \text{ containing } S, \quad \exists p \;\text{ s.t }\; \forall q, \;\; r(p,q) \in H. \tag{22.5}$$

Proof. See previous lecture. □

Even though (22.2) is false in the general case, the half-space condition (22.5) can be used to prove *Blackwell's Approachability Theorem* (BAT).

Theorem 22.1. *If* (22.4) *or* (22.5) *holds, then there exists an adaptive strategy*

$$p_t \leftarrow f_t(q_1, \dots, q_{t-1}) \tag{22.6}$$

such that for all $q_1, q_2, \dots$

$$d\left(\frac{1}{T}\sum_{t=1}^{T} r(p_t, q_t), S\right) \to 0 \tag{22.7}$$

as $T \to \infty$.

Here $d : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is the usual euclidean metric on $\mathbb{R}^d$, extended to $\mathbb{R}^d \times 2^{\mathbb{R}^d}$* by

$$d(x, S) = \inf_{s \in S} d(x, s) \tag{22.8}$$

for sets $S \subset \mathbb{R}^d$.

Proof. The proof is constructive, i.e. it explicitly gives an adaptive strategy so that (22.7) holds. This strategy is given below. See Figure 1 for an illustration.

Algorithm 1: Adaptive Strategy

for *t=1,2,...* **do**
  Compute

$$\bar{r}_t \leftarrow \frac{1}{t}\sum_{i=1}^{t} r(p_i, q_i) \tag{22.9}$$

  if $\bar{r}_t \in S$ **then**
    Choose p_{t+1} as anything (it doesn't matter).
  else
    Let $\pi_S(\bar{r}_t)$ be the projection of $\bar{r}_t$ onto S, i.e. $\pi_S(x) = \arg\inf_{y \in S} d(x, y)$. See Figure 1.
    Select the half space H_{t+1} containing S such that ∂H_{t+1} (∂ is the boundary operator) contains $\pi_S(\bar{r}_t)$ and has normal direction $\bar{r}_t - \pi_S(\bar{r}_t)$.
    Choose p_{t+1} such that $r(p_{t+1}, q) \in H_{t+1}$ for all q. This is possible because of the half-space condition (22.5).
  end
  Observe q_{t+1}.
end

We now analyze this strategy. The average payoff will update as

$$\bar{r}_{t+1} = \frac{t}{t+1}\bar{r}_t + \frac{1}{t+1} r(p_{t+1}, q_{t+1}). \tag{22.10}$$

*$2^{\mathbb{R}^d}$ is the power set of $\mathbb{R}^d$, the set of all subsets defined as $2^{\mathbb{R}^d} = \{D : D \subset \mathbb{R}^d\}$.

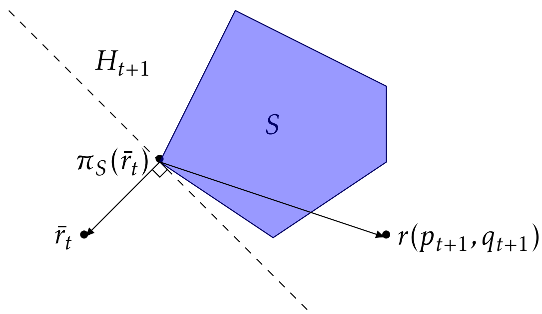

Figure 1: Blackwell Approachability algorithm.

We want to bound the distance of $\bar{r}_T$ to the set S.

$$\begin{aligned} d^2(\bar{r}_{t+1}, S) \leq d^2(\bar{r}_{t+1}, \pi_S(\bar{r}_t)) &= \left\| \frac{t}{t+1}\bar{r}_t + \frac{1}{t+1} r(p_{t+1}, q_{t+1}) - \pi_S(\bar{r}_t) \right\|_2^2 \\ &= \left\| \frac{t}{t+1}(\bar{r}_t - \pi_S(\bar{r}_t)) + \frac{1}{t+1}(r(p_{t+1}, q_{t+1}) - \pi_S(\bar{r}_t)) \right\|_2^2 \\ &= \left(\frac{t}{t+1}\right)^2 \|\bar{r}_t - \pi_S(\bar{r}_t)\|_2^2 + \left(\frac{1}{t+1}\right)^2 \|r(p_{t+1}, q_{t+1}) - \pi_S(\bar{r}_t)\|_2^2 \\ &+ \frac{2t}{(t+1)^2}(\bar{r}_t - \pi_S(\bar{r}_t)) \cdot (r(p_{t+1}, q_{t+1}) - \pi_S(\bar{r}_t)). \end{aligned} \tag{22.11}$$

Since $\bar{r}_t$ and $r(p_{t+1}, q_{t+1})$ are on different sides of the hyperplane ∂H_{t+1} (see Figure 1), which by construction contains $\pi_S(\bar{r}_t)$, the dot product in the last line is negative. Recognizing that $d(\bar{r}_t, S) = \|r(p_{t+1}, q_{t+1}) - \pi_S(\bar{r}_t)\|_2$, it follows that

$$(t+1)^2 d^2(\bar{r}_{t+1}, S) \leq t^2 d^2(\bar{r}_t, S) + \|r(p_{t+1}, q_{t+1}) - \pi_S(\bar{r}_t)\|_2^2. \tag{22.12}$$

By scaling, we can WLOG assume that S and the domain of r are inside the unit ball. Then the distance $\|r(p_{t+1}, q_{t+1}) - \pi_S(\bar{r}_t)\|_2^2$ is $\mathcal{O}(1)$.

Then, by summing (22.12) telescopically, we get

$$T^2 d^2(\bar{r}_T, S) \leq d^2(\bar{r}_1, S) + \mathcal{O}(T), \tag{22.13}$$

so, finally,

$$d(\bar{r}_T, S) \leq \mathcal{O}(\sqrt{\frac{1}{T}}). \tag{22.14}$$

This shows that the average payoff converges to the set S as $T \to \infty$. □

22.2 Solving Online Learning using Approachability

Consider an online learning setting where loss vectors $\ell^1, \ell^2, \ldots \in [0,1]^d$ are observed. We want to choose weights $w^1, w^2, \ldots \in \Delta_d$ so that

$$\forall \epsilon > 0, \quad \exists T \text{ s.t. } \frac{1}{T}\left(\sum_{t=1}^{T} \ell^t \cdot w^t - \sum_{t=1}^{T} \ell_i^t\right) \leq \epsilon \quad \forall i. \tag{22.15}$$

Note that this is the normal condition of sublinear regret when competing against the best fixed expert. The problem of choosing w^t can be reduced to an approachability problem as follows:

Reduction: The components of an approachability problem are the payoff function $r : \Delta_d \times [0,1]^d \to \mathbb{R}^d$ and a convex set $S \subset \mathbb{R}^d$, define these in terms of ℓ and w as follows:

$$r(w,l) = \langle w \cdot \ell - \ell_1, w \cdot \ell - \ell_2, \ldots, w \cdot \ell - \ell_d \rangle, \tag{22.16}$$

$$S = \mathbb{R}^d_- := \{v \in \mathbb{R}^d : v_i \leq 0 \ \forall i\}. \tag{22.17}$$

From this definition the average payoff in the approachability problem becomes

$$\frac{1}{T}\sum_{t=1}^{T} r(\ell^t, w^t) = \langle \frac{1}{T}\sum_{t=1}^{T} \ell^t \cdot w^t - \sum_{t=1}^{T} \ell_1^t, \ldots, \frac{1}{T}\sum_{t=1}^{T} \ell^t \cdot w^t - \sum_{t=1}^{T} \ell_d^t \rangle, \tag{22.18}$$

so if we can make the average payoff in the approachability problem approach the set S, we get low regret in the online learning problem. To apply Blackwell's Approachability Theorem, we verify that (22.4) holds, which is a more useful condition than (22.5) in practice.

Claim 22.2. *For all $\ell \in [0,1]^d$, there is a w such that $r(w,l) \in \mathbb{R}^d_-$.*

Proof. By choosing

$$w(\ell) = e_{i^*}, \text{ where } i_* = \arg\min_j \ell_j, \tag{22.19}$$

we get that

$$r(e_{i^*}, \ell) = \langle \ell_{i^*} - \ell_1, \ldots, \ell_{i^*} - \ell_d \rangle, \tag{22.20}$$

which is in $\mathbb{R}^d_-$ by definition. □

Remark 22.3. *The intuition behind this result is that if we observe the loss vector in advance, of course we can beat the best expert. We just choose the minimizer ourselves!*

Thus condition (22.4) holds, and we can apply Blackwell's Approachability Theorem. To conclude, the online learning problem of minimizing the regret (22.15) can be solved by applying Algorithm 1 on the approachability problem defined by (22.16) and (22.17).

22.3 Solving Approachability using Online Convex Optimization

NOTE: The following material was drawn from recent work of Bernstein and Shimkin (`http://arxiv.org/abs/1312.7658`) which gave a simplification of previous work by Abernethy, Bartlett, and Hazan (`http://jmlr.org/proceedings/papers/v19/abernethy11b/abernethy11b.pdf`).

We just saw that solving a learning problem in the simplex is can be reduced to solving an approachability problem in the dual cone to the simplex, which is $\mathbb{R}^d_-$. It turns out that a connection exists also in the other direction; online convex optimization can be used to solve an approachability problem. This provides an alternative proof for Blackwell's Approachability Theorem.

Let a bi-affine $r : X \times Y \to \mathbb{R}^d$ and a convex set $S \in \mathbb{R}^d$ define an approachability problem. Assume that r and S fulfill the half-space condition (22.5). Define the *support function* h_S of S as

$$h_S(\theta) = \max_{s \in S} \theta \cdot s. \tag{22.21}$$

Since h_S is the maximum over a set of affine functions, it is convex.

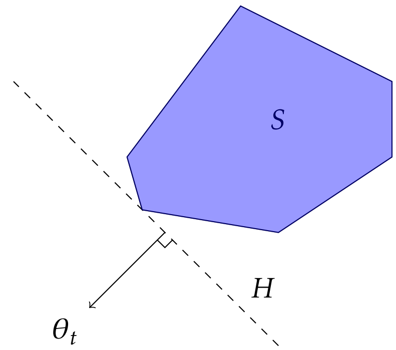

Figure 2: Choice of half space H according to a direction θ_t.

Claim 22.4.

$$d(x,S) = \max_{\|\theta\|_2 \le 1} \{\theta \cdot s - h_S(\theta)\}. \tag{22.22}$$

Proof.

$$\max_{\|\theta\|_2 \le 1} \{\theta \cdot x - h_S(\theta)\} = \max_{\|\theta\|_2 \le 1} \{\theta \cdot x - \max_{y \in S} \theta \cdot y\} = \max_{\|\theta\|_2 \le 1} \min_{y \in S} \theta \cdot (x - y). \tag{22.23}$$

Also note that for the projection of x on S,

$$\pi_S(x) = \arg\min_{y \in S} \|x - y\|_2 = \arg\min_{y \in S} \frac{(x-y)}{\|x-y\|_2} \cdot (x - y) \ge \arg\min_{y \in S} \theta \cdot (x - y), \tag{22.24}$$

for all θ such that $\|\theta\|_2 \le 1$. Since $\theta = (x-y)/\|x-y\|_2$ gives the projection,

$$\pi_S(x) = \max_{\|\theta\|_2 \le 1} \arg\min_{y \in S} \theta \cdot \|x - y\|, \tag{22.25}$$

so,

$$\max_{\|\theta\|_2 \le 1} \min_{y \in S} \theta \cdot (x - y) = \|x - \pi_S(x)\| = d(x,S). \tag{22.26}$$

□

Now, suppose we are receiving a sequence of q_t's. We want to choose p_t such that the average payoff $\frac{1}{T}\sum_{t=1}^{T} r(p_t, q_t)$ approaches S. To this end, we choose p_t so that

$$\theta_t \cdot r(p_t, q) \le \max_{s \in S} \theta_t \cdot s = h_s(\theta_t), \quad \forall q, \tag{22.27}$$

where the θ_t's are to be defined later. This choice is possible because we assume that the half-space condition (22.4) holds. Indeed, if H is the half-space tangent to S and with normal θ pointing *away* from S, then $\max_{s \in S} \theta_t \cdot s$ is attained in the tangent point and all points h of H are 'below' the tangent point in the θ-direction, as shown in Figure 2.

We now define an objective function

$$f_t(\theta) = h_S(\theta) - \theta \cdot r(p_t, q_t) \tag{22.28}$$

and choose θ_{t+1} by doing an online convex optimization update where θ_{t+1} is constrained to the euclidean unit ball. We then know that for all $\epsilon > 0$ there exists a T such that the following regret bound holds

$$\frac{1}{T}\left(\sum_{t=1}^{T} f_t(\theta_t) - \min_{\|\theta\|_2 \le 1} \sum_{t=1}^{T} f_t(\theta)\right) \le \epsilon. \tag{22.29}$$

But by the definition of f, the regret expression reads

$$\begin{aligned}
&\frac{1}{T}\left(\sum_{t=1}^{T} f_t(\theta_t) - \min_{\|\theta\|_2 \le 1} \sum_{t=1}^{T} f_t(\theta)\right) \\
&= \frac{1}{T}\left(\sum_{t=1}^{T} h_s(\theta_t) - \theta_t \cdot r(p_t, q_t)\right) - \min_{\|\theta\|_2 \le 1}\left\{h_S(\theta) - \theta \cdot \frac{1}{T}\sum_{t=1}^{T} r(p_t, q_t)\right\} \\
&= \frac{1}{T}\left(\sum_{t=1}^{T} \underbrace{h_s(\theta_t) - \theta_t \cdot r(p_t, q_t)}_{\ge 0 \text{ by } (22.27)}\right) + \underbrace{\max_{\|\theta\|_2 \le 1}\left\{\theta \cdot \frac{1}{T}\sum_{t=1}^{T} r(p_t, q_t) - h_S(\theta)\right\}}_{= d(\frac{1}{T}\sum_{t=1}^{T} r(p_t, q_t), S) \text{ by } (22.22)}.
\end{aligned} \tag{22.30}$$

It follows that

$$d\left(\frac{1}{T}\sum_{t=1}^{T} r(p_t, q_t), S\right) \le \epsilon, \tag{22.31}$$

and since ϵ can be arbitrarily small,

$$d\left(\frac{1}{T}\sum_{t=1}^{T} r(p_t, q_t), S\right) \to 0 \tag{22.32}$$

as $T \to \infty$. Thus we can make the average payoff approach S by using online convex optimization.