In this lecture we formulate and prove the celebrated *approachability theorem* of Blackwell, which extends von Neumann's minimax theorem to zero-sum games with vector-valued payoffs [1]. (The proof here is based on the presentation in [2]; a similar presentation was given by Foster and Vohra [3].) This theorem is powerful in its own right, but also has significant implications for regret minimization; as we will see in the next lecture, the algorithmic insight behind Blackwell's theorem can be used to easily develop both external and internal regret minimizing algorithms.

1 Zero-Sum Games with Vector-Valued Payoffs

We first define two-player zero-sum games with vector-valued payoffs. Each player i has an action space A_i (assumed to be finite). In a vector-valued game, the payoff to player 1 when the action pair (a_1, a_2) is played is $\mathbf{\Pi}(a_1, a_2) \in \mathbb{R}^K$, for some finite K; that is, the payoff to player 1 is a vector. Similarly, the payoff to player 2 is $-\mathbf{\Pi}(a_1, a_2)$. We use similar notation as earlier lectures: i.e., we let $\mathbf{\Pi}(s_1, s_2)$ denote the expected payoff to player 1 when each player i uses mixed action $s_i \in \Delta(A_i)$. We will typically view s_i as a vector in $\mathbb{R}^{A_i}$, with $s_i(a_i)$ equal to the probability that player i places on a_i.

The game is played repeatedly by the players. We use s_i^t to denote the mixed action chosen by player i at time t, and we let a_i^t denote the actual action played by player i at time t. We let $h^T = (\boldsymbol{a}^0, \dots, \boldsymbol{a}^{T-1})$ denote the history of the actual play up to time T.

We assume that the payoffs all lie in the unit ball (with respect to the standard Euclidean norm): $\|\Pi(a_1, a_2)\| \leq 1$ for all a_1, a_2. Since action spaces are finite, this just amounts to a rescaling of payoffs for analytical simplicity.

2 Approachability

We first develop approachability in the scalar payoff setting. We then generalize to halfspaces in the vector-valued payoff setting, and finally state Blackwell's theorem for approachability of general convex sets.

2.1 The Scalar Case

We first develop the notion of approachability in the one-dimensional (i.e., scalar payoff) setting, where $K = 1$. In this case players 1 and 2 play a standard zero-sum game, and it is well known (from von Neumann's minimax theorem) that there exists a mixed action s_1 of player 1 such that for any pure action a_2 of player 2, there holds $\Pi(s_1, a_2) \geq \mathrm{val}(\Pi)$ (where val denotes the value of the zero-sum game); and similarly, there exists s_2 such that for all a_1, there holds $\Pi(a_1, s_2) \leq \mathrm{val}(\Pi)$.

We now consider the implication of these observations for repeated play. Suppose that player 1 plays s_1 repeatedly. Then regardless of the (possibly history-dependent) strategy of player 2, the Azuma-Hoeffding inequality together with the Borel-Cantelli lemma can be easily applied to establish that:

$$\liminf_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \Pi(a_1^t, a_2^t) \geq \mathrm{val}(\Pi), \text{ almost surely.} \tag{1}$$

(To use the Azuma-Hoeffding inequality, just observe that $\Pi(s_1, a_2^t) - \Pi(a_1^t, a_2^t)$ is a martingale difference sequence.) Alternatively, observe the preceding relationship holds if player 1 plays a Hannan consistent strategy.

For any $v \leq \mathrm{val}(\Pi)$, the preceding paragraph establishes that there exists a strategy for player 1 such that, regardless of the (possibly history-dependent) strategy of player 2, there holds:

$$\liminf_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \Pi(a_1^t, a_2^t) \geq v, \text{ almost surely.} \tag{2}$$

Applying the preceding insight to player 2, the following converse holds as well. Given $\varepsilon > 0$, for any strategy of player 1 there exists a (possibly history-dependent) strategy of player 2 such that:

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} \Pi(a_1^t, a_2^t) < v + \varepsilon, \text{ almost surely.}$$

Summarizing, player 1 can guarantee that his average payoff converges to the set $[v, \infty)$ if and only if $v \leq \mathrm{val}(\Pi)$. A set $S = [v, \infty)$ is called *approachable* if there exists a strategy for player 1 such that, regardless of the strategy of player 2, condition (2) holds,

2.2 The Vector-Valued Case

To generalize the preceding development to the vector-valued case, we wish to study sets $S \subset \mathbb{R}^K$ such that player 1 can guarantee the average payoff $\frac{1}{T}\sum_{t=0}^{T-1} \mathbf{\Pi}(a_1^t, a_2^t)$ converges to the set S. Approachability generalizes to the vector-valued payoff setting in a natural way: a set S is called *approachable* if there exists a (possibly history-dependent) strategy for player 1 such that, regardless of the (possibly history-dependent) strategy of player 2, there holds:

$$\lim_{T\to\infty} d\left(\frac{1}{T}\sum_{t=0}^{T-1} \mathbf{\Pi}(a_1^t, a_2^t), S\right) = 0, \text{ almost surely,}$$

where $d(\boldsymbol{v}, S) = \inf_{\boldsymbol{y}\in S} \|\boldsymbol{y} - \boldsymbol{v}\|$ is the Euclidean distance from $\boldsymbol{v}$ to the set S.

In the scalar case, our study of approachability yields that $[v, \infty)$ is approachable if and only if $v \leq \mathrm{val}(\Pi)$, or equivalently, if and only if there exists a mixed action s_1 for player 1 such that:

$$v \leq \min_{a_2 \in A_2} \Pi(s_1, a_2).$$

We start by using this insight from the scalar case to study approachability of *halfspaces* in the case of vector-valued payoffs.

In particular, let S have the following form, where $\|\boldsymbol{V}\| = 1$:

$$S = \{\boldsymbol{u} \in \mathbb{R}^K : \boldsymbol{V} \cdot \boldsymbol{u} \geq v\}.$$

Such a set S is a halfspace in $\mathbb{R}^K$, with $K-1$-dimensional tangent plane $\{\boldsymbol{u} : \boldsymbol{V} \cdot \boldsymbol{u} = v\}$. To investigate approachability of the halfspace S, consider a scalar zero-sum game where $\hat{\Pi}(a_1, a_2) = \boldsymbol{V} \cdot \boldsymbol{\Pi}(a_1, a_2)$. In this scalar game, the set $[v, \infty)$ is approachable if and only if there exists a mixed action s_1 such that:

$$v \leq \min_{a_2 \in A_2} \hat{\Pi}(s_1, a_2).$$

But note that approachability of thet set $[v, \infty)$ in the scalar game is equivalent to approachability of the set S in the original game. We conclude the halfspace S is approachable if and only if there exists a mixed action s_1 such that:

$$v \leq \min_{a_2 \in A_2} \boldsymbol{V} \cdot \boldsymbol{\Pi}(s_1, a_2). \tag{3}$$

We are now ready to state Blackwell's approachability theorem. While approachability of halfspaces can be studied using scalar zero-sum games, Blackwell's theorem provides the analytical tool necessary to establish approachability of general convex sets S.

Theorem 1 (Blackwell [1]) *A closed, convex set $S \subset \mathbb{R}^K$ is approachable if and only if all halfspaces containing S are approachable.*

3 Proof of Blackwell's Theorem

One direction of the proof is trivial: if S is approachable, then all halfspaces containing S are also approachable. We only need to show that if all halfspaces containing S are approachable, then S is approachable. As observed above, the proof proceeds by constructing a strategy that "mixes" the optimal strategies from each of the halfspaces containing S, to build a strategy where average payoff always converges to S.

We will need some additional notation. Let $\boldsymbol{\Pi}^{T-1} = \frac{1}{T}\sum_{t=0}^{T-1} \Pi(a_1^t, a_2^t)$ denote the average payoff of player 1 up to time $T-1$. For any vector $\boldsymbol{v}$, we let $P_S(\boldsymbol{v})$ denote the *projection* of $\boldsymbol{v}$ onto S:

$$P_S(\boldsymbol{v}) = \arg\min_{\boldsymbol{y} \in S} \|\boldsymbol{y} - \boldsymbol{v}\| = \arg\min_{\boldsymbol{y} \in S} \|\boldsymbol{y} - \boldsymbol{v}\|^2.$$

(Since the last expression is a minimization problem with convex feasible region and strictly convex objective function, we conclude $P_S(\boldsymbol{v})$ is uniquely defined.) Note that $d(\boldsymbol{v}, S) = \|P_S(\boldsymbol{v}) - \boldsymbol{v}\|$.

The idea behind Blackwell's proof is simple and constructive. Suppose that the average payoff $\boldsymbol{\Pi}^{T-1}$ does not lie in S. Then Blackwell's strategy suggests first projecting $\boldsymbol{\Pi}^{T-1}$ to the set S, and then playing the optimal strategy for the halfspace containing S that is "tangent" to S at $P_S(\boldsymbol{\Pi}^{T-1})$; see Figure 1. The proof amounts to showing that such a strategy reduces the distance of the average payoff to S (in expectation).

Formally, we define a strategy for player 1 as follows. The initial action a_1^0 can be chosen according to any mixed action for player 1. If $\boldsymbol{\Pi}^{T-1} \in S$, then a_1^T can be chosen according to any

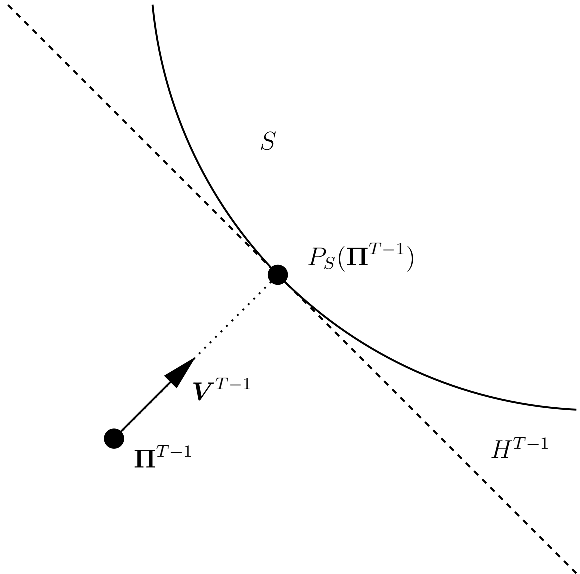

Figure 1: *Proof of Blackwell's Theorem.* The average payoff up to time $T-1$ is $\mathbf{\Pi}^{T-1}$. The figure assumes that $\mathbf{\Pi}^{T-1} \notin S$. $P_S(\mathbf{\Pi}^{T-1})$ is the projection of $\mathbf{\Pi}^{T-1}$ onto the set S, i.e., the element of S closest (in Euclidean distance) to $\mathbf{\Pi}^{T-1}$. The vector $\boldsymbol{V}^{T-1}$ is the unit vector in the direction of S, i.e., in the direction $P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}^{T-1}$. The halfspace H^{T-1} is defined by: $H^{T-1} = \{\boldsymbol{u} : \boldsymbol{V}^{T-1} \cdot \boldsymbol{u} \geq \boldsymbol{V}^{T-1} \cdot P_S(\mathbf{\Pi}^{T-1})\}$.
Since the projection minimizes the Euclidean distance to the set S, the resulting optimality condition yields that $S \subset H^{T-1}$. The Blackwell strategy is for player 1 to choose a mixed action s_1^T that guarantees, regardless of the action a_2^T of player 2, that $\mathbf{\Pi}(s_1^T, a_2^T)$ lies in the halfspace H^{T-1}.

(history dependent) mixed action for player 1. Suppose instead that $\mathbf{\Pi}^{T-1} \notin S$; then we make the following definitions:

$$\boldsymbol{V}^{T-1} = \frac{P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}^{T-1}}{\|P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}^{T-1}\|}, \quad v^{T-1} = \boldsymbol{V}^{T-1} \cdot P_S(\mathbf{\Pi}^{T-1}).$$

Let the halfspace H^{T-1} be defined by:

$$H^{T-1} = \{\boldsymbol{u} : \boldsymbol{V}^{T-1} \cdot \boldsymbol{u} \geq v^{T-1}\}.$$

Optimality of the projection implies that for any $\boldsymbol{v}$, and for any $\boldsymbol{y} \in S$:

$$(P_S(\boldsymbol{v}) - \boldsymbol{v}) \cdot (\boldsymbol{y} - P_S(\boldsymbol{x})) \geq 0.$$

Rearranging the preceding expression, we conclude that $S \subset H^{T-1}$; i.e., H^{T-1} is a halfspace containing S. By assumption H^{T-1} is approachable, so there exists a mixed action s_1^T such that:

$$\min_{a_2 \in A_2} \boldsymbol{V}^{T-1} \cdot \mathbf{\Pi}(s_1^T, a_2) \geq v^{T-1}.$$

We consider the strategy for player 1 where he plays according to s_1^T at time T. Then:

$$\begin{aligned}
d(\mathbf{\Pi}^T, S)^2 &= \|P_S(\mathbf{\Pi}^T) - \mathbf{\Pi}^T\|^2 \\
&\leq \|P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}^T\|^2 \\
&= \left\| P_S(\mathbf{\Pi}^{T-1}) - \frac{T}{T+1}\mathbf{\Pi}^{T-1} - \frac{1}{T+1}\mathbf{\Pi}(a_1^T, a_2^T) \right\|^2 \\
&= \left\| \frac{T}{T+1}(P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}^{T-1}) + \frac{1}{T+1}(P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}(a_1^T, a_2^T)) \right\|^2 \\
&= \left(\frac{T}{T+1}\right)^2 \|P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}^{T-1}\|^2 + \left(\frac{1}{T+1}\right)^2 \|P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}(a_1^T, a_2^T)\|^2 \\
&\quad + \frac{2T}{T+1}(P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}^{T-1}) \cdot (P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}(a_1^T, a_2^T)). \qquad (*)
\end{aligned}$$

The first inequality follows by the minimum-norm property of the projection, and the remainder of the derivation is elementary.

Define M as:

$$M = \sup_{\boldsymbol{v}:\|\boldsymbol{v}\|\leq 1} \|P_S(\boldsymbol{v})\|.$$

The supremum is over the unit ball; thus we wish to upper bound the norm of the projection of the unit ball onto S. Note that for $\boldsymbol{v}$ with $\boldsymbol{v} \leq 1$, we have:

$$\|P_S(\boldsymbol{v})\| \leq \|\boldsymbol{v}\| + \|P_S(\boldsymbol{v}) - \boldsymbol{v}\| \leq 1 + d(\boldsymbol{v}, S).$$

Thus:

$$M \leq 1 + \sup_{\boldsymbol{v}:\|\boldsymbol{v}\|\leq 1} d(\boldsymbol{v}, S) < \infty,$$

where finiteness follows since $d(\cdot, S)$ is continuous and the unit ball is compact.

Recalling that all payoffs $\mathbf{\Pi}(a_1, a_2)$ lie in the unit ball, and using $M < \infty$, we obtain:

$$\|P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}(a_1^T, a_2^T)\|^2 \leq (1+M)^2.$$

Applying this to $(*)$ and rearranging terms, we obtain:

$$\begin{aligned}
(T+1)^2\|P_S(\mathbf{\Pi}^T) - \mathbf{\Pi}^T\|^2 - T^2\|P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}^{T-1}\|^2 \leq \\
(1+M)^2 + 2T(P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}^{T-1}) \cdot (P_S(\mathbf{\Pi}^{T-1}) - \mathbf{\Pi}(a_1^T, a_2^T)).
\end{aligned}$$

Summing terms, we obtain:

$$\|P_S(\mathbf{\Pi}^T) - \mathbf{\Pi}^T\|^2 \leq \frac{(1+M)^2(T-1)}{T^2} + \frac{2}{T}\sum_{t=0}^{T-1} \frac{t}{T}(P_S(\mathbf{\Pi}^t) - \mathbf{\Pi}^t) \cdot (P_S(\mathbf{\Pi}^t) - \mathbf{\Pi}(a_1^{t+1}, a_2^{t+1}))$$

Now note that $t/T < 1$ for $0 \leq t \leq T-1$. For notational convenience, define $\boldsymbol{V}^t = 0$ and $v^t = 0$ if $\mathbf{\Pi}^t \in S$. Then $P_S(\mathbf{\Pi}^t) - \mathbf{\Pi}^t = \|P_S(\mathbf{\Pi}^t) - \mathbf{\Pi}^t\|\boldsymbol{V}^t$. Further, for all t we have $\|P_S(\mathbf{\Pi}^t) - \mathbf{\Pi}^t\| \leq$

$1 + M$. Thus we conclude:

$$\|P_S(\mathbf{\Pi}^T) - \mathbf{\Pi}^T\|^2 \leq \frac{(1+M)^2}{T} + \frac{2(1+M)}{T} \sum_{t=0}^{T-1} \boldsymbol{V}^t \cdot (P_S(\mathbf{\Pi}^t) - \mathbf{\Pi}(a_1^{t+1}, a_2^{t+1})).$$

By our choice of s_1^{t+1}, we know that for all t:

$$\boldsymbol{V}^t \cdot P_S(\mathbf{\Pi}^t) = v^t \leq \boldsymbol{V}^t \cdot \mathbf{\Pi}(s_1^{t+1}, a_2^{t+1}).$$

(Observe here that we are using the fact that the inequality holds regardless of what pure action player 2 plays at time $t+1$. This is where approachability of the halfspace H^t is used.) Substituting this inequality gives:

$$\|P_S(\mathbf{\Pi}^T) - \mathbf{\Pi}^T\|^2 \leq \frac{(1+M)^2}{T} + \frac{2(1+M)}{T} \sum_{t=0}^{T-1} \boldsymbol{V}^t \cdot (\mathbf{\Pi}(s_1^{t+1}, a_2^{t+1}) - \mathbf{\Pi}(a_1^{t+1}, a_2^{t+1})). \quad (4)$$

Define:

$$\mathcal{X}_t = \boldsymbol{V}^t \cdot (\mathbf{\Pi}(s_1^{t+1}, a_2^{t+1}) - \mathbf{\Pi}(a_1^{t+1}, a_2^{t+1})). \quad (5)$$

Observe that $|X_t| \leq 2$, since all payoffs $\mathbf{\Pi}(a_1, a_2)$ lie in the unit ball. Further, $\{X_t\}$ is a martingale difference sequence with respect to the history; i.e., $\mathbb{E}[X_t|h^t] = 0$. Given $\varepsilon > 0$, by the Azuma-Hoeffding inequality we have:

$$\mathbb{P}\left(\frac{1}{T}\left|\sum_{t=0}^{T-1} X_t\right| > \varepsilon\right) \leq 2e^{-T\varepsilon^2/4}.$$

By the Borel-Cantelli lemma, we conclude that, almost surely, $\frac{1}{T}\left|\sum_{t=0}^{T-1} X_t\right| \leq \varepsilon$ for all but finitely many T. Since this holds for any $\varepsilon > 0$, we conclude that, almost surely, the right hand side of (4) converges to zero as $T \to \infty$. Thus $d(\mathbf{\Pi}^T, S) \to 0$ as $T \to \infty$ almost surely, as required. □

4 Remarks

In this section we gather together several remarks regarding the theorem and its proof:

1. The algorithm of the proof may be summarized as follows:

   (a) At time $t = 0$, player 1 can play according to any mixed action s_1^0.

   (b) At time $t > 0$, player 1 plays according to any mixed action s_1^t such that:

$$(P_S(\mathbf{\Pi}^{t-1}) - \mathbf{\Pi}^{t-1}) \cdot P_S(\mathbf{\Pi}^{t-1}) \leq \min_{a_2 \in A_2} (P_S(\mathbf{\Pi}^{t-1}) - \mathbf{\Pi}^{t-1}) \cdot \mathbf{\Pi}(s_1^t, a_2). \quad (6)$$

   The inequality (6) is sometimes called the *Blackwell condition*.

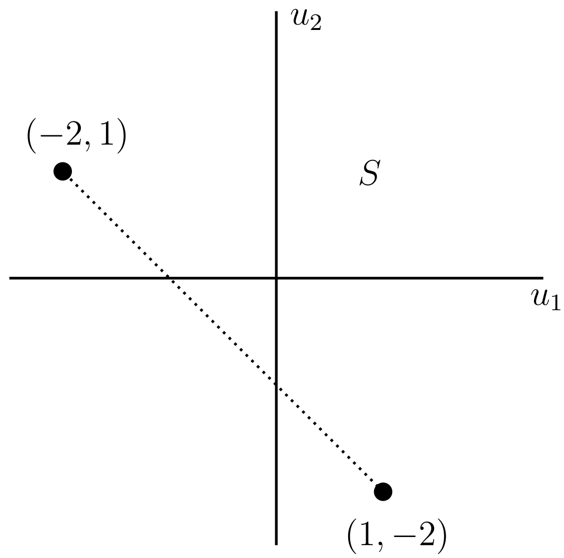

Figure 2: *Example in Remark 3.* In the example, player 1 achieves payoff $(-2, 1)$ if he plays A, and $(1, -2)$ if he plays B, regardless of player 2's action. The halfspaces where $u_1 \geq 0$ and $u_2 \geq 0$ are approachable, but their intersection $S = \{\boldsymbol{u} : u_1 \geq 0, u_2 \geq 0\}$ is not (since the convex hull of player 1's achievable payoffs lies outside S).

---

2. We can use the proof to obtain some insight into the rate of convergence of the algorithm used in the proof. Fix $T > 0$, and given $\delta > 0$, choose ε as:

$$\varepsilon = \sqrt{\frac{4}{T} \log\left(\frac{2}{\delta}\right)}.$$

Then recalling the definition of X_t in (5), the Azuma-Hoeffding inequality gives that:

$$\mathbb{P}\left(\frac{1}{T}\left|\sum_{t=0}^{T-1} X_t\right| > \varepsilon\right) \leq 2e^{-T\varepsilon^2/4} = \delta.$$

Thus, referring to (4), we conclude that (for fixed T), with probability at least $1 - \delta$, we have $d(\mathbf{\Pi}^T, S) \leq O(\sqrt{(1/T)\log(1/\delta)})$. Thus $d(\mathbf{\Pi}^T, S) \leq O(T^{-1/4})$ with high probability. This can be sharpened to $O(\sqrt{T})$ via a slightly different analysis, that uses a version of the Azuma-Hoeffding inequality for vector-valued martingales; see [2].

3. Note that, in general, *all* halfspaces containing S must be approachable for S to be approachable. For example, if S is the intersection of only finitely many halfspaces, it may not suffice that each of those halfspaces is approachable. To see this, consider a vector-valued zero-sum game where player 2 has no effect on player 1's payoff; and player 1 receives payoff $(-2, 1)$ if action A is played, and $(1, -2)$ if action B is played; see Figure 2.

We consider whether the set $S = \{\boldsymbol{u} : u_1, u_2 \geq 0\}$ is approachable. Clearly, the two halfspaces $S' = \{\boldsymbol{u} : u_1 \geq 0\}$ and $S'' = \{\boldsymbol{u} : u_2 \geq 0\}$ are approachable: the former if player 1 always plays B, and the latter if player 1 always plays A. However, the set $S = S' \bigcap S''$ is *not* approachable, since the convex hull of player 1's payoffs lies outside the set S.

4. A consequence of the preceding observation is that the theorem cannot be established by first proving a meta-theorem that "all intersections of approachable sets are approachable," since the preceding result is not true in general. However, we note that Blackwell's proof essentially amounts to establishing that by "mixing" the optimal strategies given by each halfspace containing S, one can create a strategy that approaches their intersection. Thus, in some sense, if one starts with "enough" approachable sets, then their intersection will be approachable.

5. While intersections of approachable sets need not be approachable, unions of approachable sets are always approachable; in fact, any superset of an approachable set is approachable. Thus one direction of the proof is trivial: if S is approachable, so is any halfspace containing S.

6. For a convex set S to be approachable, it suffices that we have approachability of all halfspaces "tangent" to S, i.e., whose tangent hyperplane is tangent to the set S. This follows since any halfspace containing S contains a halfspace tangent to S, and thus must be approachable by the preceding remark.

References

[1] D. Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6:1–8, 1956.

[2] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, Cambridge, United Kingdom, 2004.

[3] D. Foster and R. Vohra. Regret in the online decision problem. *Games and Economic Behavior*, 29:7–35, 1999.