CS 683 — Learning, Games, and Electronic Markets **Spring 2007**

Notes from Week 4: Approachability and internal regret

Instructor: Robert Kleinberg *12-16 Feb 2007*

1 Proof of Blackwell's approachability theorem

At the end of last week, we stated the following theorem without proof.

Theorem 1 (Blackwell's approachability theorem). *Let $\vec{\mathscr{G}}$ be a two-player game with vector payoffs in a bounded subset of $\mathbb{R}^n$, and let u be the payoff function of player 1. Let S be a nonempty closed convex subset of $\mathbb{R}^n$ such that for all halfspaces $H \supseteq S$, the set H is approachable. Then S is approachable.*

Proof. The set of possible payoffs of $\vec{\mathscr{G}}$ lies in a bounded subset of $\mathbb{R}^n$, so without loss of generality (rescaling payoffs if necessary) we may assume that all payoff vectors $u(x, y)$ satisfy $\|u(x, y)\|_2 \leq 1$. A set S is approachable if and only if its intersection with the convex hull of the set of payoff vectors is approachable, so we may also assume without loss of generality that all vectors $s \in S$ satisfy $\|s\|_2 \leq 1$.

Gives the hypothesis on $\vec{\mathscr{G}}$, we will explicitly construct an online algorithm whose average payoffs converge into S. The algorithm is simple. Let x_t, y_t denote the strategies chosen by the algorithm and adversary at time t, and let $u_t = u(x_t, y_t)$. Let

$$A_T = \frac{1}{T} \sum_{t=1}^{T} u_t$$

denote player 1's average vector payoff up to time T. If $A_T \in S$, then player 1 picks an arbitrary strategy at time $T + 1$. Otherwise, let B_T denote the point in S which is closest to A_T. Elementary geometry establishes that S is contained in the halfspace $H = \{v : v^T(A_T - B_T) \leq 0\}$, so there is a mixed strategy p such that $u(p, q) \in H$ for every mixed strategy q of player 2. The algorithm selects a random strategy x_t by sampling from this mixed strategy p.

Let d_t denote the distance from A_t to the nearest point of S (in the ℓ_2 norm). Now let's compute the expectation of d_{t+1}^2, given the transcript of play up to time t, i.e. the sequence of random variables $x_1, \ldots, x_t, y_1, \ldots, y_t$. We will use x_*, y_* as shorthand for this sequence.

$$\begin{aligned}
\mathbf{E}\left(d_{t+1}^2 \,|\, x_*, y_*\right) &\leq \mathbf{E}\left(\|A_{t+1} - B_t\|_2^2 \,|\, x_*, y_*\right) \\
&= \mathbf{E}\left(\left\|\frac{t}{t+1}(A_t - B_t) + \frac{1}{t+1}(u_{t+1} - B_t)\right\|_2^2 \,\middle|\, x_*, y_*\right) \\
&= \frac{t^2}{(t+1)^2}d_t^2 + \frac{1}{(t+1)^2}\|u_{t+1} - B_t\|_2^2 \\
&\quad + \frac{2t}{(t+1)^2}\mathbf{E}\left((u_{t+1} - B_t)^T(A_t - B_t) \,|\, x_*, y_*\right) \\
&\leq \frac{t^2}{(t+1)^2}d_t^2 + \frac{1}{(t+1)^2}\|u_{t+1} - B_t\|_2^2 \\
&\leq \frac{t^2}{(t+1)^2}d_t^2 + \frac{4}{(t+1)^2},
\end{aligned}$$

where the last line follows from our assumption that all payoff vectors and all points of S belong to the unit ball of $\mathbb{R}^n$. Now let $Z_t = t^2 d_t^2 - 4t$. We have

$$\mathbf{E}(Z_{t+1} \,|\, x_*, y_*) \leq Z_t,$$

i.e. Z_t is a supermartingale. To apply Azuma's inequality, we need to have an upper bound on $|Z_{t+1} - Z_t|$.

For a set $Q \subseteq \mathbb{R}^n$ and a vector $x \in \mathbb{R}^n$, let $\operatorname{dist}(x, Q)$ denote $\inf_{z \in Q} \|x - z\|_2$. We have

$$\begin{aligned}
Z_{t+1} - Z_t &= (t+1)^2 d_{t+1}^2 - 4(t+1) - t^2 d_t^2 + 4t \\
&= (t+1)^2 d_{t+1}^2 - t^2 d_t^2 - 4 \qquad (1)
\end{aligned}$$

Let $\zeta_t = t d_t$. We may rewrite the right side of (1) in terms of ζ_t and ζ_{t+1}:

$$(t+1)^2 d_{t+1}^2 - t^2 d_t^2 - 4 = (\zeta_{t+1} - \zeta_t)(\zeta_{t+1} + \zeta_t) - 4. \qquad (2)$$

Our assumption that all payoff vectors and all points of S lie in the unit ball implies that for all t, $d_t \geq 2$. Thus

$$\zeta_{t+1} + \zeta_t \leq 2(t+1) + 2t = 4t + 2. \qquad (3)$$

Observe that $\zeta_t = \operatorname{dist}(tA_t, tS)$. Recalling that B_t is the point of S which is closest to A_t (and therefore tB_t is the point of tS which is closest to tA_t), and observing that $(t+1)A_{t+1} = tA_t + u_{t+1}$, we see that

$$\begin{aligned}
\zeta_{t+1} &= \|(t+1)A_{t+1} - (t+1)B_{t+1}\|_2 \\
&\leq \|(t+1)A_{t+1} - (t+1)B_t\|_2 \\
&= \|t(A_t - B_t) + (u_{t+1} - B_t)\|_2 \\
&\leq \|t(A_t - B_t)\|_2 + \|u_{t+1} - B_t\|_2 \\
&= \zeta_t + \|u_{t+1} - B_t\|_2 \\
&\leq \zeta_t + 2
\end{aligned}$$

where the last line follows from the fact that u_{t+1} and B_t are both contained in the unit ball. Similarly,

$$\begin{aligned}\zeta_t &= \|tA_t - tB_t\|_2 \\ &\leq \|tA_t - tB_{t+1}\|_2 \\ &= \|(t+1)(A_{t+1} - B_{t+1}) - (u_{t+1} - B_{t+1})\|_2 \\ &\leq \|(t+1)(A_{t+1} - B_{t+1})\|_2 + \|u_{t+1} - B_{t+1}\|_2 \\ &\leq \zeta_{t+1} + 2.\end{aligned}$$

Hence

$$|\zeta_{t+1} - \zeta_t| \leq 2. \tag{4}$$

Combining (1)-(4), we obtain

$$|Z_{t+1} - Z_t| \leq 2(4t+2) + 4 = 8t + 8. \tag{5}$$

Setting $c_t = 8t + 8$, we have

$$\sum_{t=1}^{T} c_t^2 = 64 \sum_{t=1}^{T} (t+1)^2 = \frac{64}{6}(T+1)(T+2)(2T+3) - 64,$$

which is less than $25T^3$ for sufficiently large T. Setting $\sigma = \sqrt{\sum_{t=1}^{T} c_t^2} < 5T^{3/2}$ and $s = 2\sqrt{\ln(T)}$, we find that

$$\begin{aligned}\frac{1}{T^2} = e^{-\frac{1}{2}s^2} &> \Pr(Z_T > Z_0 + s\sigma) \\ &= \Pr\left(T^2 d_T^2 - 4T > 10(T^3 \ln(T))^{1/2}\right) \\ &= \Pr\left(d_T^2 > 10\left(\frac{\ln(T)}{T}\right)^{1/2} + \frac{4}{T}\right) \qquad (6) \\ &> \Pr\left(d_T > 4\left(\frac{\ln(T)}{T}\right)^{1/4}\right). \qquad (7)\end{aligned}$$

Summing over $T = 1, 2, \ldots$, we find that the expected number of T which satisfy $d_T > 4\left(\ln(T)/T\right)^{1/4}$ is finite. By the Borel-Cantelli Lemma, the number of such T is finite almost surely. Thus $d_T \to 0$ almost surely. □

Remark 1. In addition to proving that a closed convex set S is approachable if and only if every halfspace containing S is approachable, the proof actually established two stronger assertions which are worth mentioning separately.

1. If a closed convex set S is approachable at all, then there is an algorithm which ensures that the distance of the average payoff vector from S converges to zero at a rate of $O\left((\log(T)/T)^{1/4}\right)$.

2. This algorithm can be implemented efficiently, as long as we have efficient algorithms for implementing the following two operations:

    (a) Given a point $a \in \mathbb{R}^n$, find the point $b \in S$ which is closest to a.

    (b) Given a halfspace $H \supseteq S$, find a mixed strategy p for player 1, such that $u(p, q) \in H$ for every mixed strategy q for player 2.

2 Blackwell's theorem implies no-internal-regret learning algorithms

Recall that a *no-regret algorithm* for the best expert problem is one whose regret after T trials is $o(T)$. Similarly a *no-internal-regret algorithm* for the best expert problem is one whose internal regret after T trials is $o(T)$.

Last week we saw how to use Blackwell's theorem to derive the existence of no-regret algorithms for the best expert problem. Here, we recall that proof and recast its notation in a more linear-algebraic format.

Theorem 2. *There is a no-regret algorithm for the maximization version of the best expert problem with n experts and $[0, 1]$-valued payoffs.*

Proof. Consider the game $\vec{\mathscr{G}}$ in which player 1's strategy set is $[n]$ and player 2's strategy set is the set of all function $g : [n] \to [0, 1]$. For a pair (k, g), the payoff vector $u(i, g)$ is equal to the vector whose j-th component is $g(j) - g(k)$. We will identify such a function g with the column vector $(g(1), g(2), \ldots, g(n))^\mathsf{T} \in \mathbb{R}^n$. Likewise, we will identify a mixed strategy $p \in \mathbf{\Delta}([n])$ with the column vector $(p(1), p(2), \ldots, p(n))^\mathsf{T} \in \mathbb{R}^n$. Note that under these interpretations, the expected payoff vector obtained by playing p against g is

$$\begin{pmatrix} g(1) - \sum_i p(i)g(i) \\ \vdots \\ g(n) - \sum_i p(i)g(i) \end{pmatrix} = g - \mathbf{1}(p^\mathsf{T} g),$$

where $\mathbf{1} \in \mathbb{R}^n$ is the vector whose components are all equal to 1.

We want to prove that there is an algorithm whose average payoff vector after T trials approaches the negative orthant $S = \{(x_1, \ldots, x_n) \,|\, x_1, \ldots, x_n \leq 0\}$. We will prove that S is approachable, which establishes the theorem. According to Blackwell's approachability theorem, it suffices to prove that every halfspace containing S is approachable. Every such halfspace is defined by an inequality of the form $a^\mathsf{T} x \leq b$ where $b \geq 0$ and $a_i \geq 0$ for all i. To prove that $\{x \,|\, a^\mathsf{T} x \leq b\}$ is approachable, we will actually prove something stronger:

$$\forall a \in \mathbb{R}^n_+ \;\; \exists p \in \mathbf{\Delta}([n]) \;\; \forall g \in \mathbb{R}^n \;\; a^\mathsf{T}(g - \mathbf{1}(p^\mathsf{T} g)) = 0. \tag{8}$$

Indeed, (8) follows easily by taking p to be the vector $a/(a^\mathsf{T}\mathbf{1})$. ∎

The proof of the existence of no-internal-regret algorithms is very similar, but the algebra is trickier. We begin with a lemma.

Lemma 3. *If M is an $n \times n$ matrix satisfying:*

1. $M_{ij} \geq 0$ *for all* $i \neq j$
2. $\mathbf{1}^\mathsf{T} M = 0$

then there is a nonzero vector p such that $Mp = 0$ and $p_i \geq 0$ for all i.

Proof. Let I denote the $n \times n$ identity matrix. The hypothesis on M implies that for sufficiently small $\varepsilon > 0$, the matrix $L = I + \varepsilon M$ has non-negative entries, and its diagonal entries are strictly positive. It follows that for any vector $q \in \mathbb{R}^n_+ \setminus \{0\}$ the product Lq lies in $\mathbb{R}^n_+ \setminus \{0\}$. Thus the function $q \mapsto Lq/(\mathbf{1}^\mathsf{T} Lq)$ is a continuous mapping from $\mathbf{\Delta}([n])$ to itself. By Brouwer's fixed point theorem this mapping has a fixed point, i.e. a vector $p \in \mathbf{\Delta}([n])$ satisfying $Lp = p(\mathbf{1}^\mathsf{T} Lp)$. Observe that

$$\mathbf{1}^\mathsf{T} Lp = (\mathbf{1}^\mathsf{T} I - \varepsilon \mathbf{1}^\mathsf{T} M)p = \mathbf{1}^\mathsf{T} p,$$

which is equal to 1 for $p \in \mathbf{\Delta}([n])$. Hence $Lp = p$. This implies $\varepsilon Mp = 0$. □

Remark 2. Instead of using Brouwer's fixed point theorem, we could have deduced the final step using the Perron-Frobenius Theorem.

Theorem 4 (Perron-Frobenius). *If A is an $n \times n$ matrix with non-negative real entries, then*

- *There is an eigenvalue λ_{max} of A that is real and non-negative.*
- *There is at least one non-negative eigenvector corresponding to λ_{max}.*
- *For any other complex eigenvalue λ of A, we have $|\lambda| \leq \lambda_{max}$.*
- *If A is irreducible (i.e. the digraph on $[n]$ whose edges are pairs (i,j) such that $A_{ij} > 0$ is strongly connected), then λ_{max} has a 1-dimensional eigenspace and $|\lambda| < \lambda_{max}$ for all other complex eigenvalues λ.*

Theorem 5. *There is a no-internal-regret algorithm for the maximization version of the best expert problem with n experts and $[0,1]$-valued payoffs.*

Proof. This time, we design the game to have vector payoffs in the space $\mathbb{R}^{n \times n}$ of all $n \times n$ real matrices. Once again, player 1's strategy set is $[n]$ and player 2's strategy set is the set of functions from $[n]$ to $[0,1]$. In this game, however, the payoff vector $u(k,g)$ is defined to be the matrix $U = (U_{ij})$ given by

$$U_{ij} = \begin{cases} g(j) - g(k) & \text{if } i = k \\ 0 & \text{otherwise.} \end{cases}$$

Let $D : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ be the function which maps a column vector v to the diagonal matrix $D(v)$ whose i-th diagonal entry is v_i. Observe that if p is a mixed strategy of player 1, the payoff $u(p, g)$ is equal to the matrix

$$u(p, q) = gp^\intercal - \mathbf{1}g^\intercal D(p).$$

(Proof: check that both sides are linear functions of p, and they are equal whenever p is one of the standard basis vectors.)

The theorem amounts to saying that the set S of matrices with non-positive entries is approachable. Every halfspace in $\mathbb{R}^{n \times n}$ which contains S is defined by a linear inequality of the form $\mathrm{Tr}(A^\intercal X \leq b)$ where $b \geq 0$ and A is a matrix with non-negative entries. As before, we will prove that each such halfspace is approachable by proving the stronger claim that

$$\forall A \in \mathbb{R}_+^{n \times n} \;\; \exists p \in \mathbf{\Delta}([n]) \;\; \forall g \in \mathbb{R}^n \;\; \mathrm{Tr}(A^\intercal(gp^\intercal - \mathbf{1}g^\intercal D(p))) = 0. \tag{9}$$

We have

$$\mathrm{Tr}(A^\intercal gp^\intercal) = \mathrm{Tr}(p^\intercal A^\intercal g) = \mathrm{Tr}(g^\intercal Ap),$$

using the fact that the trace of a product of matrices is preserved under cyclic permutations of the factors, and also under transposition of the product matrix. Similarly,

$$\mathrm{Tr}(A^\intercal \mathbf{1} g^\intercal D(p)) = \mathrm{Tr}(g^\intercal D(p) A^\intercal \mathbf{1}).$$

Now we see that it suffices to construct $p \in \mathbf{\Delta}([n])$ such that

$$Ap = D(p)A^\intercal \mathbf{1}. \tag{10}$$

Using the identity $D(v)w = D(w)v$, which is valid for all pairs of vectors $v, w \in \mathbb{R}^n$, we can rewrite (10) as

$$\begin{aligned} Ap &= D(A^\intercal \mathbf{1})p \\ \left[A - D(A^\intercal \mathbf{1})\right] p &= 0. \end{aligned} \tag{11}$$

Observe that $M = A - D(A^\intercal \mathbf{1})$ satisfies the conditions of Lemma 3. The non-negativity of the off-diagonal entries is obvious, and the identity $\mathbf{1}^\intercal M = 0$ follows from the calculation

$$\begin{aligned} \mathbf{1}^\intercal M &= \mathbf{1}^\intercal A - \mathbf{1}^\intercal D(A^\intercal \mathbf{1}) \\ &= \left[A^\intercal \mathbf{1} - D(A^\intercal \mathbf{1})\mathbf{1}\right]^\intercal \\ &= \left[A^\intercal \mathbf{1} - D(\mathbf{1})A^\intercal \mathbf{1}\right]^\intercal \\ &= 0. \end{aligned}$$

Applying Lemma 3, we conclude that there is a nonzero vector p with non-negative entries which satisfies (11). We may rescale p if necessary to obtain a vector in $\mathbf{\Delta}([n])$, thus completing the proof of (9). □