Universität Bielefeld/IMW

Working Papers
Institute of Mathematical Economics

Arbeiten aus dem
Institut für Mathematische Wirtschaftsforschung

Nr. 182

On Optimal Strategies in Repeated Zero–Sum Games with Lack of Information on one Side

by

Martin Heuer

January 1990

H. G. Bergenthal

Institut für Mathematische Wirtschaftsforschung
an der
Universität Bielefeld
Adresse/Address:
Universitätsstraße
4800 Bielefeld 1
Bundesrepublik Deutschland
Federal Republic of Germany

0. Introduction

This paper presents a method to calculate recursively the value of a finitely repeated zero–sum game with incomplete information on one side. The analysis is not done directly by the recursive formula, which may be regarded as the informed player's view of the problem, but we rather take the uninformed player's point of view and compute an optimal strategy for him in the first place.

The idea is applied to a game that has received great attention in the literature since the convergence behaviour of its value function involves the density and the distribution function of the standard normal distribution (see Mertens and Zamir [76]). In their analysis the normal distribution springs forth from some differential equation. Here the N–stage game is explicitly solved and one finds that the normal distribution appears due to the central limit theorem.

1. The Model

The data of the game are listed below:

- a finite set R
  (set of types for player 1)
- for every type $r \in \mathbb{R}$ a mxn–matrix A^r
  (payoff matrices)
- a probability distribution $p \in \Delta(R)$
- a natural member N
  (number of stages)

The game is determined by the following rules:

- At state zero $r \in R$ is selected according to p. Both players know the distribution p but only player 1 learns his type.
- At stage $t = 1,\dots,N$ both players independently and simultaneously pick out actions $i_t \in \{1,\dots,m\}$ resp. $j_t \in \{1,\dots,n\}$.
- Afterwards they learn their opponent's choice of action (and nothing else).
- Both players have perfect recall.
- After stage N player 1 receives from player 2 the amount $\frac{1}{N} \sum_{t=1}^{N} A^r(i_t,j_t)$.

Let $H = \{1,\dots,m\} \times \{1,\dots,n\}$. In the following the identities $h = (i,j)$, $h_t = (i_t,j_t)$ and $h^t = (h_1,\dots,h_t)$ are always tacitly assumed.

A strategy of player 1 is given by a sequence of stochastic kernels $\sigma = (\sigma_1,\dots,\sigma_N)$,

$$\sigma_t \mid R \times H^{t-1} \Rightarrow \{1,\dots,m\}$$

(i.e. $\sigma_t(r,h^{t-1};\cdot)$ is a probability distribution on $\{1,\dots,m\}$ for all $(r,h^{t-1}) \in R \times H^{t-1}$) and a strategy of player 2 consists of a sequence $\tau = (\tau_1,\dots,\tau_N)$

$$\tau_t \mid H^{t-1} \Rightarrow \{1,\dots,n\}$$

The sets of strategies are denoted by

$$\Sigma^N = \prod_{t=1}^{N} \Sigma_t \quad , \quad T^N = \prod_{t=1}^{N} T_t .$$

Using the strategies $\sigma = (\sigma_1,\dots,\sigma_N)$, $\tau = (\tau_1,\dots,\tau_N)$ the players generate the distribution $\underline{P}^p_{(\sigma,\tau)}$ on $R \times H^N$:

$$\underline{P}^p_{(\sigma,\tau)}(r,h^N) = p(r) \prod_{t=1}^{N} \sigma_t(r,h^{t-1}; i_t)\, \tau_t(h^{t-1}; j_t)$$

and a payoff function on $\Sigma^N \times T^N$ is naturally defined as

$$\alpha^p_N(\sigma,\tau) = E_{\underline{P}^p_{(\sigma,\tau)}} \left(\frac{1}{N} \sum_{t=1}^{N} A_t \right) = E^p_{(\sigma,\tau)} \left(\frac{1}{N} \sum_{t=1}^{N} A_t \right)$$

A_t being a random variable on $R \times H^N$,
$A_t(r,h^N) = A^r(h_t)$.

Hence we have a non–cooperative two person zero–sum game

$$\Gamma_N(p) = (\Sigma^N, T^N, \alpha^p_N).$$

Γ_N (p) is defined in terms of mixed behaviour strategies. One could as well define the game in terms of behaviour strategies (i.e. functions $\sigma_t : R \times H^{t-1} \to \{1,\dots,m\}$ resp. $\tau_t : H^{t-1} \to \{1,\dots,n\}$) and consider its mixed extension (i.e. the probability distributions on the sets of pure strategies). In this case the min–max theorem guarantees the existence of a value. But the two approaches are equivalent in a certain sense (see e.g. Kuhn [53]) so that we may borrow the existence of a value $v_N(p)$ of $\Gamma_N(p)$ from the alternative approach.

The set of equilibrium points of $\Gamma_N(p)$ is denoted by $NE_N(p)$. Due to the rectangular property of equilibrium points in zero–sum game we may also write $NE_N(p)$ as the product of the player's optimal strategies:

$$NE_N(p) = NE_N^1(p) \times NE_N^2(p).$$

2. Properties of Equilibria

Definition 1:

Let $(\sigma,\tau) \in \Sigma^N \times T^N$. The vector $\xi_{(\sigma,\tau)} = (\xi_{(\sigma,\tau)}(r))_{r \in R}$,

$$\xi_{(\sigma,\tau)}(r) = \begin{cases} E^p_{(\sigma,\tau)}\left(\frac{1}{N}\sum_{t=1}^{N} A_t \mid r\right) & \text{if } p(r) > 0 \\ 0 & \text{if } p(r) = 0 \end{cases}$$

is called vector payoff of (σ,τ).

Thus the payoff function can be represented as follows:

$$\alpha^p_N(\sigma,\tau) = p \cdot \xi_{(\sigma,\tau)}$$

Lemma 2:

Let $(\sigma,\tau) \in NE_N(p)$. Then $\xi_{(\sigma,\tau)} \geq \xi_{(\sigma',\tau)} \quad \forall\, \sigma' \in \Sigma^N$.

Proof:

Suppose there were a strategy σ' such that

$\xi_{(\sigma,\tau)}(r_0) > \xi_{(\sigma',\tau)}(r_0)$ for at least one $r_0 \in R$.

In this case the strategy σ'',

$$\sigma''(r,\dots) = \begin{cases} \sigma'(r,\dots) & , \ r = r_0 \\ \sigma(r,\dots) & , \ r \neq r_0 \end{cases}$$

would yield a payoff exceeding the value of the game against player 2's equilibrium strategy. The contradiction proves the lemma.

q.e.d.

Proposition 3:

Let $(\sigma,\tau) \in NE_N(p)$, $\xi = \xi_{(\sigma,\tau)}$. Then

$$p \cdot \xi = v_N(p)$$

$$q \cdot \xi \geq v_N(q) \qquad \forall\, q \in \Delta(R).$$

Proof:

The equation just reflects the fact that the equilibrium strategies produce the equilibrium payoff. Suppose the inequality were violated for some $q \in \Delta(R)$. From the preceding lemma we know that $\xi \geq \xi_{(\sigma',\tau)} \quad \forall\, \sigma' \in \Sigma^N$. Consequently

$$q \cdot \xi_{(\sigma,\tau)} \leq q \cdot \xi < v_N(q).$$

Thus applying strategy τ in $\Gamma_N(q)$ player 2 could do better than the value permits.

q.e.d.

Corollary 4:

The function v_N is concave.

Define the sets

$$W_N = \{x \in \mathbb{R}^R : q \cdot x \geq v_N(q) \quad \forall\, q \in \Delta(R)\}$$

The boundary of W_N is given by

$$V_N = \{x \in W_N : \exists\, q \in \Delta(R) : q \cdot x = v_N(q)\}$$

(the set of supergradients of v_N)

$$V_N(p) = \{x \in V_N : p \cdot x = v_N(p)\}$$

(the superdifferential of v_N at p)

Remark that the statement "x is superdifferential of v_N at p" is equivalent to " $-p$ is a normal vector of W_N at x".

Every strategy $\sigma=(\sigma_1,\ldots,\sigma_N)\in\Sigma^N$ of player 1 can be regarded as a pair $(f,F)\in\Sigma^1\times(\Sigma^{N-1})^H$, that is a behaviour strategy for the first stage together with a mapping from H into the set of all $(N-1)$ –stage strategies. The canonical bijection between Σ^N and $\Sigma^1\times(\Sigma^{N-1})^H$ is given by the identity on Σ^1 and the equation

$$\sigma_{t+1}(r,(h,h^{t-1});i)=F(h)\,(r,h^{t-1};i)\qquad\forall\, r\in R,\ (h,h^{t-1})\in H^t,\ i\in\{1,\ldots,m\}$$
$$\forall\, t=1,\ldots,N-1$$

Analogously, a strategy $\tau=(\tau_1,\ldots,\tau_N)\in T^N$ is identified with a pair $(g,G)\in T_1\times(T^{N-1})^H$ by the equation

$$\tau_{t+1}(h,h^{t-1};j)=G(h)\,(h^{t-1};j)$$

The following representation of the payoff function reflects this point of view:

Lemma 5:

$$\alpha_N^P\,((f,F),(g,G))$$
$$=\ \frac{1}{N}\sum_r p(r)\sum_{h=(i,j)} f(r;i)\,g(j)\,(A^r(h)+(N-1)\,\alpha_{N-1}^{p^i}\,((F(h),G(h)))$$

with

$$p^i(r)=\underline{P}^p_{(f,g)}\,(r\,|\,i)\ =\frac{p(r)\,f(r;i)}{\sum_{r'} p(r')\,f(r';i)}$$

Proof:

Follows from the fact that

$$\underline{P}^p_{((f,F),(g,G))}\,(r,h^{N-1}\,|\,h)=\underline{P}^{p^i}_{(F(h),G(h))}\,(r,s,h^{N-1})$$

q.e.d.

Theorem 6:

Let $((f,F),(g,G) \in NE_N(p)$. If $\underline{P}^p_{(f,g)}(h) > 0$, then $(F(h),G(h)) \in NE_{N-1}(p^i)$.

Proof:

Suppose that $F(h) \notin NE^1_N(p^i)$ for some $h = (i,j)$ with $\underline{P}^p_{(f,g)}(h) > 0$. Since (g,G) must be an optimal reply to (f,F) we deduce with regard to the representation of the payoff function in Lemma 5 that

$$\alpha^{p^i}_{N-1}\,(F(h),G(h)) < v_{N-1}(p^i).$$

Now player 1 can improve his payoff by switching from F(h) to some equilibrium strategy $\sigma' \in NE^1_{n-1}(p^i)$, leaving the remainder of his strategy unchanged. Consequently (f,F) cannot be an equilibrium strategy.

q.e.d.

Corollary 7:

Let $((f,F),(g,G)) \in NE_N(p)$.

Then $\xi_{(f,F),(g,G)} \in V_N(p)$ and if $\underline{P}^p_{(f,g)}(h) > 0$, $\xi_{F(h),G(h)} \in V_{N-1}(p^j)$.

Corollary 8: (The recursive formula)

$$v_N(p) = \max_f \min_g \frac{1}{N} \sum_{r \in R} p(r) \sum_{h=(i,j)} f(r;i)\, g(j)\, (A^r(h) + (N-1)\, v_{N-1}(p^i))$$

Proposition 9:

For every $\xi \in V_N$ there is a strategy τ of player 2 enforcing ξ, i.e. $\xi_{(\sigma,\tau)} \leq \xi \quad \forall \sigma \in \Sigma^N$.

Proof:

If $\{\xi\} = V_N(p)$ for some p, ξ is enforced by any equilibrium strategy of player 2 in $\Gamma_N(p)$. Let $\xi \in V_N(p)$, $|V_N(p)| > 1$ (i.e. v_N is not differentiable at point p). v_N is piecewise linear on $\Delta(R)$ (see Ponssard and Sorin [78], section 3). Thus any (super–) gradient corresponding to a region of linearity including p is a supergradient of v_N at p, and $V_N(p)$ is simply the convex hull of all these supergradients. So let $\xi = \sum_k \lambda^k \xi^k$ be a finite convex combination of the ξ^k, each ξ^k being enforced by a strategy τ^k. Then ξ is enforced by $\tau = (\tau_1, \ldots, \tau_N)$,

$$\tau_T(h^{T-1}; j_T) = \frac{\sum_k \lambda^k \prod_{t=1}^{T} \tau_t^k (h^{t-1}; j_t)}{\sum_k \lambda^k \prod_{t=1}^{T-1} \tau_t^k (h^{t-1}; j_t)}$$

because $\underline{P}^p_{(\sigma,\tau)} = \sum_k \lambda^k \underline{P}^p_{(\sigma,\tau^k)} \quad \forall \sigma \in \Sigma^N.$

q.e.d.

An analogous argument shows that it is sufficient to know player 1's optimal strategies only for those $\Gamma_N(p)$ corresponding to the kinks of the value function:

Proposition 10:

Let $p = \sum_k \lambda^k p^k$, $v_N(p) = \sum_k \lambda^k v_N(p^k)$, $\sum_k \lambda^k = 1$, $\lambda^k \geq 0$, and let σ^k be an optimal strategy for player 1 in the game $\Gamma_N(p^k)$.

Then $\sigma = (\sigma_1,\ldots,\sigma_N)$

$$\sigma_T(r,h^{T-1};i_T) = \frac{\sum_k \lambda^k \frac{p^k(r)}{p(r)} \prod_{t=1}^{T} \sigma_t^k(r,h^{t-1};i_t)}{\sum_k \lambda^k \frac{p^k(r)}{p(r)} \prod_{t=1}^{T-1} \sigma_t^k(r,h^{t-1};i_t)}$$

is an optimal strategy of $\Gamma_N(p)$.

With regard to corollary 7 any equilibrium strategy of player 2 in $\Gamma_N(p)$ can be represented by a distribution y on his set of actions $\{1,\ldots,n\}$ and a tuple $X = (\xi^{i,j})_{(i,j)\in H}$ with $\xi^{i,j} \in V_{N-1}$, $\xi^{i,j}$ being the vector payoff player 2 is going to enforce from stage 2 to stage N after observing action i of player 1 and employing action j at the first stage. If the value function v_{N-1} of the $(N-1)$–stage game is differentiable at p^i he has no choice. $\xi^{i,j}$ is always the unique (super–)gradient of v_{N-1} at p^i. But even if he has a choice he does not lose anything by making $\xi^{i,j}$ independent of his own action j. Suppose there is an optimal strategy (y,X) for player 2 he can replace the $\xi^{i,j}$ by the expected vector payoff $\xi^i = \sum_j y_i \xi^{i,j}$ enforcing the same N–stage vector payoff ξ as before. In view of proposition 9 one is free to choose any $\xi^i \in V_{N-1}$.

Consequently

$$v_N(p)$$

$$= \min_{y\in\Delta(\{1,\ldots,n\})} \ \min_{X\in(V_{N-1})^m} \ \max_f \ \frac{1}{N}\sum_r p(r)$$

$$\Big(\sum_{i,j} f(r;i)\, y_j\, A^r(i,j) + (N-1)\,\xi^i(r)\Big)$$

$$= \min_{y,X} \ \frac{1}{N}\sum_r p(r) \ \max_i \Big(\sum_j y_j\, A^r(i,j) + (N-1)\,\xi^i(r)\Big)$$

If (y,X) represents an equilibrium strategy of player 2 he will get the vector payoff ξ,

$$\xi(r) = \max_i \frac{1}{N} \left(\sum_j y_j A^r(i,j) + (N-1)\, \xi^i(r) \right)$$

Generally $i_0 = \arg\max_i \frac{1}{N}(\ \dots\ )$ will not be unique. If there were only one maximizer i_0 given (y,X) in most cases player 2 would improve by modifying (y,X) such that $\frac{1}{N}\left(\sum_j y_j A^r(i_0,j) + (N-1)\, \xi^{i_0}(r)\right)$ is reduced at the cost of some other i_1 until the terms corresponding to i_0 and i_1 are equal. Moreover a unique maximizer i_0 would imply a "completely revealing" move of player 1. He would always choose action i_0 if his type is r such that the posterior probability of r given any action $i \neq i_0$ would be zero.

Exploring player 2's potentiality in the N–stage game, having solved the (N – 1)–stage game it seems a sensible idea to try to solve the equation

$$\sum_j y_j A^r(i,j) + (N-1)\, \xi^i(r) = N\, \xi(r) \qquad \forall\, i = 1,\dots,m$$

This is not always possible since completely revealing moves cannot be excluded, and if there is a solution it will in general not be unique, which implies that player 2 can choose which vector payoff he wants to attain (dependent on the prior distribution p).

In order to understand the importance of the first example, we have to mention a general result on repeated games with incomplete information (see e.g. Sorin [80]):

Let $A(p) = \sum_{r \in R} p(r)\, A^r$ be the so called non–revealing game, $u(p) = \operatorname{val} A(p)$ its value and define $v_\infty = \operatorname{cav} u(p)$ to be the concavification of the function u. Then

$$\lim_{N \to \infty} v_N(p) = v_\infty(p)$$

or, more precisely

$$0 \leq v_N(p) - v_\infty(p) \leq \frac{C}{\sqrt{N}}$$

for some constant $C \in \mathbb{R}$.

Example 1: (cf. Zamir [71/72], Mertens and Zamir [76])

$$A^1 = \begin{matrix} & L & R \\ T & 3 & -1 \\ B & -3 & 1 \end{matrix} \qquad A^2 = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

In this case $v_\infty(p) \equiv 0$, such that the above result implies that $0 \leq v_n(p) \leq \frac{C}{\sqrt{n}}$. Zamir [71/72] has shown that $\frac{p(1-p)}{\sqrt{n}} \leq v_n(p) \leq \frac{\sqrt{p(1-p)}}{\sqrt{n}}$. Especially $\frac{C}{\sqrt{n}}$ is the best overall bound for the difference $v_n - v_\infty$, there is a game with order of speed of convergence $\frac{1}{\sqrt{n}}$. This result has been refined by Mertens and Zamir [76], who showed that

$$\lim_{n \to \infty} \sqrt{n}\, v_n(p) = \frac{1}{\sqrt{2\Pi}} e^{-\frac{x_p^2}{2}}$$

where x_p is the p–quantible of the standard normal distribution, i.e.

$$\frac{1}{\sqrt{2\Pi}} \int_{-\infty}^{x_p} e^{-\frac{x^2}{2}} dx = p.$$

The proof is rather technical and does not give the intuition behind the result. It is based on a general result about the variation of martingales in [0,1]. In fact both results have been derived without computing any value $v_n(p)$. We are now in a position to make up for that and we will find that the result by Mertens and Zamir follows from the explicit representation of the value function. The value and the optimal strategies for both players will be described in terms of the binominal distribution. Let

$$b(k;n) = b(k;n, \tfrac{1}{2}) = \binom{n}{k} 2^{-n}$$

be the probability function of the binominal distribution with probability of success $\frac{1}{2}$, and let

$$B(k;n) = \sum_{m=0}^{k} b(m;n)$$

be its distribution function. In accordance with the interpretation of b(k;n) and B(k;n) we additionally define

$$b(k;n) = 0 \text{ if } k < 0 \text{ or } k > n \text{ and } B(k;n) = 0 \text{ if } k < 0,\ B(k;n) = 1 \text{ if } k > n.$$

Theorem 11:

The value functions v_n are piecewise linear, so they are well defined by the values at their kinks. The points of non–differentiability are given by

$$p_{k,n} = B(k-1;n)$$

and the values by

$$v_n(p_{k,n}) = \tfrac{1}{2}\, b(k-1;\, n-1)$$

An optimal strategy $(f,F) \in NE_n^1(p_{k,n})$ is defined as follows:

$$f(r_1;T) = \frac{1}{2}\,\frac{B(k-1;\, n-1)}{B(k-1;n)}$$

(Consequently we have $f(r_1;B) = \frac{1}{2}\,\frac{B(k;n-1)}{B(k-1;n)}$)

$$f(r_2;T) = \frac{1}{2}\,\frac{1 - B(k-2;n-1)}{1 - B(k-1;n)}$$

$$F(T) \in NE_{n-1}^1\,(p_{k-1,n-1})$$

$$F(B) \in NE_{n-1}^1\,(p_{k,n-1}) \quad ,k = 0,\dots,n+1.$$

(remark that according to the previous definition

$$p_{-1,n} = p_{0,n} = 0 \text{ and } p_{n+2,n} = p_{n+1,n} = 1 \,)$$

Moreover $NE^1_n(p_{k,n})$ turns out to be a singleton.

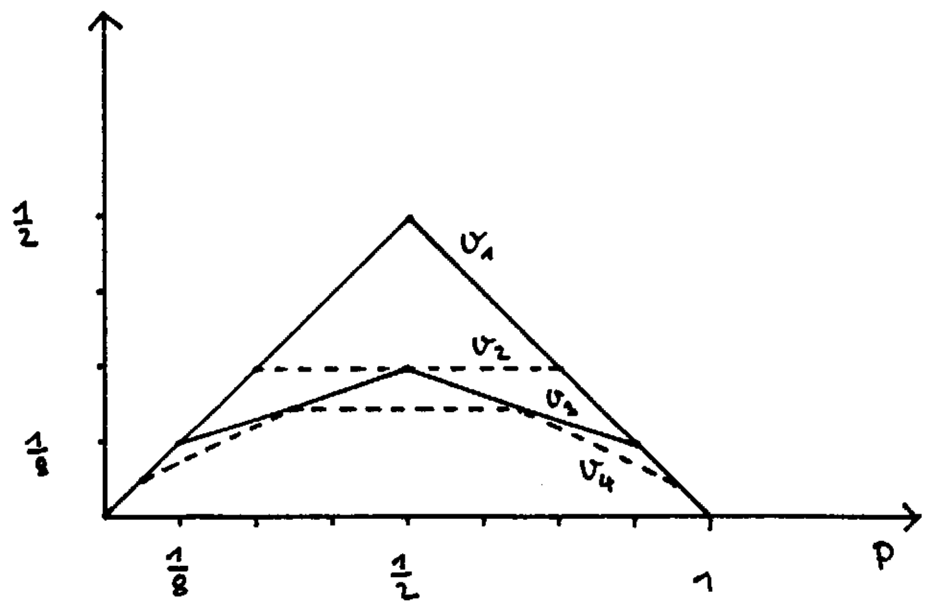
An optimal strategy for player 2 in $\Gamma_n(p)$, $p_{k,n} \leq p \leq p_{k+1,n}$, is given by

$$g(L) = \frac{1}{2} - \frac{1}{4} B(k-1; n-1) \quad , \quad k = 0,\dots,n$$

$$G(T) \in NE^2_{N-1}(p_{k,n-1}) \cap NE^2_{N-1}(p_{k+1,n-1})$$

$$G(B) \in NE^2_{N-1}(p_{k-1,n-1}) \cap NE^2_{N-1}(p_{k,n-1}) \quad k = 0,\dots,n+1.$$

Of course the intersection $NE^2_{N-1}(p_{k,n-1}) \cap NE^2_{N-1}(p_{k+1,n-1})$ is always non-empty. The instruction means that player 2 has to enforce the vector payoff corresponding to the interval $[\, p_{k,n-1}, p_{k+1,n-1}]$.



Before proving theorem 11 let us state some consequences. The equilibrium vector payoff of the n–stage game corresponding to the interval $[p_{k,n}, p_{k+1,n}]$, $\xi_{k,n} = (\xi^1_{k,n}, \xi^2_{k,n})$, satisfies

$$\xi_{k,n} \cdot (p_{k,n}, 1-p_{k,n}) = v_n(p_{k,n})$$

$$\xi_{k,n} \cdot (p_{k+1,n}, 1-p_{k+1,n}) = v_n(p_{k+1,n})$$

Explicitly:

(1) $\xi^2_{k,n} + (\xi^1_{k,n} - \xi^2_{k,n})\, B(k-1;n) = \frac{1}{2}\, b(k-1;n-1)$

(2) $\xi^2_{k,n} + (\xi^1_{k,n} - \xi^2_{k,n})\, B(k;n) = \frac{1}{2}\, b(k;n-1)$

Subtracting (1) from (2) yields:

$$(\xi^1_{k,n} - \xi^2_{k,n})\, b(k;n) = \frac{1}{2}\,(b(k;n-1) - b(k-1;n-1))$$

$$\Rightarrow \quad (\xi^1_{k,n} - \xi^2_{k,n}) = \frac{b(k;n-1) - b(k-1;n-1)}{2b(k;n)}$$

$$= \frac{\binom{n-1}{k} - \binom{n-1}{k-1}}{\binom{n}{k}}$$

$$= 1 - \frac{2k}{n} \qquad , k = 0, \ldots, n+1$$

$$\xi^2_{k,n} = \frac{1}{2}\, b(k;n-1) - (1 - \tfrac{2k}{n})\, B(k;n)$$

$$= \frac{2k}{n}\, B(k;n) - \frac{1}{2}(B(k;n-1) + B(k-1;n-1)) + \frac{1}{2}\, b(k;n-1)$$

$$= \frac{2k}{n}\, B(k;n) - B(k-1;n-1)$$

$$\xi^1_{k,n} = \xi^2_{k,n} + 1 - \frac{2k}{n}$$
$$= 1 - B(k-1,n-1) - \frac{2k}{n}(1 - B(k;n))$$

Adding (1) and (2) one obtains

$$2\,\xi^2_{k,n} + 2(\xi^1_{k,n} - \xi^2_{k,n})\,B(k;n+1) = b(k;n)$$

such that

$$v_{n+1}(p_{k+1,n+1}) = \tfrac{1}{2}(v_n(p_{k,n}) + v_n(p_{k+1,n})) = v_n(p_{k+1,n+1})$$

$\xi^1_{k,n} - \xi^2_{k,n} = 1 - \frac{2k}{n}$ implies that

$v_n'(p) = 1 - \frac{2k}{n}$ for $p \in (p_{k,n}, p_{k+1,n})$

Proof of theorem 11:

(by induction over n)

n = 1: The one stage game is transformed into a matrix game:

	L	R
T T	$2+p$	$-2+p$
T B	$-2+5p$	$2-3p$
B T	$2-5p$	$-2+3p$
B B	$-2-p$	$2-p$

E.g. the strategy B T for player 1 means:
Play top in the type is r_1 and play bottom if the type is r_2.
One checks easily that for $p \le \frac{1}{2}$ the mixed strategies

$$\left[\frac{1-2p}{2-2p}, \frac{1}{2-2p}, 0, 0\right], \left[\frac{1}{2}, \frac{1}{2}\right]$$

and for $p \ge \frac{1}{2}$ the strategies

$$\left[0, \frac{1}{2p}, 0, \frac{2p-1}{2p}\right], \left[\frac{1}{4}, \frac{3}{4}\right]$$

constitute an equilibrium. In terms of behaviour strategies:

$$f(r_1;T) = \begin{cases} 1 & , p \le \frac{1}{2} \\ \frac{1}{2p} & , p \ge \frac{1}{2} \end{cases} \qquad f(r_2;T) = \begin{cases} \frac{1-2p}{2-2p} & , p \le \frac{1}{2} \\ 0 & , p \ge \frac{1}{2} \end{cases}$$

$$g(L) = \begin{cases} \frac{1}{2} , & p \ge \frac{1}{2} \\ \frac{1}{4} , & p \le \frac{1}{2} \end{cases}$$

Consequently

$$v_1(p) = \min\{p, 1-p\}$$

$n \to n+1$:

At first we show that player 2 can enforce the vector payoff $\xi_{k,n+1}$ in the $(n+1)$–stage game by playing $(y, 1-y)$ at stage 1 regardless of player 1's action at stage 1 provided he is able to attain $\xi_{k,n}$ and $\xi_{k+1,n}$ in the n–stage game.

Let $y = \frac{1}{2} - \frac{1}{4} B(k-1;n)$

Then y must satisfy

$$(n+1)\ \xi_{k,n+1} = (4y-1,\ 4y-2) + n\ \xi_{k,n} \quad \text{(if player 1 plays top)}$$

and

$$(n+1)\ \xi_{k,n+1} = (1-4y,\ 2-4y) + n\ \xi_{k-1,n} \quad \text{(if player 1 plays bottom)}$$

The first equation is verified as follows:

$$\begin{aligned} & 4y - 2 + n\ \xi^2_{k,n} \\ = \quad & -B(k-1;n) + 2k\ B(k;n) - n\ B(k-1, n-1) \\ = \quad & -B(k-1;n) + 2k\ B(k;n+1) + k\ b(k;n) - n\ B(k-1;n) - \tfrac{n}{2}\ b(k-1;\ n-1) \\ = \quad & 2\ k\ B(k;\ n+1) - (n+1)\ B(k-1;n) \\ = \quad & (n+1)\ \xi^2_{k,n+1} \end{aligned}$$

$$4y - 1 + k\,\xi^1_{k,n}$$
$$= 4y - 2 + n\,\xi^2_{k,n} + 1 + n - 2k$$
$$= (n+1)\,\xi^2_{k,n+1} + (n+1) - 2k$$
$$= (n+1)\,\xi^1_{k,n+1}$$
$$\Rightarrow (n+1)\,\xi_{k,n+1} = (4y-1, 4y-2) + \xi_{k,n}$$

The second equation is checked analogously:

$$4y - 2 + n\,\xi^2_{k,n}$$
$$= -B(k-1;n) + 2k\,B(k-1;n) - n\,B(k-2, n-1)$$
$$= -B(k-1;n) + (2k-2)\,B(k-1;n) - n\,b(k-2; n-1)$$
$$= 2 - 4y + n\,\xi^2_{k-1,n}$$

$$4y - 1 + n\,\xi^1_{k,n}$$
$$= 4y - 2 + n\,\xi^2_{k,n} + 1 + n - 2k$$
$$= 2 - 4y + n\,\xi^2_{k-1,n} + n - (2k-1)$$
$$= 1 - 4y + n\,\xi^2_{k-1,n} - 2(k-1)$$
$$= 1 - 4y + n\,\xi^1_{k-1,n}$$
$$\Rightarrow (4y-1, 4y-2) + n\,\xi_{k,n} = (1-4y, 2-4y) + n\,\xi_{k-1,n}$$

So far we have shown that player 2 can guarantee the payoff $v_{n+1}(p)$ in $\Gamma_{n+1}(p)$, it is not yet clear that he cannot do better.

If we suppose for a moment that v_{n+1} is the value function of Γ_{n+1} (and that (g,G) is an equilibrium strategy for player 2) we can deduce player 1's optimal strategy in a game $\Gamma_{n+1}(p_{k,n+1})$. In $\Gamma_{n+1}(p_{k+1,n})$ player 2 can decide to enforce the vector payoff $\xi_{k,n+1}$ or $\xi_{k+1,n+1}$ (or any convex combination of these two).

If he chooses $\xi_{k,n+1}$ from the second stage on he must head for $\xi_{k-1,n}$ if player 1 plays bottom and for $\xi_{k,n}$ if player 1 plays top.

According to corollary 7 this implies that the conditional probability $\underline{P}(r_1|B)$ is located within interval $[p_{k-1,n}, p_{k,n}]$ and $\underline{P}(r_1|B)$ lies within the interval $[p_{k,n}, p_{k+1,n}]$.

$$
\xi_{k,n+1}
\begin{array}{l}
\overset{T}{\nearrow}\ \xi_{k,n} \qquad \leftrightarrow [p_{k,n}, p_{k+1,n}] \\
\\
\underset{B}{\searrow}\ \xi_{k-1,n} \qquad \leftrightarrow [p_{k-1,n}, p_{k,n}]
\end{array}
$$

If player 2 decides to enforce $\xi_{k+1,n+1}$ the situation is represented as follows:

$$
\xi_{k+1,n+1}
\begin{array}{l}
\overset{T}{\nearrow}\ \xi_{k+2,n} \qquad \leftrightarrow [p_{k+1,n}, p_{k+2,n}] \\
\\
\underset{B}{\searrow}\ \xi_{k+1,n} \qquad \leftrightarrow [p_{k,n}, p_{k+1,n}]
\end{array}
$$

Since both strategies of player 2 are equilibrium strategies in $\Gamma_{n+1}(p_{k+1,n})$ the posterior probabilities must satisfy both restrictions simultaneously:

$$\underline{P}(r_1|T) = p_{k+1,n}$$

$$\underline{P}(r_1|B) = p_{k,n}$$

Player 1's first stage strategy f is designed in such a way that it creates exactly these posteriors.

f is completely determined by the following lemma:

Lemma 12:

Let $p = \sum_{i=1}^{m} \lambda_i p_i$ be a convex combination of probabilities $p_i \in \Delta(R)$. Define $f \in \Sigma_1$ by

$$f(r;i) = \lambda_i \frac{p_i(r)}{p(r)}$$

Then $\underline{P}(i) = \lambda_i$, $p^i = p_i$

i.e. action i is played with a total probability of λ_i and the conditional probability on R given i, p^i equals p_i.

With respect to theorem 6 it is clear that (f,F) must satisfy
$F(T) \in NE_N^1 (p_{k+1,n})$ and $F(B) \in NE_N^1 (p_{k,n})$.

The proof is complete if we show that the strategy (f,F) that was derived under the hypothesis that $v_{n+1}(p)$ is the value of $\Gamma_{n+1}(p)$ actually guarantees the amount $v_{n+1} (p_{k+1,n+1})$ in the game $\Gamma_{n+1} (p_{k+1,n+1})$.

Since

$$v_{n+1} (p_{k+1,n+1}) = \frac{1}{2} (v_n(p_{k+1,n}) + v_n(p_{k,n}))$$

it suffices to check that the first stage payoff equals $v_{n+1}(p_{k+1,n+1})$ independently of player 2's action (cf. corollary 8).

Suppose player 2 plays left:

$$p_{k+1,n+1}(3\, f(r_1;T) - 3\, f(r_1;B))$$

$$+ (1 - p_{k+1,n+1})\, (2\, f(r_2;T) - 2\, f(r_2;B))$$

$$= \quad B(k;n+1)\, \frac{3}{2} \left[\frac{B(k\,;n)}{B(k\,;n+1)} - \frac{B(k-1;n)}{B(k;n+1)} \right]$$

$$+ (1 - B(k;n+1)) \left[\frac{1 - B(k;n)}{1 - B(k;\, n+1)} - \frac{1 - B(k-1;n)}{1 - B(k;n+1)} \right]$$

$$= \frac{3}{2}(B(k;n) - B(k-1;n)) + B(k-1;n) - B(k;n)$$

$$= \frac{1}{2}\, b(k;n)$$

$$= v_{n+1}(p_{k+1;n+1})$$

A similar computation shows that player 1 also gets $v_{n+1}(p_{k+1;n+1})$ if player 2 decides to play right.

q.e.d.

Theorem 13: (Mertens and Zamir [76])

$$\lim_{n \to \infty} \sqrt{n}\, v_n(p) = \frac{1}{\sqrt{2\Pi}}\, e^{-\frac{x_p^2}{2}}$$

x_p being the p–quantile of the standard normal distribution.

Proof:

Define

$$m(p,n) = \min \{k : B(k;n) > x_p\}$$

$$m'(p,n) = \min \{k : \frac{2k - n}{\sqrt{n}} > x_p\}$$

$$x_p(n) = \frac{2\, m(p,n) - n}{\sqrt{n}}$$

$$x_p'(n) = \frac{2\, m'(p,n) - n}{\sqrt{n}}$$

As an immediate consequence of these definitions we obtain

$$\lim_{n \to \infty} B(m(p,n);n) = p$$

$$\lim_{n \to \infty} x_p'(n) = x_p$$

and due to the central limit theorem it follows that

$$\lim_{n\to\infty} B(m'(p,n);n) = p$$

$$0 = \left| \lim_{n\to\infty} \left(B(m(p,n);n) - B(m'(p,n);n)\right) \right|$$

$$= \left| \lim_{n\to\infty} \int_{x'_p(n)}^{x_p(n)} e^{-\frac{t^2}{2}}\, dt \right|$$

$$= \left| \lim_{n\to\infty} \int_{x_p}^{x_p(n)} e^{-\frac{t^2}{2}}\, dt \right|$$

Consequently $x_p(n)$ converges as well and $\lim_{n\to\infty} x_p(n) = x_p$.

$$\lim_{n\to\infty} \frac{\sqrt{n-1}\quad v_n(p_{m(p,n)},n)}{\frac{2}{\sqrt{2\Pi(n-1)}}\; e^{-\frac{x_p(n-1)^2}{2}}}$$

$$= \lim_{n\to\infty} \frac{b(m(p,n)-1;n-1)}{\frac{2}{\sqrt{2\Pi(n-1)}}\; e^{-\frac{x_p(n-1)^2}{2}}}$$

$$= 1$$

due to the local central limit theorem and the convergence resp. boundedness of $x_p(n)$. Consequently

$$\lim_{n \to \infty} \sqrt{n-1}\, v_n\,(p_{m(p,n),n}) = \frac{1}{\sqrt{2\Pi}}\, e^{-\frac{x_p^2}{2}}$$

resp.
$$\lim_{n \to \infty} \sqrt{n}\, v_n\,(p_{m(p,n),n}) = \frac{1}{\sqrt{2\Pi}}\, e^{-\frac{x_p^2}{2}}$$

By definition of $m(p,n)$ and $p_{k,n}$ we have $p \in [\, p_{m(p,n),n}, p_{m(p,n)+1,n})$

Since $v_n' \equiv 1 - \frac{2\ m(p,n)}{n}$ on this interval:

$$\sqrt{n}\ |\ v_n(p) - v_n(p_{m(p,n),n})\ |$$

$$= \sqrt{n}\,(p - p_{m(p,n),n})\ |\ 1 - \frac{2\ m(p,n)}{n}\ |$$

$$= (p - p_{m(p,n),n})\ x_p(n)$$

so that finally

$$\lim_{n \to \infty} \sqrt{n}\ |\ v_n(p) - v_n(p_{m(p,n),n})\ |$$

$$= \lim_{n \to \infty} (p - p_{m(p,n),n}) \cdot \lim_{n \to \infty} x_p(n)$$

$$= \lim_{n \to \infty} (p - B(m(p,n) - 1;n)) \cdot x_p$$

$$= 0$$

resp.

$$\lim_{n \to \infty} \sqrt{n}\, v_n(p) = e^{-\frac{x_p^2}{2}}$$

q.e.d.

The next example is an exercise taken from Sorin [80], but here we overfulfill the task by solving the game completely.

Example 2: (cf. Sorin [80])

$$A^1 = \begin{matrix} \\ T \\ B \end{matrix} \begin{matrix} L & R \\ \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix} \end{matrix} \qquad A^2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$A(p) = p\, A^1 + (1-p)\, A^2 = \begin{bmatrix} 2p & 1-p \\ 2p & p \end{bmatrix}$$

$$u(p) = \text{val } A(p) = \begin{cases} 2p & , p \le \frac{1}{3} \\ 1-p & , \frac{1}{2} \le p \le \frac{1}{2} \\ p & , p \ge \frac{1}{2} \end{cases}$$

$$v_\infty(p) = \text{cav } u(p) = \begin{cases} 2p & , p \le \frac{1}{3} \\ \frac{1}{2} + \frac{p}{2} & , p \ge \frac{1}{3} \end{cases}$$

n = 1:

The value of the one–stage game is again computed by transforming it into a matrix game:

	L	R
T T	2p	1 − p
T B	2p	0
B T	2p	1
B B	2p	0

B T is a dominant strategy for player 1. Thus

$$v_1(p) = \begin{cases} 2p, & p \le \frac{1}{2} \\ 1, & p \ge \frac{1}{2} \end{cases}$$

That means in the one stage game player 2 can choose among the vector payoffs (2,0) and (1,1).

An optimal strategy for player 2:

n = 2:

Choosing $\xi^T = (2,0)$ and $\xi^B = (1,1)$ the equation

$$(2\,y, 1-y) + \xi^T = (1+y, 0) + \xi^B$$

can be solved. One finds $y = 0$, and player 2 is able to enforce the vector payoff $\frac{1}{2}((0,1) + (2,0)) = (1,\frac{1}{2})$ in the two–stage game. Of course he can also attain the vector payoffs (2,0) and (1,1) of the one stage game by playing the one–stage strategy twice. If $p \le \frac{1}{3}$, (2,0) gives him the best payoff. If $p > \frac{1}{3}$ he will strive for $(1,\frac{1}{2})$ which always dominates (1,1). So he is able at least to attain

$$v_2(p) = \begin{cases} 2p, & p \le \frac{1}{3} \\ \frac{1}{2} + \frac{p}{2}, & p \ge \frac{1}{3} \end{cases}$$

Player 2 has no chance to improve his payoff further since v_n is always a monotone decreasing sequence converging to v_∞ (see eg. Sorin [80]). In this case we find that v_2 already equals v_∞ such that v_2 really is the value of the two–stage game.

An optimal strategy for player 2 is given by

$$\tau_1(L) = \begin{cases} 1, & p < \frac{1}{3} \\ 0, & p \geq \frac{1}{3} \end{cases}$$

$$\tau_2(T;L) = 1 \qquad\qquad \tau_2(B;L) = \begin{cases} 1, & p < \frac{1}{3} \\ 0, & p \geq \frac{1}{3} \end{cases}$$

$n > 2$:

Since $v_n = v_\infty \ \forall\, n \geq 2$ the question is:

How can player 2 enforce the vector payoffs (2,0) and $(1,\frac{1}{2})$ in the n–stage game?

If he wants (2,0) he only has to repeat the strategy of the one–stage game.

If he wants $(1,\frac{1}{2})$ he must solve

$$n(1,\tfrac{1}{n}) \geq (2y, 1-y) + (n-1)\,\xi^T$$

$$n(1,\tfrac{1}{n}) \geq (1+y, 0) + (n-1)\,\xi^B$$

for a suitable choice of $\xi^T, \xi^B \in V_{n-1} = V_2$.

Taking $\xi^T = \lambda \cdot (2,0) + (1-\lambda)\,(1,\frac{1}{2})$ and $\xi^B = (1,\frac{1}{2})$ and one gets $\lambda = \frac{1}{n-1}$ and $y = 0$ such that

$$n(1,\tfrac{1}{2}) = (0,1) + (2,0) + (n-2)\,(1,\tfrac{1}{2}) \text{ (if player 1 plays T)} \qquad (3)$$

$$n(1,\tfrac{1}{2}) \geq (1,0) + (n-1)\,(1,\tfrac{1}{2}) \qquad \text{(if player 1 plays B)} \qquad (4)$$

In this example the situation is clear enough that we don't have to carry out the convex combination of vector payoffs according to proposition 9. If n is even the problem is exceptionally easy. Since the vector payoff $(1,\frac{1}{2})$ is precisely attained within two stages player 2 must only repeat the strategy of $\Gamma_2(p)$:

$$\tau_{2t-1}(\ldots;L) = \begin{cases} 1\ , p < \frac{1}{3} \\ 0\ , p \geq \frac{1}{3} \end{cases}$$

$$\tau_{2t}(\ldots T;L) = 1 \qquad\qquad \tau_{2t}(\ldots B;L) = \begin{cases} 1\ , p < \frac{1}{3} \\ 0\ , p \geq \frac{1}{3} \end{cases}$$

$$\forall\ t = 1,\ldots,\frac{n}{2}$$

If n is uneven player 2 may start as he did in the preceding case, i.e. he employs the above strategy for $t = 1,\ldots,\frac{n-3}{2}$. If $p \geq \frac{1}{3}$ he still has to achieve the average vector payoff $(1,\frac{1}{2})$ during the last three stages. That means we still need an optimal strategy for the three–stage game. With respect to (3) and (4) one finds

$$\tau_{n-2}(\ldots;L) = \begin{cases} 1\ , p < \frac{1}{3} \\ 0\ , p \geq \frac{1}{3} \end{cases}$$

If player 1 chooses T at this stage player 2 must obtain the vector payoff $\frac{1}{2}((2,0) + (1,\frac{1}{2}))$ during the last two stages. Since he cannot get $(1,\frac{1}{2})$ in the one stage game we cannot do with pure strategy choices. It is easily checked that he can guarantee the desired payoff by

$$\tau_{n-1}(\ldots T;L) = \begin{cases} 1, & p < \frac{1}{3} \\ \frac{1}{2}, & p \geq \frac{1}{3} \end{cases}$$

$$\tau_n(\ldots T\ T;L) = 1 \qquad \tau_n(\ldots T\ B;L) = \begin{cases} 1, & p < \frac{1}{3} \\ \frac{1}{2}, & p \geq \frac{1}{3} \end{cases}$$

If player 1 chooses B at stage $n-2$ we are again in a well known position. According to (4) player 2 must obtain the vector payoff $(1,\frac{1}{2})$ during the last two stages; i.e. he has to employ an optimal strategy in $\Gamma_2(p)$.

$$\tau_{n-1}(\ldots B;L) = \begin{cases} 1, & p < \frac{1}{3} \\ 0, & p \geq \frac{1}{3} \end{cases}$$

$$\tau_n(\ldots B\ T;L) = 1 \qquad \tau_n(\ldots B\ B;L) = \begin{cases} 1, & p < \frac{1}{3} \\ 0, & p \geq \frac{1}{3} \end{cases}$$

An optimal strategy for player 1:
If $p < \frac{1}{3}$ he always plays T, his dominant strategy in A(p).

If $p \geq \frac{1}{3}$ the posteriors player 1 must create are determined by theorem 6 and inequality (4). Theorem 6 says that $\underline{P}(r_1|T) = \frac{1}{3}$ and $\underline{P}(r_1|B) \in [\frac{1}{3},1]$. The second statement is not very helpful, but it is made more precise by inequality (4): If he chooses B his type must be r_1, i.e. $\underline{P}(r_1|B) = 1$. According to Lemma 12

$$\sigma_1(r_1;T) = \tfrac{3}{2}(1-p)\cdot\frac{\frac{1}{3}}{p} = \frac{1-p}{2p}$$
$$\sigma_1(r_2;T) = 1$$
$$(p = \tfrac{3}{2}(1-p)\cdot\tfrac{1}{3} + \tfrac{3}{2}(p-\tfrac{1}{3})\cdot 1)$$

From stage 2 on player 1 must act optimally in the game $\Gamma_{n-1}(\frac{1}{3})$ (if he has played top) or $\Gamma_{n-1}(1)$ (if he has played bottom). Since $v_{n-1}(\frac{1}{3}) = u_{n-1}(\frac{1}{3})$ he can defend his equilibrium payoff by playing optimally in the matrix game $A(\frac{1}{3})$ at ever stage after choosing action T at stage 1. Consequently

$$\sigma_t(T,...;T) = 1$$

$$\sigma_t(B,...;T) = 0 \qquad \forall\, t = 2,...,n$$

References

[53] Kuhn, H.W.: Extensive Games and the Problem of Information, Contributions to the Theory of Games II, Princeton

[76] Mertens, J.F. and S. Zamir: The Normal Distribution and Repeated Games, IJGT 5, 187 – 197

[80] Ponssard, J.P. and S. Sorin: The LP–Formulation of Finite Zero–Sum Games with Incomplete Information, IJGT 9, 99 – 105

[80] Sorin, S.: An Introduction to Two–Person Zero–Sum Repeated Games with Incomplete Information, IMSS–Economics, Stanford University, TR 312

[71/72] Zamir, S.: On the Relation between Finitely and Infinitely Repeated Games with Incomplete Information, IJGT 1, 179 – 198