ε-BAYESIAN SOLUTIONS OF ZERO-SUM GAMES IN EXTENSIVE FORM

J.P. Ponssard

April 1974 WP-74-11

Abstract

The paper introduces a procedure to select an "ε-Bayesian" optimal reply after a non-optimal move in zero-sum games defined in extensive form. The procedure was suggested by John Harsanyi. The "ε-Bayesian solution" for the class of zero-sum sequential games with incomplete information is derived.

"ε-Bayesian Solutions"

of

Zero-Sum Games in Extensive Form

1. Introduction

The traditional approach to games in extensive form had always been to put them in normal form and then derive the optimal behavioral strategies from the optimal mixed strategies. This procedure, while perfectly legitimate, had the obvious drawback of not going much insight into the interpretation of the extensive solution of the game. Indeed, while formal properties of optimal mixed strategies have been studied at length, not much was known about optimal behavioral strategies (with the notable exception of games with perfect information).

However, in recent years, there has been some renewed interest in games in extensive form [A-M, P, W] and the study of some examples pointed out that indeed optimal behavioral strategies do have a significantly different rationale (if they have any) than the mixed strategies. For instance, it can be shown that optimal behavioral strategies do not guarantee any security level conditional of what is learned during the game. Nevertheless, they seem to maximize conditional expectations given the other players optimal strategies (see [A-M]). Thus, their rationale would be better interpreted in an equilibrium framework than in a minimax one. The objective of this paper is to develop such an interpretation in a Bayesian context. Such an exercise is appealing for at least one reason. From a practical standpoint, it is worthwhile to compare the recommendations of game theory with those of decision theory and inconsistencies should be resolved or at least thoroughly understood.

Now, it is quite obvious that optimal behavioral strategies should maximize a player's conditional expectations at each information set provided this information set may be obtained with some positive probability (since otherwise a Bayesian best reply at this information set would generate not only a preferable conditional expectation but a preferable unconditional one as well, hence a contradiction). Thus, our main task will be concerned with the question of what to do after a non-optimal move.

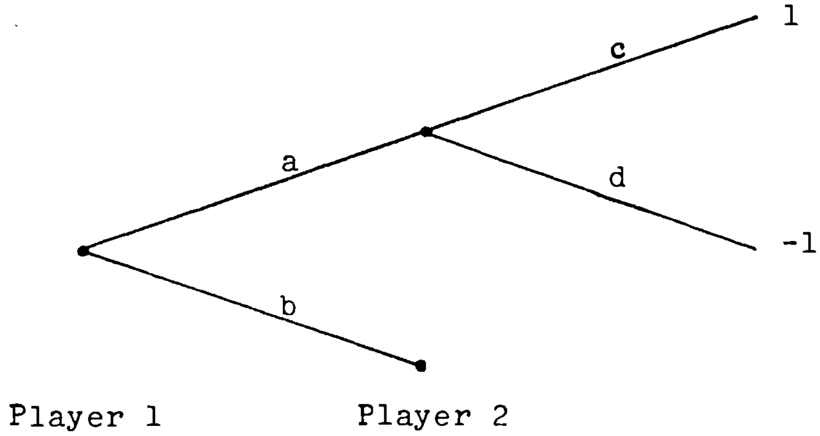
In principle game theory delineates a set of "optimal replies" for non-optimal moves. We wish to understand the rationale of such "optimal replies". An interpretation will be proposed: a player confronted with a non-optimal move should look at the game as the limiting case of an ε-game in which this non-optimal move had to be played with some small probability ε. Letting ε go to zero, "ε-optimal" replies will be derived. Ordinarily the subset of ε-optimal replies will be strictly included in the set of the optimal replies of the original game. Furthermore, it will be shown that ε-optimal replies are the limits of Bayesian replies for the ε-game. This procedure will define an "ε-Bayesian solution".

These ideas are illustrated by means of examples in the next section. Then in section 3 we derive the ε-Bayesian solutions for the class of sequential zero-sum games with incomplete information [P-Z].

2. The Main Ideas

2.1 An Introductory Example

Consider the following zero-sum game in extensive form in which Player 1 is the maximizer and Player 2 the minimizer.



A set of optimal behavioral strategies for this game is (b;d) but note that (b;αc + (1 - α) d)[1/] would do just as well as far as the value is concerned if $0 \leq \alpha \leq \frac{1}{2}$. Define a Bayesian optimal behavioral strategy as one which maximizes the player's conditional expectation at each information set. Then clearly, (b,d) is the Bayesian solution of this game.

Note that a strategy in which $\alpha \neq 0$ might still be interpreted as a threat: it is a commitment which should deter Player 1 of ever playing move a. This interpretation has serious drawbacks; first, threats should certainly play no role in a zero-sum context, second if Player 2 may commit himself then this should be explicitly

---

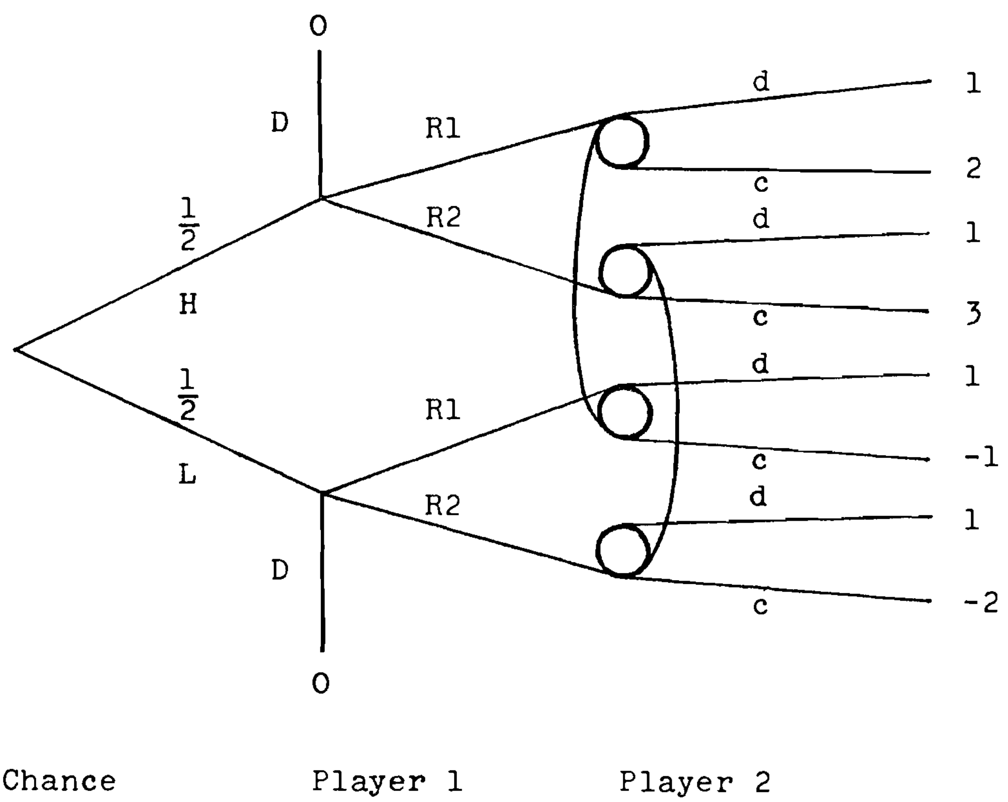
[1/] (b;αc + (1 - α) d) means that Player 1 selects move b and Player 2 selects move c with probability α and move d with probability (1 - α). This notation is used consistently.

modelled in the extensive form. Consequently this interpretation does not appear very convincing.

This example was presented to point out that we are likely to have some difficulties after non-optimal moves. Admittedly these difficulties are easily bypassed in this case. Our next example will show that these difficulties may be more serious but it will also introduce a general procedure to deal with them.

2.2 The Main Example

Consider the following example which may be interpreted as a one stage simplified poker game. Player 1 receives one card which may be low (L) or high (H), each with probability $\frac{1}{2}$. Then he may drop (D), raise 1 (R1) or raise 2 (R2). If Player 1 raised, Player 2 may drop (d) or call (c). The corresponding payoffs are shown in the tree, Player 1 is the maximizer.



To generalize the idea of a Bayesian solution to this game, the first difficulty concerns the definition of conditional expectations given that a player is at some information set. However, given Player 1's optimal strategy, then one may derive conditional probabilities on (H,L) depending on which move is played and then look for an optimal behavioral strategy for Player 2 which maximizes his conditional expectations. This procedure will work except for Player 1's non-optimal moves. There the problem of defining a Bayesian optimal behavioral strategy seems to be self-defeating since the conditional expectation does not exist. Let us take a closer look at the example.

There are two extremal sets of optimal behavioral strategies:

$$\alpha_1 = ((R2|H)^{1/} \quad \text{or} \quad (\tfrac{1}{3} D + \tfrac{2}{3} R2|L) \; ,$$
$$(\tfrac{2}{3} d + \tfrac{1}{3} c|R2) \quad \text{or} \quad (\tfrac{1}{2} d + \tfrac{1}{2} c|R1)) \; ,$$

$$\alpha_2 = ((R2|H) \quad \text{or} \quad (\tfrac{1}{3} D + \tfrac{2}{3} R2|L) \; ,$$
$$(\tfrac{2}{3} d + \tfrac{1}{3} c|R2) \quad \text{or} \quad (\tfrac{1}{3} d + \tfrac{2}{3} c|R1)) \; .$$

Given Player 1's optimal behavioral strategy, we may derive the following conditional probabilities:

$$\text{Prob }(H|D) = \frac{\text{Prob }(D|H)\text{ Prob }(H)}{\text{Prob }(D|H)\text{ Prob }(H) + \text{Prob }(D|L)\text{ Prob }(L)}$$
$$= \frac{0 \cdot \frac{1}{2}}{0 \cdot \frac{1}{2} + \frac{1}{3} \cdot \frac{1}{2}} = 0 \; ,$$

1/ (R2|H) means "play move D if move H is played". This notation is used consistently.

and similarly,

$$\text{Prob } (H|R2) = \frac{3}{5} \quad .$$

Consequently, if move R2 is played the conditional expectation of calling is

$$\frac{3}{5} \cdot 3 + \frac{2}{5} \cdot (-2) = 1 \quad .$$

Thus in terms of conditional expectation Player 2 is indifferent between move d or move c and his optimal reply at this stage $(\frac{2}{3}\, d + \frac{1}{3}\, c|R2)$ may indeed be qualified as Bayesian. So far, so good.

Note, however, that move R1 is non-optimal and that we may not define a conditional expectation given R1. But any optimal reply for Player 2, which has to be a convex combination of α_1 and α_2, will imply a randomization between d and c and so if we insist that this strategy be Bayesian, this will imply that the conditional expectations associated with c or d be equal. This in turn implies that

$$\text{Prob } (H|R1) = \frac{2}{3} \quad ,$$

(so that $\text{Prob } (H|R1) \cdot (2) + \text{Prob } (L|R1) \cdot (-1) = 1$).

At this point we have two possible interpretations; either to interpret Player 2's optimal behavioral strategy given R1 as a threat (see §2.1) or to insist on a Bayesian interpretation.

Let us try to pursue the logic of the Bayesian interpretation.

If this case "makes sense", it implies that

$$\text{Prob }(H|R1) = \frac{2}{3}$$

and thus if move R1 is played, then it was played with probability k, say, with a low card and 2k with a high card. So that

$$\text{Prob }(H|R1) = \frac{\text{Prob }(R1|H)\text{ Prob }(H)}{\text{Prob }(R1|H)\text{ Prob }(H) + \text{Prob }(R1|L)\text{ Prob }(L)}$$

$$= \frac{2k \cdot \frac{1}{2}}{2k \cdot \frac{1}{2} + k \cdot \frac{1}{2}} = \frac{2}{3} .$$

This suggests that if this non-optimal move was to be played, it should still be played according to some criterium![1/]

Let us then define an ε-game, the rules of which will converge to the rules of the original game as ε goes to zero. In the ε-game each personal information set should be obtained with a probability of at least ε. Hence the two constraints

$$\text{Prob }(R1|H)\text{ Prob }(H) + \text{Prob }(R1|L)\text{ Prob }(L) \geq \varepsilon \quad ,$$

$$\text{Prob }(R2|H)\text{ Prob }(H) + \text{Prob }(R2|L)\text{ Prob }(L) \geq \varepsilon \quad .$$

Letting ε go to zero the optimal behavioral strategies of the ε-game will converge to some optimal behavioral strategies in the original game. For ε small enough, the application of this procedure generates the following unique ε-optimal strategy (this will be proved in the next section).

[1/] This disturbing implication is the motivation of our analysis.

$$\alpha_\varepsilon = ((\tfrac{4}{3}\, \varepsilon R1 + (1 - \tfrac{4}{3}\, \varepsilon)\ R2|H) \quad \text{or}$$

$$(\tfrac{1}{3}(1 + \tfrac{2}{3}\, \varepsilon)\ D + \tfrac{2}{3}\, \varepsilon R1 + \tfrac{2}{3}(1 - \tfrac{4}{3}\, \varepsilon)\ R2|L) \quad ;$$

$$(\tfrac{2}{3}\ d + \tfrac{1}{3}\ c|R2) \quad \text{or} \quad (\tfrac{4}{9}\ d + \tfrac{5}{9}\ c|R1)) \quad .$$

Note that

$$\text{Prob } (R1|H) = 2 \text{ Prob } (R1|L)$$

and so

$$\text{Prob } (H|R1) = \tfrac{2}{3} \quad .$$

Thus, it looks like our assumption that R1 had to be played was the missing part of the Bayesian interpretation puzzle and, once this requirement is introduced, then the picture gets focused. Indeed, as ε goes to zero, α_ε converges to $(\frac{2}{3}\ \alpha_1 + \frac{1}{3}\ \alpha_2)$ and so Player 2's ε-optimal reply to move R1 converges to a subset of the optimal replies in the original game. We shall call this subset the "ε-Bayesian solution".

From a practical point of view, our answer to the problem of non-optimal move is certainly not entirely satisfactory and exogenous considerations should play a more significant role in the analysis [L-R, §4-11]. However, it is hoped that these purely endogenous theoretical considerations may help to develop a better understanding of the subject.

3. The ε-Bayesian Solution of Zero-Sum Sequential Games with Incomplete Information

This class of games was introduced in [P-Z], the value and the optimal behavioral strategies were explicitly derived. The objective of this section is to provide the ε-Bayesian solution of these games according to the ideas developed so far. The main difficulty concerns the degeneracy of the set of optimal behavioral strategies after a non-optimal move.

3.1 Definition of the Game

The game consists of four steps:

Step 0: Chance chooses a move $k \varepsilon K$ according to a probability distribution $p^0 = (p_k^0)_{k \varepsilon K}$. Player 1 is informed of the move chosen by Chance, Player 2 is not.

Step 1: Player 1 chooses a move $i \varepsilon I$. Player 2 is informed of the move chosen by Player 1.

Step 2: Player 2 chooses a move $j \varepsilon J$.

Final Step: Player 1 receives an amount a_{ij}^k (a real number) from Player 2.

(Assume that K, I, J are all finite sets).

3.2 Definition of the ε-Game

Let x_k^i = Prob (move i is played | move k was played). The following rule will be added to the original game. Let ε be a small positive number. It is required that

$$\forall i \in I \quad \sum_{k \in K} x_k^i p_k^0 \geq \varepsilon \quad .$$

Note that it is not assumed that each pure strategy be played with some small probability. Such an assumption would be more appropriate for the normal form and ordinarily will generate different results. Instead, we require that each of Player 2's information sets be obtained with a probability of at least ε. Thus the x_k^i's are weighted by the initial probabilities (p_k^0) and so a non-optimal move might not be played with the same probability for each k in order to satisfy the constraint.

3.3 The Solution of the ε-Game

For all probability distributions $p \varepsilon P = \{p = (p_k)_{k \varepsilon K} \mid p_k \geq 0$ $\Sigma p_k = 1\}$, define the function u,

$$u(p) = \underset{i \varepsilon I}{\text{Max}}\ \underset{j \varepsilon J}{\text{Min}} \sum_{k \varepsilon K} p_k a_{ij}^k \quad ,$$

and denote by $\bar{u}$ the concavification of u; i.e., the smallest concave function f which satisfies $f(p) \geq u(p)$ for all $p \varepsilon P$.

Denote by Λ_0 the set of all supporting hyperplanes to $\bar{u}$ at $p = p_0$ and for all $\lambda \varepsilon \Lambda_0$ and $i \varepsilon I$ denote by $\delta_i(\lambda)$, a non negative real number such that

$$\delta_i(\lambda) = \underset{p \varepsilon P}{\text{Min}}\ \underset{j \varepsilon J}{\text{Min}} \sum_{k \varepsilon K} p_k(\lambda_k - a_{ij}^k) \quad .$$

Clearly, δ_i is a continuous function of λ.

Theorem 3.3.1 For $\varepsilon \leq \varepsilon_0$, the ε-game has a value $V(p_0)$ and

$$V(p_0) = \bar{u}(p_0) - \varepsilon \operatorname{Max}_{\lambda\varepsilon\Lambda_0} \sum_{i\varepsilon I} \delta_i(\lambda) \quad .$$

We shall prove that each Player has a strategy which will guarantee him at least $V(p_0)$.

Let y_i^j = Prob (move j is played | move i was played). Denote by λ_0 a supporting hyperplane in Λ_0 such that

$$\sum_{i\varepsilon I} \delta_i(\lambda_0) = \operatorname{Max}_{\lambda\varepsilon\Lambda_0} \sum_{i\varepsilon I} \delta_i(\lambda) \quad ,$$

and, to simplify notation, let $\delta_i^0 = \delta_i(\lambda_0)$.[1/] For all i in I define a G_i game the $|K|x|J|$ payoff matrix of which is $||a_{ij}^k - (\lambda_0^k - \delta_i^0)||$.

Let y_i^j = Prob (move j is played | move i was played) and $y_i = (y_i^j)_{j\varepsilon J}$.

Lemma 3.3.2 The behavioral strategy $y = (y_i)_{i\varepsilon I}$ such that each y_i is optimal in each G_i respectively is optimal in the ε-game.

Proof: For each $i\varepsilon I$, the hyperplane $\lambda_0 - \delta_0^i$ is a supporting hyperplane to $\operatorname{Min}_{j\varepsilon J} \sum_{k\varepsilon K} p_k a_{ij}^k$. Hence, there exists a strategy $\bar{y}_i = (y_i^j)_{j\varepsilon J}$ such that (see lemma 2 in [P])

$$\forall p\varepsilon P \quad \sum_{k\varepsilon K} \sum_{j\varepsilon J} p_k a_{ij}^k \bar{y}_i^j \leq \sum_{k\varepsilon K} p^k(\lambda_0^k - \delta_i^0) \quad ,$$

1/ δ_i^0 may be interpreted as the penalty incurred to Player I by playing move i. For a non optimal move, $\delta_i^0 > 0$. Note that this penalty is the same for all k.

or $$\forall p \varepsilon P \quad \sum_{k\varepsilon K} \sum_{j\varepsilon J} p_k(a^k_{ij} - (\lambda^k_0 - \delta^0_i))\bar{y}^j_i \leq 0 \quad .$$

By definition of δ^0_i, there exists some $p^i \varepsilon P$ such that

$$\forall j \varepsilon J \quad \sum_{k\varepsilon K} p^i_k(a^k_{ij} - (\lambda^0_k - \delta^0_i) \geq 0 \quad ,$$

so that the value of G_i is zero and $\bar{y}_i$ is an optimal strategy in G_i. It follows that, for each $i\varepsilon I$,

$$\sum_{k\varepsilon K} \sum_{j\varepsilon J} p^0_k a^k_{ij} \bar{y}^j_i \leq (\sum_{k\varepsilon K} p^0_k \lambda^0_i) - \delta^0_i \quad .$$

But $\sum_{k\varepsilon K} p^0_k \lambda^0_k = \bar{u}(p^0_k)$ and so strategy $\bar{y}$ ensures that Player 2 will not lose more than

$$\begin{aligned} &\sum_{i\varepsilon I} \text{Prob (move i is played)}[\bar{u}(p^0) - \delta^0_i] \\ &= \bar{u}(p^0) - \sum_{i\varepsilon I} \delta^0_i \left[\sum_{k\varepsilon K} x^i_k p^0_k\right] \\ &\leq \bar{u}(p^0) - \varepsilon \sum_{i\varepsilon I} \delta^0_i \quad . \ || \end{aligned}$$

If λ^0 maximizes $\sum_{i\varepsilon I} \delta_i(\lambda)$ for $\lambda \varepsilon \Lambda_0$ then it can be shown[1/] that for $\varepsilon \leq \varepsilon_0$ there exists a convex combination $(\delta^0_i)_{i\varepsilon I}$ and points

---

1/ The proof is quite technical and will not be reproduced here.

p^i in P which satisfy

$$\sum_{i\epsilon I} \gamma_i^0 p^i = p^0 ,$$

$$\sum_{i\epsilon I} \gamma_i^0 = 1 ,$$

if $\delta_0^i > 0$ then $\gamma_i^0 = \varepsilon_0$,

if $\delta_0^i = 0$ then $\gamma_i^0 \geq \varepsilon_0$.

Lemma 3.3.3 The behavioral strategy for Player 1,

$$(k\epsilon K)(i\epsilon I) \quad x_k^i = \gamma_i^0 p_k^i / p_k^0 ,$$

is optimal in the ε_0-game.

Proof: Given Player 1's strategy, it is easily seen that Prob (move k was played | move i is played) = p_k^i. It follows that Player 1 cannot get less than

$$\sum_{i\epsilon I} \text{Prob (move i is played)} \min_{j\epsilon J} \sum_{k\epsilon K} p_k^i a_{ij}^k$$

$$= \sum_{i\epsilon I} \gamma_i^0 \Big(\sum_{k\epsilon K} p_k^i [\lambda_k^0 - \delta_i^0] \Big)$$

$$= \sum_{k\epsilon K} p_k^0 \lambda_k^0 - \sum_{k\epsilon I} \gamma_i^0 \delta_i^0$$

$$= \overline{u}(p^0) - \varepsilon_0 \sum_{i\epsilon I} \delta_i^0 \quad . \quad ||$$

The combination of the two lemmas proves the theorem. Letting ε go to zero, we obtain the ε-Bayesian solution of the original game.

References

[A-M] Aumann, R.J. and M. Maschler, "Some Thoughts on the Minimax Principle", Management Science, Vol. 18, No. 5.

[N-M] Von Neumann, J. and A. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, Princeton, N.J., 1944, Third Edition, 1953.

[W] Wilson, R., "Computing Equilibria of Two Person Games from the Extensive Form", W-P No. 176, May 1970, Graduate School of Business, Stanford University.

[P-Z] Ponssard, J.P. and S. Zamir, "Zero Sum Sequential Games with Incomplete Information", The Int. J. of Game Theory 2, No. 2.

[P] Ponssard, J.P., "Zero Sum Games with 'Almost' Perfect Information", forthcoming in Management Science.

[L-R] Luce, R.D. and H. Raiffa, "Games and Decisions", John Wiley, 1957.