AN ALGORITHM FOR SOLVING S-GAMES AND DIFFERENTIAL S-GAMES

Jerzy A. Filar[1] and T. E. S. Raghavan[2]

Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD. 21218[1]

Department of Mathematics, University of Illinois at Chicago Circle, Chicago, IL. 60680[2]

Summary

We present an algorithm for solving S-Games. Our algorithm can be used to compute approximately the value of the game as well as ε-optimal strategies of the two players. For games with similar structure to S-games which do not necessarily possess a value, the algorithm can sometimes be used as a heuristic procedure for determining the existence of a minimax solution. Further, it is shown that a certain simple class of differential games (we call them "differential S-games") can be viewed as static games and solved by the above procedure.

1. Introduction

We shall consider a two-person, zero-sum game played as follows: Player P_2 chooses a point s in a compact set S in R^n while player P_1 chooses a coordinate i of R^n. The payoff to P_1 is the i^{th} coordinate s_i of s. Such games were first studied by Blackwell and Girshick [2], pp. 47-51 and are often referred to as S-games.

In this note we present an algorithm (Section 2) for approximately computing the value and ε-optimal strategies, for both players in an S-game. Our algorithm reduces the problem to a sequence of minimization problems. Thus the computational strength of the procedure is determined by the strength of the existing techniques for solving the associated minimum problems. If the set S is not known to be compact our algorithm can lead to a heuristic procedure for testing the existence of the value.

We expect that one fruitful area of application for our algorithm will be that of finding solutions for what we call "Differential S-Games." These are differential games in which only one player (the minimizer) controls the law of motion, while the other player can decide only which out of a fixed number of functionals will be used to compute the payoff. Thus a differential S-game is an essentially control theoretic problem[3] complicated by the uncertainity as to which performance index is being used during the process. In Section 3 we show that under suitable assumptions these games can be analyzed as S-games, and consequently that our algorithm can be applied. Of course, the effectiveness of the algorithm will now depend on our ability to solve the associated sequence of optimal control problems. A simple example is given in Section 4.

2. An Algorithm for S-Games

Let Γ be an S-game as described in the introduction. For such a game Blackwell and Girshick [2] established the following result.

<u>Theorem 2.1.</u> The game has a value, v(Γ). Both players have optimal randomized strategies and further, player P_2 possesses an optimal strategy which randomizes on at most n points of S.

We shall now outline a procedure which shows that within some ε>0, an approximation to the value and to ε-optimal strategies for both players can be obtained by solving a sequence of appropriate matrix games, and a corresponding sequence of minimization problems. Our algorithm is a generalization of Troutt's (see [8], pp. 343-345) algorithm for non-convex S_n-games.

Let $\lambda = (\lambda_1 \dots \lambda_n)$ be an arbitrary probability vector, and define for each $s \in S$

$$H(\lambda,s) = \sum_{i=1}^{n} \lambda_i s_i. \qquad (1)$$

Here $s_i = i^{th}$ coordinate of s. Our algorithm will be as powerful as our ability to solve the following minimization problem:

$$\underset{s \in S}{\text{minimize}}\ H(\lambda,s). \qquad (2)$$

In the next two sections it will be seen that this algorithm is potentially at least, applicable to a large class of problems.

<u>Notation</u>: If P_1 uses a strategy $\lambda = (\lambda_1 \dots \lambda_n)$ and P_2 uses a strategy $\eta = (\eta_1 \dots \eta_k)$ on the points $s^1, s^2, \dots s^k$ of S (i.e., chooses s^j with probability η_j), then the expected payoff is simply

$$\Phi(\lambda,\eta) = \sum_{i=1}^{n} \sum_{j=1}^{k} \lambda_i \eta_j s_i^j, \qquad (3)$$

where $s_i^j = i^{th}$ coordinate of s^j. Further, let $A(s^1, s^2, \dots, s^k)$ be an n×k matrix whose j^{th} column is the point $s^j \in S$. Also, let v(A) stand for the value of a matrix game A.

The Algorithm

<u>Step 1.</u> [Initialization]. For each coordinate $i = 1,2,\dots,n$ solve the minimization problem (2) with $\lambda = e_i$, the i^{th} vector of the standard basis of R^n. This yields a set of n points in S, say, $\xi^1, \xi^2, \dots, \xi^n$. Now construct the matrix $A^1 = A(\xi^1, \xi^2, \dots, \xi^n)$, and compute its value, $v(A^1)$ and a pair of optimal strategy

[3] For an introduction to the theories of differential games and optimal control we refer the reader to Isaacs, [5] and Berkowitz [1], respectively.

CH1788-9/82/0000-0172$00.75

vectors λ^1 and η^1 for players P_1 and P_2 in the matrix game A^1.

Step 2. Solve the minimization problem (2) with $\lambda = \lambda^1$. This yields a point s^2 such that $H(\lambda^1,s^2) = \min H(\lambda^1,s)$. If $H(\lambda^1,s^2) = v(A^1)$, stop. Otherwise, construct the matrix game

$$A^2 = A^1(\xi^1,\xi^2,\ldots,\xi^n,s^2) \text{ and solve it.}$$

Step 3. Repeat Step 2 with $\lambda = \lambda^2$ (λ^2 is the P_1's optimal strategy in A^2), and so on.

The above procedure generates a sequence of solutions $v(A^m),\lambda^m,\eta^m$ to the matrix games A^m as well as a sequence of points $\xi^1,\ldots,\xi^n, s^2,s^3,s^4,\ldots$, in S. The relationship between these sequences and the solution of the original game is summarized in the following result.

Theorem 2.2.

(i) If the algorithm terminates in k iterations, then $v(\Gamma) = v(A^k),\lambda^k, \eta^k$ are optimal for P_1 and P_2, η^k is to be regarded as a randomization on the points of S corresponding to the columns of A^k.

(ii) $v(A^k) \to v(\Gamma)$ as $k \to \infty$.

(iii) Given $\varepsilon > 0$, there exists a positive integer M such that λ^m,η^m are ε-optimal for players P_1 and P_2 for all $m \geq M$.

Proof.

(i) The algorithm terminating in k steps means that $H(\lambda^k,s) \geq v(A^k)$ for all $s \in S$. But $\Phi(\lambda,\eta^k) \leq v(A^k)$ for all probability vectors λ since η^k is optimal in the matrix game A^k.

(ii) It follows that

$$H(\lambda^k,s^{k+1}) = \min_s H(\lambda^k,s) \leq v(\Gamma).$$

Since λ^{k+1} is optimal in the matrix game A^{k+1} and since $v(\Gamma) \leq v(A^{k+1})$ we have that for every k

$$H(\lambda^k,s^{k+1}) \leq v(\Gamma) \leq v(A^{k+1}) \leq H(\lambda^{k+1},\xi) \quad (4)$$

whenever $\xi \in \{\xi^1,\ldots,\xi^n,s^2,s^3,\ldots s^{k+1}\}$. Since the vectors λ^k are probability vectors, $\|\lambda^{k_\nu} - \lambda^{k_\nu+1}\|_\infty \to 0$ for some subsequence $\{\lambda^{k_\nu}\}_{\nu=1}^\infty$. Thus from (4) we obtain

$$0 \leq v(A^{k_\nu+1}) - v(\Gamma) \leq H(\lambda^{k_\nu+1},s^{k_\nu+1}) - H(\lambda^{k_\nu},s^{k_\nu+1})$$

$$\leq \|\lambda^{k_\nu+1} - \lambda^{k_\nu}\|_\infty \sum_{i=1}^n |s_i^{k_\nu+1}|. \quad (5)$$

Hence $v(A^{k_\nu}) \to v(\Gamma)$ as $\nu \to \infty$. Since $v(A^k) \downarrow$ limit, this limit must be $v(\Gamma)$.

(iii) The fact that η^m becomes ε-optimal randomized strategy for P_2 for m sufficiently large follows immediately from (ii), since P_1 has no restrictions placed on him in the matrix games A^m.

For player P_1, assume that there does not exist a positive integer M such that (iii) is satisfied for all $m \geq M$. This implies that there exists a subsequence $\{m_r\}_{r=1}^\infty$ of positive integers such that for every r,

$$H(\lambda^{m_r},s^{m_r+1}) < v(\Gamma) - \varepsilon. \quad (6)$$

Further, the subsequence of probability vectors $\{\lambda^{m_r}\}_{r=1}^\infty$ must have a convergent subsequence. Without loss of generality assume that $\|\lambda^{m_r} - \lambda^0\|_\infty \to 0$ as $r \to \infty$. Now it is possible to to find r and ℓ such that $m_r < m_\ell$ and

$$-\varepsilon/2 < H(\lambda^{m_r},s^{m_r+1}) - H(\lambda^{m_\ell},s^{m_r+1}) < \varepsilon/2. \quad (7)$$

Now using the fact that

$$v(\Gamma) \leq v(A^{m_\ell}) \leq H(\lambda^{m_\ell},s^{m_r+1})$$

we obtain a contradiction to (6), namely

$$v(\Gamma) - \varepsilon/2 \leq H(\lambda^{m_r},s^{m_r+1}). \quad (8)$$

□

We shall now briefly discuss the situation where the set S is not known to be compact. It could happen that the game still possesses a minimax solution, and our algorithm could sometimes still be useful as a heuristic procedure for "guessing" the solution. This is a consequence of the following result which can be very easily proved.

Corollary 2.3. Consider an "S-game" in which S is not necessarily compact. Let $\bar{\theta}_k = v(A^k)$ and suppose that $\underline{\theta}_k = \min_S H(\lambda^k,s)$ exists for every k. If the sequences $\{\underline{\theta}_k\}_{k=1}^\infty$ and $\{\bar{\theta}^k\}_{k=1}^\infty$ converge to a common limit θ, then the value exists and equals θ. Thus the algorithm of Theorem 2.2 still yields an approximate solution.

Remark 2.4. Note that as soon as our algorithm produces an interval $[\underline{\theta}_k,\bar{\theta}^k]$ which is small and $\underline{\theta}_{k+1}$ which is near $\underline{\theta}_k$, we are in possession of a pair (λ^{k+1},η^k) of reasonably good strategies for P_1 and P_2 in the following sense: The expected gain of P_1 will be at least $\underline{\theta}_{k+1}$ when he uses λ^{k+1} while the expected loss of P_2 will be no more than $\bar{\theta}^k$ when he uses η^k. With such a pair of strategies at hand, the question of whether the value actually exists may be of little practical interest.

3. A Class of Differential S-Games

In this section we introduce a special class of differential games which can be analyzed as static S-games and hence we shall call them Differential S-Games. Such a game can also be viewed as a game between a controller in the usual control theoretic sense and an antagonistic opponent who can decide which of the finitely many possible payoff functionals will be used to compute the controller's loss. That choice of the payoff functional remains unknown until the end of the game. More precisely, we consider a differential game Γ with R^m as state space played by two players P_1, P_2 over the fixed time interval $I = [0,1]$. At each time $t \in I$, P_2 picks an element $v(t)$ from a compact subset V of R^ℓ in such a way that $v(t)$ is measurable. Player P_1, on the other hand, picks out a number i from the set $N = \{1,2,\ldots,n\}$ at the beginning of the game.

The "law of motion" of the game is specified by the differential equation

$$\frac{dx}{dt} = f(t,x,v(t)). \tag{9}$$

Here $x \in R^m$ and $f: I \times R^m \times V \to R^m$ is a continous function satisfying a Lipschitz condition

$$\| f(t,x_1,v) - f(t,x_2,v) \| \leq K \| x_1 - x_2 \| \tag{10}$$

whenever $x_1, x_2 \in R^m$, $t \in I$, $v \in V$; K is some fixed positive number.

The equation (9) now has a unique solution $x(t)$ corresponding to any given initial condition $x(0) = x_0$; the resulting solution shall be called the trajectory corresponding to $v(\cdot)$. We may now compute the payoff $H(i,v)$ which P_2 will pay to P_1 at the end of the game, by

$$H(i,v) = \int_0^1 h(t,x(t),i,v(t))\, dt. \tag{11}$$

Here $h: I \times R^m \times N \times V \to R$ is a continuous uniformly bounded function.

We shall associate with every pure control the n-component vector

$$\underset{\sim}{H}(v) = (H(1,v),H(2,v),\ldots,H(n,v)).$$

It is shown in [4] that the set $\mathcal{V}$ of pure controls can be chosen so that the set

$$S = \{\underset{\sim}{H}(v) \mid v \in \mathcal{V}\}$$

is a compact set in R^n.

It is now clear that our game is equivalent to an S-game as formulated in Section 1. Namely, player P_2 chooses a point s of a compact set S in R^n while player P_1 chooses a coordinate i of R^n. The payoff to P_1 is the i^{th} coordinate s_i of s, that is, $H(i,v)$.

Let $\lambda = (\lambda_1 \ldots \lambda_n)$ be a probability vector, and define for every $v \in \mathcal{V}$

$$H(\lambda,v) = \int_0^1 \left(\sum_{i=1}^{n} \lambda_i h(t,x(t),i,v(t)) \right) dt. \tag{12}$$

The algorithm of Section 2 is now applicable to the differential S-game and will be as powerful as our ability to solve the following optimal control problem in pure controls:

$$\underset{v \in \mathcal{V}}{\text{minimize}}\ H(\lambda,v)$$

$$\text{subject to:}\quad \frac{dx}{dt} = f(t,x(t),v(t)), \quad x(0) = x_0. \tag{13}$$

4. A Simple Example

We shall now illustrate the working of our algorithm. The example solved below shows that this algorithm can sometimes yield approximate optimal solutions even in games which do not precisely fit the theory of Sections 2 and 3, provided that the conceptual ingredients of an S-game are present.

The law of motion is given by

$$\frac{dx}{dt} = v \quad , \quad x(0) = 1. \tag{14}$$

Player P_1 has only two choices, 1 or 2 and

$$h(t,x,i,v) = \begin{cases} 10v^2 \text{ if } i = 1 \\ x^2 + v^2 \text{ if } i = 2 \end{cases}$$

The expression (12) of Section 3 now becomes

$$H(\lambda,v) = \int_0^1 [\lambda_1 10v^2(t) + \lambda_2(x^2(t)+v^2(t))]dt. \tag{15}$$

Since in this example $H(\lambda,v)$ is convex in v and linear in λ it follows from Sion's Theorem [7] that we can take $\mathcal{V}$ to be the whole of L_2 [0,1] and be assured that the value exists. In order to apply our algorithm the optimal control problem (13) has to be solvable. If the minimization in (13) is taken over the piecewise continuous functions instead of $\mathcal{V}$ then this problem can be solved by a standard application of the Maximum Principle (see Rozonoer [6] pp. 1291). However, since piecewise continuous functions are dense in L_2 [0,1], the latter solution is also optimal for (13). That solution is of the following form:

For any $\lambda = (\lambda_1,\lambda_2)$ define the constants $\gamma = \left(\frac{\lambda_2}{1+9\lambda_1}\right)^{\frac{1}{2}}$,

$A_1 = \frac{1}{e^{2\gamma}+1}$, $A_2 = 1 - A_1$. The optimal trajectory for a particular λ (with $\lambda_2 > 0$) is

$$\bar{x}(t) = A_1 e^{\gamma t} + A_2 e^{-\gamma t}, \tag{16}$$

while the optimal control is given by

$$\bar{v}(t) = \gamma(A_1 e^{\gamma t} - A_2 e^{-\gamma t}). \tag{17}$$

If $\lambda_2 = 0$, quite clearly $\bar{v}(t) \equiv 0$.

We shall now summarize the numerical results of the first five iterations:

$$A^5 = \begin{bmatrix} 0 & 1.7080 & .4268 & .9597 & .6659 & .8055 \\ 1 & .7616 & .8218 & .7765 & .7952 & .7850 \end{bmatrix}$$

The optimal strategies for players P_1 and P_2 in this matrix game are $\lambda^5 = (.0681, .9319)$, $\eta^5 = (0,0,0,0, .1368, .8632)$.

Let $\underline{\theta}_k = H(\lambda^k, v^{k+1}) = \min_v H(\lambda^k, v)$ and $\bar{\theta}_k = v(A^k)$ for $k = 1,2,\ldots$.

In this example the lower value $\underline{\theta}_k$ and the upper value $\bar{\theta}_k$ are given by

$\bar{\theta}_k$:	.8775	.8037	.7907	.7875	.7864
$\underline{\theta}_k$:	.7616	.7729	.7847	.7852	.7862

Thus after five iterations we know that $v(\Gamma)$ lies in the interval [.7862, .7864] (also see Remark 2.4).

The vector η^5 may be regarded as the "nearly" optimal set of weights which player II can assign to the six pure controls corresponding to the columns of A^5. The exact expression for the fifth and the sixth of these is obtained by substituting the appropriate values of γ, A_1 and A_2 in (17).

Acknowledgement

We wish to thank A.J. Goldman from the Johns Hopkins University and D.J. Wilson from the University of Melbourne for reading an earlier version of this manuscript and for their helpful comments.

References

[1] L.D. Berkovitz, Optimal Control Theory, Springer-Verlag, New York, 1974.

[2] D. Blackwell and M.A. Girshick, Theory of Games and Statistical Decisions, Wiley, New York, 1954.

[3] R.J. Elliott and N.J. Kalton, The Existence of Value in Differential Games, Memoirs Amer. Math. Soc. 126, (1972-73).

[4] J.A. Filar, and T.E.S. Raghavan, An Algorithm for Solving S-Games and Differential S-Games, SIAM J. Control, 1982, to appear.

[5] R. Issacs, Differential Games, Wiley, New York, London, 1965.

[6] L.I. Rozonoer, L.S. Pontryagin, Maximum Principle in the Theory of Optimum Systems I, Automation and Remote Control 20, #10, 1959, pp. 1288-1302.

[7] M. Sion, On General Minimax Theorem, Pacific J. Math. 8, 1958, pp. 171-176.

[8] M.D. Troutt, Algorithms for non-convex S_n-games, Math. Programming, 14, 1978, pp. 332-348.