

REPEATED TWO-PERSON GAMES OVER LARGE TIME INTERVALS

A. Yu. Levin

1970

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-197001.82856>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

UDC 518.9

MATHEMATICS

A. Yu. LEVIN

REPEATED TWO-PERSON GAMES OVER LARGE TIME INTERVALS

(Presented by Academician A. N. Kolmogorov, 21 X 1969)

The theory of antagonistic two-person games after Borel and von Neumann has developed and been generalized in various directions; on this, see, for example, ⁽¹⁾. The approach proposed below has apparently not been discussed, although it seems quite natural. Its essence, briefly speaking, consists in taking the time factor into account.

Consider an encounter of two players, during which they play between themselves a number of rounds of one and the same zero-sum matrix game. Each round lasts a positive time; along with the payoff matrix $A = \|a_{ij}\|$, there is also given a matrix $T = \|t_{ij}\|$, where t_{ij} is the mean duration of a round when players I and II choose, respectively, the i -th and j -th pure strategies. We shall suppose (although this is possible and not very essential) that the corresponding variances σ_{ij}^2 of the durations of the rounds are finite. The total time of the encounter is either fixed in advance or, more generally, is a random variable independent of the course of the game. (Thus the number of rounds in the encounter is also a random variable, which, however, depends on the course of the game.) We shall be interested in the case where the mathematical expectation t of the duration of the encounter is large in comparison with all $t_{ij}\sigma_{ij}$. This will make it possible to ignore the "boundary effect" of the last round, which—if it has not ended by the moment the encounter ends—may, depending on the rules, either be annulled, or be played to the end, or lead to some partial payoff. The boundary effect greatly complicates the finding of exact optimal strategies; however, if t is regarded as large and the matter is approached from an asymptotic point of view, the contribution of a single round may be neglected.

It is assumed that both players, as usual, seek to maximize the guaranteed value of the mathematical expectation of their payoff over the encounter. What strategies should they use?

It is clear that the standard von Neumann optimal strategies for the game with matrix A are, generally speaking, unsuitable here, since they do not take the time factor into account: small frequent winnings may turn out to be more advantageous than large but rare ones. There are two basic cases in which the

strategies optimal in the game with matrix A remain optimal (more precisely, asymptotically optimal; see below) here as well: a) when the game with matrix A is harmless; b) when all t_{ij} are equal. These are important, but nevertheless particular, cases; in many real game situations the duration of the rounds varies greatly depending on the strategies chosen by the players.

It may seem that the problem reduces to solving the game with matrix $H = \|h_{ij}\|$, where $h_{ij} = a_{ij}/t_{ij}$ are the payoffs “per unit time.” This is only partly true. Namely, it is easy to see that if the matrix H has a saddle point h_{kl} , then the k -th and l -th pure strategies of players I and II respectively are asymptotically optimal; in this case I wins, in one encounter, on average $\sim h_{kl}t$. But in the most interesting case, when H has no saddle point, the optimal mixed strategies of the game with matrix H are not, generally speaking, asymptotically optimal for the game under consideration.

The method for determining asymptotically optimal strategies in the general case is based on solving the following scalar equation:

$$(\text{value of the game with matrix } A - \lambda T) = 0. \quad (1)$$

Since all $t_{ij} > 0$, the left-hand side of (1) is strictly decreasing in λ , and equation (1), obviously, has a unique root λ_0 . Strategies optimal in the sense of Neumann for the game with matrix $A - \lambda_0 T$ are asymptotically optimal for the game described above; moreover, the average payoff of player I per encounter is equal to $\lambda_0 t + O(1)$, where $O(1)$ is majorized by a quantity independent of t .

The term asymptotically optimal means that the use of these strategies leads to an average payoff per encounter which differs by no more than $O(1)$ from the average payoff under strictly optimal strategies; here $O(1)$ again depends only on t . Also, in a separately taken encounter the probability of the inequality

$$\left| \frac{\text{payoff of I}}{t} - \lambda_0 \right| > \varepsilon \quad (\varepsilon > 0 \text{ arbitrary})$$

under optimal behavior of the players tends to zero as t grows (for fixed distributions of payoffs and durations of the games). We note that strictly optimal strategies are not, generally speaking, homogeneous in time; to find them (which apparently is a difficult problem) it is necessary to know the distribution functions of the duration of the encounter and of the individual games, as well as the rules for the last game.

Let us briefly outline the proof. Since the game with matrix $A - \lambda_0 T$ is harmless, according to the remark made above, the optimal strategies of this game are asymptotically optimal in an encounter with payoff matrix $A - \lambda_0 T$ and duration matrix T . In this case the average payoff of I in such an encounter is, obviously, $O(1)$ for arbitrarily large t . Now compare an encounter with matrices A, T and an encounter with matrices $A - \lambda_0 T, T$. For any particular choice by the

players of their mixed strategies, in the first case I will receive on average $\lambda_0 t$ (up to an accuracy of $O(1)$) more per encounter than in the second. Indeed, after each individual game, except for the unfinished one, I receive on average additionally $\lambda_0 t_{ij}$, where i, j are the numbers of the realized pure strategies; but the mathematical expectation of the sum of the t_{ij} over all games played during the encounter coincides with t , up to an accuracy of $O(1)$. Comparing these considerations (which admit a careful justification), we arrive at the required result.

Equation (1), in view of the monotonicity of the left-hand side, can be solved approximately by "bisection." The initial interval is determined from the matrix H according to the easily verified inequality

$$\max_i \min_j h_{ij} \leq \lambda_0 \leq \min_j \max_i h_{ij}.$$

If the optimal strategies in the game with matrix $A - \lambda_0 T$ are unique, then after a certain number of steps it will be possible to "feel out" the active pure strategies of the players. Denoting the corresponding submatrices of A and T by A_0, T_0 (in the case of uniqueness they are square), one can then find λ_0 as the root, lying in the obtained interval, of the algebraic equation $\det \|A - \lambda T\| = 0$.

With the exception of the last remark, what has been said extends also to infinite games. In this case, in the general case, one should naturally speak not of (asymptotically) optimal strategies, but of (asymptotically) ε -optimal strategies.

In conclusion we give a numerical example. Let

$$A = \begin{pmatrix} 1 & -2 & 4 \\ 0 & 3 & -2 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}.$$

If I uses in a match the strategy $(1/2; 1/2)$, which is optimal for him in the ordinary sense, i.e., without taking time into account, then, under II' s best response (the 2nd strategy), I wins on average $\sim 1/t$ per match ($t \gg 1$). In the game with matrix

$$H = \left\| \frac{a_{ij}}{t_{ij}} \right\| = \frac{1}{15} \begin{pmatrix} 15 & -15 & 20 \\ 0 & 9 & -5 \end{pmatrix},$$

the strategy $(2/7; 5/7)$ is optimal for I in the sense of Neumann; however, if I uses it, he will even lose—on average $\sim 1/18t$ per match (if II chooses the 3rd strategy). Solving equation (1), we find

$$\lambda_0 = \frac{1}{6}(45 - \sqrt{1929}) \approx 0.18.$$

Thus, under the asymptotically optimal strategies of I and II, which are approximately equal to $(0.47; 0.53)$ and $(0; 0.59; 0.41)$, the mathematical expectation of I's gain per match is $\sim 0.18t$.

The author is grateful to V. M. Granin and E. N. Sadovskii for the discussion.

Voronezh
State University

Received
20 X 1969

REFERENCES

¹ R. D. Luce, H. Raiffa, *Games and Decisions*, IL, 1961.

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.