U. S. AIR FORCE

PROJECT RAND

RESEARCH MEMORANDUM

ON A GEOMETRICAL GAME CONNECTED WITH SEQUENTIAL ANALYSIS

R. E. Bellman

RM-645

ASTIA Document Number ATI 210711

11 July 1951

Assigned to ______________________

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

ON A GEOMETRICAL GAME CONNECTED WITH SEQUENTIAL ANALYSIS

R. Bellman

Summary

We consider a game which has its origins in some strategic bombing models. The method of solution is one which may be applicable to a wider class of problems.

§1. Introduction

We wish to consider the following one-person -vs.-nature geometrical game:

"At any point (x, y) in the positive quadrant, a player has a choice of three moves which we shall designate by A, B or C. Associated with each move is the following set of probabilities:

(1) A: q_A = probability that the game terminates

p_{A1} = probability that the player receives x, if the game has not been terminated

p_{A2} = probability that the player receives 0

p_{A3} = probability that the player receives σx, if the game has not been terminated

B: q_B = probability that the game terminates

p_{B1} = probability that the player receives y

p_{B2} = probability that the player receives 0

p_{B3} = probability that the player receives σy, if the game has not been terminated

C: q_C = probability that the game terminates

p_{C1} = probability that the player receives x + y

p_{C2} = probability that the player receives 0

p_{C3} = probability that the player receives $\sigma(x + y)$, if the game has not been terminated

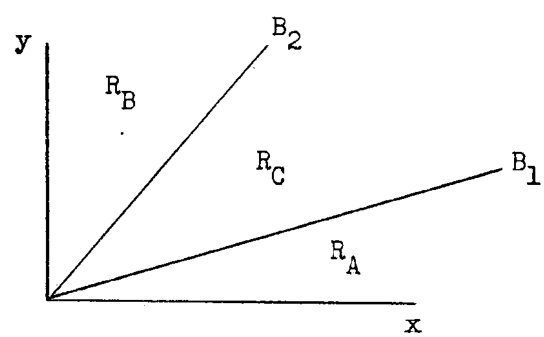
"Furthermore, in each case, if the game does not terminate, the x and y quantities he receives is subtracted from the original coordinates, and the game continues from the new point. Thus, if A is played, and the game continues, after having received 0, σx, or x, he continues from (x, y), $((1 - \sigma)x, y)$ or (0, y) respectively.

"The player's score is the sum of the amounts he receives prior to the termination of the game or to landing at (0, 0).

"The problem is to determine the course of play which maximizes the expected score."

There are two ways of presenting a solution. The first consists of describing the possibly infinite sequence of moves which yield this maximum expectation. The other consists merely in describing the best action to take at any given position. The second solution may be most aptly described by giving the location of the three point sets, R_A, R_B, R_C, which have the property that if $P \in R_A$, then A is the best first move, and similarly for the others. Intuitively, we would expect that these three sets be regions of the following form

(2)



,

since points near the x and y axes certainly belong to R_A and R_B, respectively. Actually, we shall show that this is true, where the boundary curves are straight lines.

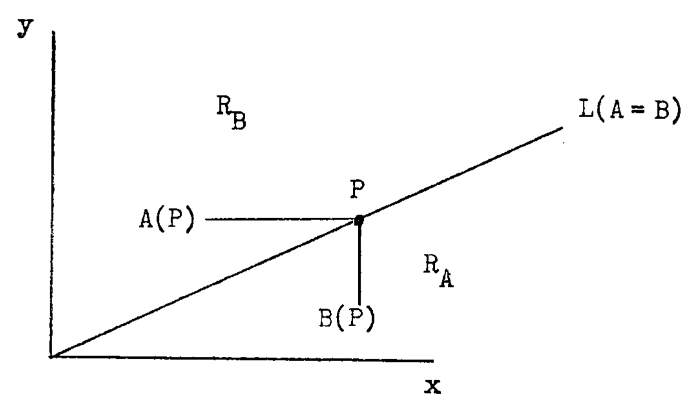
The game described above may be generalized to one in n dimensions with a correspondingly more complicated set of strategies. It would seem that the method we present here would extend, via an inductive argument, to show the existence of regions determining best first moves. However, the proof would be extremely long and detailed and we feel that another method would be preferable. Furthermore, our method seems to break down if we allow the three equal σ's above in the description of the game to become unequal. We would hypothesize that the result is still true - that three regions of the above simple type exist - but it seems quite difficult to prove.

By way of illustrating the type of argument we shall employ, we consider first a simple 2-move game where the solution is much simpler. This solution was obtained by M. Shiffman and the author.

§2. A 2-move game.*

To illustrate one of the methods we employ, let us consider the game where only the moves A and B are permissible. In this case, we would expect the positive (x, y) quadrant to be divided into two regions R_A and R_B as in the diagram below:

(1)



*The fact that f(x, y) is actually an attained maximum, and not merely an upper bound, may be proved à priori, or using the properties of the suspected solution. A complete discussion of the question of existence of the maximum in the general case has been given by H. N. Shapiro in a paper to appear shortly.

To determine the regions, it is sufficient to determine the boundary line L which consists of the points for which A or B are equally good first moves. If we let A(P) denote the point obtained upon using A, it is clear that $A(P) \in R_B$, and similarly $B(P) \in R_A$. Hence, if A is used at P, it is always followed by B, and if B is used it is always followed by A. The equation A = B implies, therefore, AB = BA. This last equation yields a simple result.

Let f(x, y) be the expectation attained using an optimal method of play. In order to simplify the exposition, let us assume that $p_{A1} = p_{A2} = 0$, $p_{B1} = p_{B2} = 0$. It is clear from the description of the game that any optimal procedure must have the property that after any number of moves have been made, the continuation from that point on is optimal. From this we see that f satisfies the functional equation

$$(2) \qquad f(x, y) = \text{Max} \begin{Bmatrix} p_1(\sigma x + f(\tau x, y), \\ p_2(\sigma y + f(x, \tau y), \end{Bmatrix}$$

(where $p_1 = (1 - q_A)p_{A3}$, $p_2 = (1 - q_B)p_{B3}$), $\tau = 1 - \sigma$.

At P if we employ A and then B, we have

$$(3) \qquad \begin{aligned} f(x, y) &= p_1(\sigma x + f(\tau x, y)) \\ &= p_1(\sigma x + p_2(\sigma y + f(\tau x, \tau y)) \\ &= \sigma(p_1 x + p_1 p_2 y) + p_1 p_2 f(\tau x, \tau y) \ . \end{aligned}$$

Similarly, B and then A yields

$$(4) \qquad f(x, y) = \sigma(p_2 y + p_1 p_2 x) + p_1 p_2 f(\tau x, \tau y) \ .$$

Equating the two results, we obtain as the equation of L:

(5) $$\frac{xp_1}{1 - p_1} = \frac{yp_2}{1 - p_2} .$$

The method above applies equally well if we allow different σ's for A and B. The line L is then given by

(6) $$\frac{\sigma_1 p_1}{1 - p_1} x = \frac{\sigma_2 p_2}{1 - p_2} y .$$

The procedure we have used yields very simply the equation of L, provided that the regions R_A and R_B are as pictured. It is now necessary to prove that R_A and R_B are actually regions where A and B are used. Since this will be done below, we will not give the proof here.

The method applies equally to the case where p_{A1}, p_{A2}, q_{A1}, q_{A2} are not taken to be zero.

§3. The 3-move game.

Again for the sake of simplicity, we shall assume that

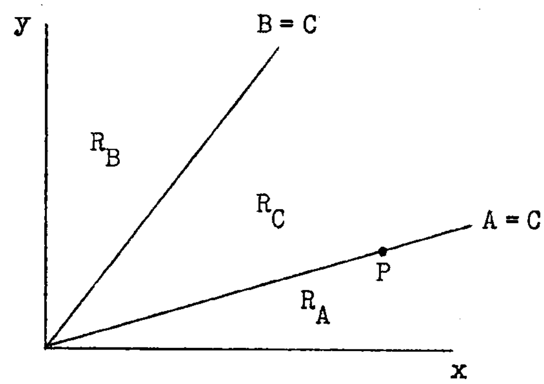
(1) $$p_{A1} = p_{A2} = 0, \quad p_{B1} = p_{B2} = 0, \quad p_{C1} = p_{C2} = 0 .$$

Furthermore, to avoid trivial cases, we assume that $p_{A3}, p_{B3} > p_{C3} > p_{A3}p_{B3}$. For typesetting purposes, set

(2) $$(1 - q_A)p_{A3} = p_1, \quad (1 - q_B)p_{B3} = p_2, \quad (1 - q_C)p_{C3} = p_{12} .$$

From what has been obtained in the 2-move game, together with what we expect intuitively, we are lead to hypothesize the following situation:

(3)



Considerably more difficulty is encountered when we attempt to determine the lines A = C, B = C. Consider the point P on A = C. If we use A, then A(P) may lie in R_B or R_C, depending upon the magnitude of σ. Hence A = C may be determined by AC = C or AB = C, and in a like manner, B = C may lead to BA = C or BC = C. It turns out, furthermore, that all four possibilities may occur. It would not seem to be difficult to examine separately each of those four possibilities and see in which regions of the $(p_1, p_2, p_{12}, \sigma)$ space each occurs. It turns out, however, that each of these cases again has sub-cases, so that a proof obtained in this manner would be quite long and tedious, and not very constructive. In place of this pedestrian procedure we employ a continuity method which permits us to observe how one case merges into another. This continuity method consists of fixing p_1, p_2 and p_{12} and allowing σ to vary between 0 and 1. At $\sigma = 1$, the regions R_A, R_B and R_C are easily obtained. As σ decreases from 1 to 0, it is easy to see how the regions vary.

§4. Some Preliminary Results:

In this section, we derive the expressions for f(x, y) at the points, assuming they exist, where AC, AB, BA, BC and C are each the best first or first two moves. The calculation is simplified by the circumstance that if C is ever a best first move then it must be repeated indefinitely. This follows from the fact that C applied to (x, y) lands one at the point $(\tau x, \tau y)$. Clearly the best move at $(\tau x, \tau y)$ is the same as the best move at (x, y). In a similar way we see that if AB or BA are ever the best first two moves, then each sequence is repeated periodically. Using these remarks, we can now readily compute the required maximum expectations.

C: $$f(x, y) = p_{12}\left(\sigma(x + y) + f(\tau x, \tau y)\right) = \frac{p_{12}\sigma(x + y)}{1 - \tau p_{12}} \ .$$

AC: $$f(x, y) = p_1\left(\sigma x + \frac{p_{12}\sigma(\tau x + y)}{1 - \tau p_{12}}\right) = \frac{p_1\sigma(x + p_{12}y)}{1 - \tau p_{12}} \ .$$

AB: $$f(x, y) = p_1\left[\sigma x + p_2\left(\sigma y + f(\tau x, \tau y)\right)\right] = \frac{p_1\sigma x + p_1p_2\sigma y}{1 - \tau p_1p_2} \ .$$

BA: $$f(x, y) = \frac{p_2\sigma y + p_1p_2\sigma x}{1 - \tau p_1p_2}$$

BC: $$f(x, y) = \frac{p_2(\sigma y + \sigma p_{12}x)}{1 - \tau p_{12}}$$

Let us now compute the equations of some important lines:

AC = C: $$(p_1 - p_{12})x = p_{12}(1 - p_1)y$$

BC = C: $$p_{12}(1 - p_2)x = (p_2 - p_{12})y$$

Notice the important and useful fact that these lines are independent of σ.

AB = C: $$x\left(\frac{p_1}{1-\tau p_1p_2} - \frac{p_{12}}{1-\tau p_{12}}\right) = y\left(\frac{p_{12}}{1-\tau p_{12}} - \frac{p_1p_2}{1-p_1p_2\tau}\right)$$

This line is only of interest to us if

(1) $$\frac{p_1}{1-\tau p_1p_2} > \frac{p_{12}}{1-\tau p_{12}} .$$

This inequality holds at $\tau = 0$.

BA = C: $$x\left(\frac{p_{12}}{1-\tau p_{12}} - \frac{p_1p_2}{1-\tau p_1p_2}\right) = y\left(\frac{p_2}{1-\tau p_1p_2} - \frac{p_{12}}{1-\tau p_{12}}\right) .$$

Again this line is important at $\tau = 0$, and is of interest for other values of τ only if

(2) $$\frac{p_2}{1-\tau p_1p_2} > \frac{p_{12}}{1-\tau p_{12}} .$$

Writing these lines in the simpler forms:

BA = C: $$x(p_{12} - p_1p_2) = y\left[p_2 - p_{12} + \tau(p_{12}p_1p_2 - p_2p_{12})\right]$$

AB = C: $$x(p_1 - p_{12} + \tau(p_{12}p_1p_2 - p_1p_{12})) = y(p_{12} - p_1p_2) .$$

Since $p_{12}p_1p_2 - p_2p_{12} < 0$, $p_{12}p_1p_2 - p_1p_{12} < 0$, as τ increases from 0 to 1, AB = C rotates towards $\pi_2 = 0$, and BA = C towards $\pi_1 = 0$.

At $\tau = 0$, we have

AB = C: $$x(p_1 - p_{12}) = y(p_{12} - p_1p_2)$$

BA = C: $$x(p_{12} - p_1p_2) = y(p_2 - p_{12}) .$$

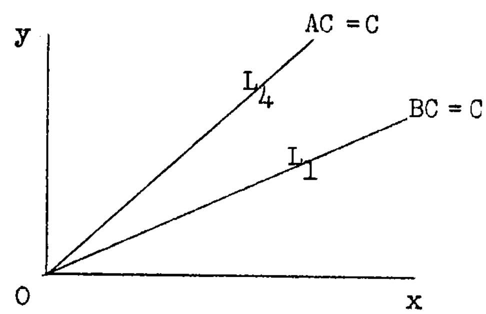
Comparing AB = C with AC = C, we see that at $\tau = 0$, AC = C lies below AB = C, and BC = C lies above BA = C.

In the next section, we apply these results to problems of determining the regions.

§5. Determination of the region R_A, R_B and R_C.

There are two cases which confront us at once. Either AC = C is above or below BC = C. Let us show that if AC = C is above BC = C, then C is never used.

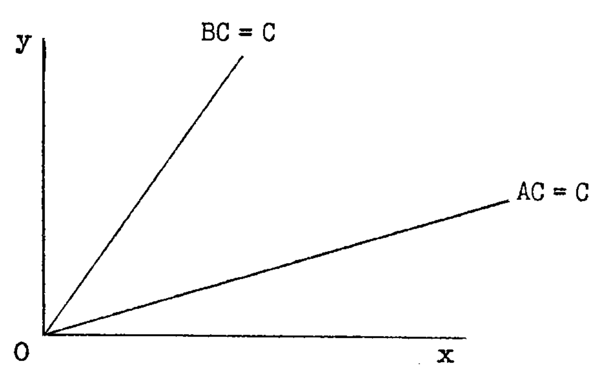
(1)



In the region between L_4 and Ox, AC superior to C, which implies that C is never a best first move. Similarly in the region bounded by Oy and L_1, C is never a best first move. Hence, in this case, C is never a best first move, which means that it is never used.

Let us consider, then, the remaining case, AC = C below BC = C.

(2)



We know from the discussion of the previous section that AB = C is above AC = C and BA = C is below BC = C. However, their relative position depends upon p_1, p_2, p_{12} and τ. At $\tau = 0$, the slope of BA = C is

$$BA = C: \qquad \frac{y}{x} = \frac{p_{12} - p_1 p_2}{p_2 - p_{12}}$$

and that of AB = C is

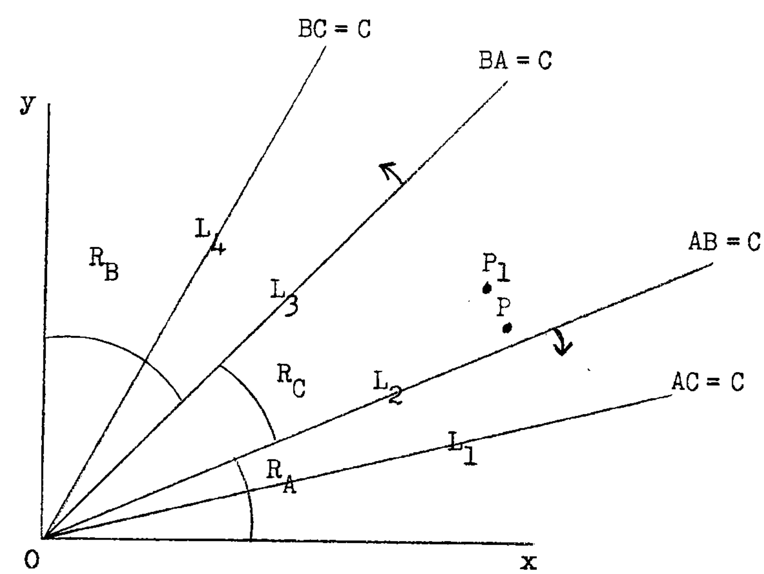
$$AB = C: \qquad \frac{y}{x} = \frac{p_1 - p_{12}}{p_{12} - p_1 p_2} .$$

From these slopes we see readily that either case may occur at $\tau = 0$, BA = C above AB = C or conversely.

Let us assume that AB = C is above BA = C to begin with. Then the argument used above shows that C is never used. If C is never used, we are reduced to the two-move game where the boundary is determined by AB = BA. Furthermore, as long as this condition persists, C is never used.

The only difficult case is, then, that where BA = C is above AB = C at $\tau = 0$ or for some τ. Let us assume that this occurs at $\tau = 0$. The diagram is then

(3)



At $\tau = 0$, the R_A, R_B and R_C regions are as marked, since A is always followed by B, B by A, and C by itself. Consequently, we need only compare AB with C and

BA with C, and we have assumed that BA = C is above AB = C.

Let $\tau > 0$ but not so large that L_2 has crossed L_1 or L_3 the line L_4. Incidentally, we note that for some values of p_1, p_2, p_{12}, L_2 may never overtake L_1, nor L_3, L_4. We then assert that R_A, R_B and R_C are as in (3).

The proof that there is a region where A is always used first is very simple. Use of A repeatedly yields

$$A^{\infty}: \qquad f \geq \frac{\sigma x p_1}{1 - \tau p_1} .$$

If B were the best first move we should have

$$(5) \qquad \begin{aligned} f(x, y) &= p_2(\sigma y + f(x, \tau y)) \\ &\leq p_2(\sigma y + f(x, y)) \\ &\leq \frac{p_2 \sigma y}{1 - p_2} . \end{aligned}$$

For

$$(6) \qquad \frac{\sigma x p_1}{1 - \tau p_1} > \frac{p_2 \sigma y}{1 - p_2} ,$$

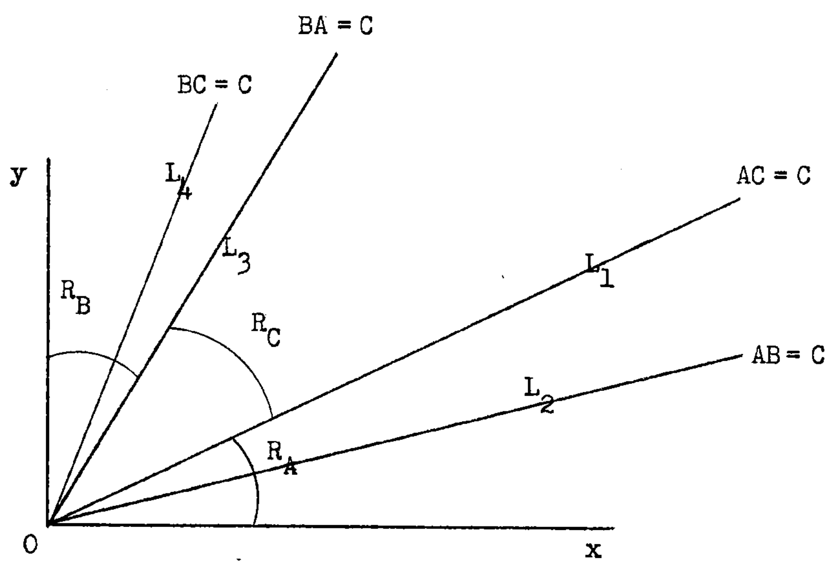
A is a better first move than B.

The proof that OL_2x is an R_A region now proceeds as follows. If the region bounded by L_5: $p_2 y / 1 - p_2 \leq x p_1 / 1 - \tau p_1$ includes OL_2x, we are finished. If not, let P be a point slightly above L_5 and below L_2. We know that C is not used, so that it is only a matter of comparing A and B. If B is used, then B(P) will be below L_5 and hence B will imply BA. Since BA < C < AB, we see that B is not used first at P. This allows us to enlarge the known R_A region up to the line L_2.

We now wish to show that the middle region is R_C. Consider the point P in OL_3L_2. If B is used at P, then BA is used at P, provided that P is chosen close enough to L_2. Since $BA < C$ in OL_3L_2, it follows that B is never used first at P. Rotate the point P slightly to P_1. If B is used first at P_1, then if $B(P_1)$ falls in R_A, B is followed by A. If $B(P_1)$ falls near P, then B is followed by A or C, where at the moment we do not know which. Therefore at P_1 either BA or BC is used. But $BA < C$ and $BC < C$. Rotating the point P further and further, we see that B is never used first in OL_3L_2. Similarly, A is never used first. Hence $OL_3L_2 = R_C$.

This proof is valid until one or the other of the lines L_2 or L_3 crosses an L_1 or L_4 respectively. Assume that L_2 crosses first. Then the diagram would be

(4)



Let us show that OL_1x is now R_A. Clearly C is never used here since $C < AC$. Hence only B or A are possible. We show as above that B is never used. Hence $OL_1x = R_A$, and similarly $OL_3y = R_B$.

Now for the middle region. Again the same proof as above holds. It is clear that the method of proof works equally well regardless of the portion of the lines.

RB:rgb