4.7 Approximating the Solution: Fictitious Play.

The method of fictitious play may be used to approximate the value and optimal strategies of a finite game. It is a sequential procedure that approximates the value as closely as desired, giving upper and lower bounds that converge to the value and strategies that achieve these bounds. Its disadvantage is its slow rate of convergence.

Let $A(i,j)$ be an m by n payoff matrix. The method starts with an arbitrary initial pure strategy $1 \le i_1 \le m$ for Player I. Alternatively from then on, each player chooses his next pure strategy as a best reply assuming the other player chooses among his previous choices at random equally likely. For example, if $i_1, \ldots, i_k$ have already been chosen for some $k \ge 1$, then j_k is chosen as that j that minimizes the expectation $(1/k)\sum_{\ell=1}^{k} A(i_\ell, j)$. Similarly, if $j_1, \ldots, j_k$ have already been chosen, i_{k+1} is then chosen as that i that maximizes the expectation $(1/k)\sum_{\ell=1}^{k} A(i, j_\ell)$. To be specific, we define

$$s_k(j) = \sum_{\ell=1}^{k} A(i_\ell, j) \qquad \text{and} \qquad t_k(i) = \sum_{\ell=1}^{k} A(i, j_\ell) \tag{1}$$

and then choose

$$j_k = \text{argmin } s_k(j) \qquad \text{and} \qquad i_{k+1} = \text{argmax } t_k(i) \tag{2}$$

If the maximum of $t_k(i)$ is assumed at several different values of i, then it does not matter which of these is taken as i_{k+1}. To be specific, we choose i_{k+1} as the smallest value of i that maximizes $t_k(i)$. Similarly j_k is taken as the smallest j that minimizes $s_k(j)$. In this way, the sequences i_k and j_k are defined deterministically once i_1 is given.

Notice that $\overline{V}_k = (1/k)t_k(i_{k+1})$ is an upper bound to the value of the game since Player II can use the strategy that chooses j randomly and equally likely from $j_1, \ldots, j_k$ and keep Player I's expected return to be at most $\overline{V}_k$. Similarly, $\underline{V}_k = (1/k)s_k(j_k)$ is a lower bound to the value of the game. It is rather surprising that these upper and lower bounds to the value converge to the value of the game as k tends to infinity.

Theorem. *If V denotes the value of the game, then $\underline{V}_k \to V$, $\overline{V}_k \to V$, and $\underline{V}_k \le V \le \overline{V}_k$, for all k.*

This approximation method was suggested by George Brown (1951), and the proof of convergence was provided by Julia Robinson (1951). The convergence of $\underline{V}_k$ and $\overline{V}_k$ to V is slow. It is thought to be of order $1/k$. In addition, the convergence is not monotone. See the example below.

A modification of this method in which the selection of the i_k and j_k is made simultaneously rather than sequentially, is often used, but it is not as fast.

It should be mentioned that as a practical matter, choosing at each stage a best reply to an opponent's imagined strategy of choosing among his previous choices at random is not a good idea. See Exercise 7. On the other hand, Alfredo Baños (1968) describes a sequential method for Player I, say, to choose mixed strategies such that liminf of the average payoff is at least the value of the game no matter what player II does. This choice of mixed strategies is based only upon Player I's past pure strategy choices and the past observed payoffs, but not otherwise on the payoff matrix or upon the opponent's pure strategy choices. It would be nice to devise a practical method of choosing a mixed strategy depending on all the information contained in the previous plays of the game that performs well.

EXAMPLE. Take as an example the game with matrix

$$A = \begin{pmatrix} 2 & -1 & 6 \\ 0 & 1 & -1 \\ -2 & 2 & 1 \end{pmatrix}$$

This is the game solved in Section 4.6. It has value .5, and optimal mixed strategies, $(.25, .75, 0)$ and $(.5, .5, 0)$ for Player I and Player II respectively. It is easy to set up a program to perform the calculations. In particular, the computations, (1), may be made recursively in the simpler form

$$s_k(j) = s_{k-1}(j) + A(i_k, j) \qquad \text{and} \qquad t_k(i) = t_{k-1}(i) + A(i, j_k) \tag{3}$$

We take the initial $i_1 = 1$, and find

k	i_k	$s_k(1)$	$s_k(2)$	$s_k(3)$	$\underline{V}_k$	j_k	$t_k(1)$	$t_k(2)$	$t_k(3)$	$\overline{V}_k$
1	1	2	$-\mathbf{1}$	6	−1	2	−1	1	**2**	2
2	3	**0**	1	7	0	1	**1**	1	0	0.5
3	1	2	**0**	13	0	2	0	**2**	2	0.6667
4	2	2	**1**	12	0.25	2	−1	3	**4**	1
5	3	**0**	3	13	0	1	1	**3**	2	0.6
6	2	**0**	4	12	0	1	**3**	3	0	0.5
7	1	**2**	3	18	0.2857	1	**5**	3	−2	0.7143
8	1	4	**2**	24	0.25	2	**4**	4	0	0.5
9	1	6	**1**	30	0.1111	2	3	**5**	2	0.5556
10	2	6	**2**	29	0.2	2	2	**6**	4	0.6
11	2	6	**3**	28	0.2727	2	1	**7**	6	0.6364
12	2	6	**4**	27	0.3333	2	0	**8**	8	0.6667
13	2	6	**5**	26	0.3846	2	−1	9	**10**	0.7692
14	3	**4**	7	27	0.2857	1	1	**9**	8	0.6429
15	2	**4**	8	26	0.2667	1	3	**9**	6	0.6

The initial choice of $i_1 = 1$ gives $(s_1(1), s_1(2), s_1(3))$ as the first row of A, which has a minimum at $s_1(2)$. Therefore, $j_1 = 2$. The second column of A has $t_1(3)$ as the maximum, so $i_2 = 3$. Then the third row of A is added to the s_1 to produce the s_2 and so on. The minimums of the s_k and the maximums of the t_k are indicated in boldface. The largest of the $\underline{V}_k$ found so far occurs at $k = 13$ and has value $s_k(j_k)/k = 5/13 = 0.3846\ldots$. This value can be guaranteed to Player I by using the mixed strategy (5/13,6/13,2/13), since in the first 13 of the i_k there are 5 2's, 6 2's and 2 3's. The smallest of the $\overline{V}_k$ occurs several times and has value .5. It can be achieved by Player II using the first and second columns equally likely. So far we know that $.3846 \le V \le .5$, although we know from Section 4.6 that $V = .5$.

Computing further, we can find that $\underline{V}_{91} = 44/91 = .4835\ldots$ and is achieved by the mixed strategy $(25/91, 63/91, 3/91)$. From row 9 on, the difference between the boldface numbers in each row seems to be bounded between 4 and 6. This implies that the convergence is of order $1/k$.

Exercise 4.6. Carry out the fictitious play algorithm on the matrix $A = \begin{pmatrix} 1 & -1 \\ 0 & 2 \end{pmatrix}$ through step $k = 4$. Find the upper and lower bounds on the value of the game that this gives.

Exercise 4.7. Suppose the game with matrix, $\begin{pmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{pmatrix}$ is played repeatedly. On the first round the players make any choices.

(a) Thereafter Player I makes a best response to his opponent's imagined strategy of choosing among her previous choice at random. If Player II knows this, what should she do? What are the limiting average frequencies of the choices of the players?

(b) Suppose Player II is required to play a best response to her opponent's previous choices. What should Player I do, and what would his limiting average payoff be?

References.

Alfredo Baños (1968) "On Pseudo Games", *Ann. Math. Statist.* **39**, 1932-1945.

George W. Brown (1951) "Iterative solution of games by fictitious play", in *Activity Analysis of Production and Allocation,* T. C. Koopmans ed., John Wiley, New York, 374-376.

Julia Robinson (1951) "An iterative method of solving a game", *Ann. Math.* **54**, 296-301

Solutions.

4.6. We take the initial $i_1 = 1$, and find

k	i_k	$s_k(1)$	$s_k(2)$	$\underline{V}_k$	j_k	$t_k(1)$	$t_k(2)$	$\overline{V}_k$
1	1	1	$\mathbf{-1}$	-1	2	-1	**2**	2
2	2	**1**	1	.5	1	0	**2**	1
3	2	**1**	3	.3333	1	1	**2**	0.6667
4	2	**1**	5	0.25	1	**2**	2	.5

The largest of the $\underline{V}_k$, namely .5, is equal to the smallest of the $\overline{V}_k$. So the value of the game is 1/2. An optimal strategy for Player I is found at $k = 2$ to be $\boldsymbol{p} = (.5, .5)$. An optimal strategy for Player II is $\boldsymbol{q} = (.75, .25)$ found at $k = 4$. Don't expect to find the value of a game by this method again!

4.7. (a) The upper left payoff of $\sqrt{2}$ was chosen so that there would be no ties in the fictitious play. So Player II knows exactly what Player I will do and will be able to guarantee a zero payoff at each future stage. If Player II's relative frequency, q_k, of column 1 by stage k goes above $1/(\sqrt{2}+1)$, Player I will play row 1, causing Player II to play column 2, thus causing q_k to decrease. Thus q_k converges to $1/(\sqrt{2}+1)$, which in fact is Player II's optimal strategy for the game. Similarly, Player I's relative frequency, p_k, of row 1 converges to $1/(\sqrt{2}+1)$, which is his optimal strategy.

(b) Player I should play the same pure strategy as his opponent at each stage, gaining either $\sqrt{2}$ or 1 at each stage. The same argument as in (a) shows that Player II's average relative frequency of column 2 converges to $1/(\sqrt{2}+1)$, so Player I's limiting average payoff is

$$\frac{1}{\sqrt{2}+1} \cdot \sqrt{2} + \frac{\sqrt{2}}{\sqrt{2}+1} = 2(2-\sqrt{2}),$$

twice the value of the game.