An Evolutionary Duel Tournament

Ath. Kehagias

1 Intro

We study an evolutionary duel tournament which consists of the following components

1. The *basic duel game* or *match*, which is played between two players. We study two variants: the *discounted* and *undiscounted* ones.

2. The *meeting*, which consists of all possible matches between N players.

3. The evolutionary *tournament*, which consists of one meeting per *generation* of players. The outcomes of each meeting are used to update the strategies used by the players in the meeting of next generation.

Our main goal is the following: there exist M strategies and each of N players will use one of these strategies. The representation of strategies in the next generation depends on their performance in the current generation. More precisely, the performance of a strategy is proprtional to the total payoff accumulated by all the players who have used this strategy. It is hoped that, as time progresses, better performing strategies will be more heavily represented and will eventually "exterminate" strategies with worse performance.

2 The Discounted Basic Game

2.1 Rules of the Game

The discounted base game is played between players P_1 and P_2; player P_n has *kill probability* p_n. In each round the two players shoot simultaneously, with P_n's shooting probability x_n being determined by the *strategy* he has adopted (a player can also choose to *abstain*, i.e., to not shoot). The game state is $\mathbf{s} = s_1 s_2$, where $s_n = 0$ (resp. $s_n = 1$) means P_n is dead (resp. alive). The state set is

$$S = \{00, 01, 10, 11\}$$

and each state $\mathbf{s}$ yields *round payoff* $q_n(\mathbf{s})$ to P_n. We have

$$\begin{aligned} &q_1(00) = 0, \quad q_1(01) = 0, \quad q_1(10) = 1, \quad q_1(11) = 1, \\ &q_2(00) = 0, \quad q_2(01) = 1, \quad q_2(10) = 0, \quad q_1(11) = 1. \end{aligned}$$

The idea is that each player receives one payoff unit for each turn in which he is alive. The *total payoff* is

$$Q_n(\mathbf{s}(0)\mathbf{s}(1)\mathbf{s}(2)\ldots) = \sum_{t=0}^{\infty} \gamma^t q_n(\mathbf{s}(t))$$

(i.e., it round payoff is discounted by a factor $\gamma \in (0,1)$). The *expected total payoff* is

$$\overline{Q}_n = \mathbb{E}(Q_n(\mathbf{s}(0)\mathbf{s}(1)\mathbf{s}(2)\ldots))$$

where the expectation is taken with respect to the probability determined by the players' strategies.

2.2 Payoff Equation under Stationary Strategies

Suppose that P_n uses a stationary strategy $\sigma(\mathbf{s}) = (x_{n1}, x_{n2})$, where

$$x_{nm} = \Pr(P_n \text{ shoots at } P_m)$$

with "P_n shoots at P_n" being understood as "P_n abstains". Since we must have $x_{n1}+x_{n2} = 1$, we can parametrize P_n's stationary strategy by a single number x_n, the probability of shooting at the other player. We define

$$V_{\mathbf{s}}^n(x_1, x_2) = \text{"}P_n\text{'s expected payoff when starting at state } \mathbf{s} \text{ and shoooting probs are } x_1, x_2\text{"}.$$

We have $\sum_{t=0}^{\infty} \gamma^t = -\frac{1}{\gamma-1}$

$$V_{00}^1 = 0, \quad V_{01}^1 = 0, \quad V_{10}^1 = \sum_{t=0}^{\infty} \gamma^t \cdot 1 = \frac{1}{1-\gamma}.$$

For V_{11}^1 we have the following equation

$$\begin{aligned} V_{11}^1 &= x_1 p_1 x_2 p_2 V_{00}^1 + (x_1(1-p_1) + (1-x_1)) x_2 p_2 V_{01}^1 + x_1 p_1 (x_2(1-p_2) + (1-x_2)) V_{10}^1 \\ &\quad + (x_1(1-p_1) + (1-x_1))(x_2(1-p_2) + (1-x_2))\left(1 + \gamma V_{11}^1\right) \end{aligned}$$

which becomes

$$V_{11}^1 = x_1 p_1 (x_2(1-p_2) + (1-x_2)) \frac{1}{1-\gamma} + (x_1(1-p_1) + (1-x_1))(x_2(1-p_2) + (1-x_2))\left(1 + \gamma V_{11}^1\right)$$

Solving, when $\max(x_1, x_2) > 0$ we get

$$V_{11}^1(x_1, x_2) = \frac{(1 - x_2 p_2)(1 - \gamma(1 - x_1 p_1))}{(1-\gamma)(1 - \gamma(1 - x_1 p_1)(1 - x_2 p_2))}$$

and when $\max(x_1, x_2) = 0$

$$V_{11}^1(x_1, x_2) = \frac{1}{1-\gamma}.$$

Repeating the process for V_{11}^2 we get

$$V_{00}^2 = 0, \quad V_{01}^2 = 0, \quad V_{10}^2 = \frac{1}{1-\gamma}$$

$$V_{11}^2(x_1, x_2) = \frac{(1 - x_1 p_1)(1 - \gamma(1 - x_2 p_2))}{(1-\gamma)(1 - \gamma(1 - x_1 p_1)(1 - x_2 p_2))}$$

2.3 Stationary Equilibria

We will now show that the game has two Nash equilibria in stationary strategies.

1. When $\max(x_1, x_2) > 1$:

$$\begin{aligned} \frac{dV_{11}^1}{dx_1} &= D_{x_1}\left(\frac{(1 - x_2 p_2)(1 - \gamma(1 - x_1 p_1))}{(1-\gamma)(1 - \gamma(1 - x_1 p_1)(1 - x_2 p_2))}\right) = \frac{(1 - x_2 p_2)\gamma p_1 p_2 x_2}{(1-\gamma)(1 - \gamma(1 - x_2 p_2)(1 - x_1 p_1))^2} > 0 \\ \frac{dV_{11}^2}{dx_2} &= D_{x_2}\left(\frac{(1 - x_1 p_1)(1 - \gamma(1 - x_2 p_2))}{(1-\gamma)(1 - \gamma(1 - x_1 p_1)(1 - x_2 p_2))}\right) = \frac{(1 - x_1 p_1)\gamma p_1 p_2 x_1}{(1-\gamma)(1 - \gamma(1 - x_2 p_2)(1 - x_1 p_1))^2} > 0 \end{aligned}$$

   Hence, whenever $x_{-n} > 0$, P_n has motive to increase x_n as much as possible. Consequently $(\widehat{x}_1, \widehat{x}_2) = (1, 1)$ is a NE.

2. When $(\widehat{x}_1, \widehat{x}_2) = (0, 0)$, we have $\frac{(1-\gamma)}{(1-\gamma)(1-\gamma)} - \frac{(1-\gamma(1-x_1p_1))}{(1-\gamma)(1-\gamma(1-x_1p_1))}$

$$\begin{aligned} V_{11}^1(0,0) - V_{11}^1(x_1, 0) &= \frac{(1-\gamma)}{(1-\gamma)(1-\gamma)} - \frac{(1 - \gamma(1 - x_1 p_1))}{(1-\gamma)(1 - \gamma(1 - x_1 p_1))} = 0 \\ V_{11}^2(0,0) - V_{11}^2(0, x_2) &= 0 \end{aligned}$$

   Hence, whenever P_n has no motive to move away from $x_n = 0$. Consequently $(\widehat{x}_1, \widehat{x}_2) = (0, 0)$ is a NE.

2.4 Nonstationary Equilibria

Let σ_G be the no-shoot strategy with grim retaliation. We will show that (σ_G, σ_G) is a NE. Suppose both players are using σ_G; then P_1 gets the payoff

$$\widehat{V}_{11}^1 = V_{11}^1(0,0) = \frac{1}{1-\gamma}.$$

Now P_1 considers deviating. If he shoots once at some round, then P_2 will keep shooting back in all subsequent rounds and then P_1's best strategy is to also keep shooting. He only needs to consider the deviating strategy in which he shoots at the first round, with probability one. This deviation yields:

$$\begin{aligned}\widetilde{V}_{11}^1 &= p_1 V_{11}^1(1,0) + (1-p_1) V_{11}^1(1,1) \\ &= \frac{1-\gamma(1-p_1)(1-p_2) - p_2(1-p_1)}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))}\end{aligned}$$

Then we have

$$\begin{aligned}\widehat{V}_{11}^1 - \widetilde{V}_{11}^1 &= \frac{1}{1-\gamma} - \frac{1-\gamma(1-p_1)(1-p_2) - p_2(1-p_1)}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))} \\ &= \frac{p_2(1-p_1)}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))} > 0\end{aligned}$$

Hence (σ_G, σ_G) is a NE for all γ, p_1, p_2 values.

3 The Undiscounted Basic Game

3.1 Rules of the Game

The undiscounted game is played just like the discounted one, but the payoffs are different. We have a *kill payoff* $a > 0$ and a *death cost* $-b < 0$ and there is no discounting. The *round payoff* $q_n(\mathbf{s})$ to P_n is

$$\begin{aligned}&q_1(00) = a-b, \quad q_1(01) = -b, \quad q_1(10) = a, \quad q_1(11) = 0, \\ &q_2(00) = a-b, \quad q_2(01) = a, \quad q_2(10) = -b, \quad q_1(11) = 0.\end{aligned}$$

The *total payoff* is

$$Q_n(\mathbf{s}(0)\,\mathbf{s}(1)\,\mathbf{s}(2)\,...) = \sum_{t=0}^{\infty} q_n(\mathbf{s}(t))$$

and the *expected total payoff* is

$$\overline{Q}_n = \mathbb{E}(Q_n(\mathbf{s}(0)\,\mathbf{s}(1)\,\mathbf{s}(2)\,...)).$$

3.2 Payoff under Stationary Strategies

With the same notations as previously we have

$$V_{00}^1 = a-b, \qquad V_{01}^1 = -b, \qquad V_{10}^1 = a.$$

The payoff equation for V_{11}^1 is:

$$\begin{aligned}V_{11}^1 &= x_1 p_1 x_2 p_2 V_{00}^1 + (x_1(1-p_1) + (1-x_1)) x_2 p_2 V_{01}^1 + x_1 p_1 (x_2(1-p_2) + (1-x_2)) V_{10}^1 \\ &\quad + (x_1(1-p_1) + (1-x_1))(x_2(1-p_2) + (1-x_2)) V_{11}^1\end{aligned}$$

or

$$\begin{aligned}V_{11}^1 &= x_1 p_1 x_2 p_2 (a-b) - (x_1(1-p_1) + (1-x_1)) x_2 p_2 b + x_1 p_1 (x_2(1-p_2) + (1-x_2)) a \\ &\quad + (x_1(1-p_1) + (1-x_1))(x_2(1-p_2) + (1-x_2)) V_{11}^1\end{aligned}$$

Solution is:

$$V_{11}^1 = \frac{x_1p_1a - x_2p_2b}{x_1p_1 + x_2p_2 - x_1p_1x_2p_2}$$

Similarly, for P_2, we get

$$V_{11}^2 = \frac{x_2p_2a - x_1p_1b}{x_1p_1 + x_2p_2 - x_1p_1x_2p_2}$$

3.3 Stationary Equilibria

The stationary unique NE is $(\widehat{x}_1, \widehat{x}_2) = (1,1)$ because of the following. When $\max(x_1, x_2) > 1$:

$$\begin{aligned}\frac{dV_{11}^1}{dx_1} &= D_{x_1}\left(\frac{x_2p_2b - x_1p_1a}{x_1p_1x_2p_2 - x_1p_1 - x_2p_2}\right)\\ &= -p_1x_2p_2\frac{-a + x_2p_2b - b}{(x_1p_1x_2p_2 - x_1p_1 - x_2p_2)^2} = \frac{p_1x_2p_2\,(a + b\,(1 - x_2p_2))}{(x_1p_1x_2p_2 - x_1p_1 - x_2p_2)^2} > 0\end{aligned}$$

$$\begin{aligned}\frac{dV_{11}^2}{dx_1} &= D_{x_2}\left(-\frac{x_2p_2a - x_1p_1b}{x_1p_1x_2p_2 - x_1p_1 - x_2p_2}\right)\\ &= -p_2x_1p_1\frac{-a + x_1p_1b - b}{(x_1p_1x_2p_2 - x_1p_1 - x_2p_2)^2} = \frac{p_2x_1p_1\,(a + b\,(1 - x_1p_1))}{(x_1p_1x_2p_2 - x_1p_1 - x_2p_2)^2} > 0\end{aligned}$$

So P_1 (resp. P_2) must take $x_1 = 1$ (resp. $x_2 = 1$). On the other hand, wWhen $x_1 = x_2 = 0$ and P_1 deviates to $x_1' > 0$ we get

$$V_{11}^1(0,0) - V_{11}^1(x_1, 0) = -a < 0$$

So P_1 (resp. P_2) has motive to deviate from $x_1 = 0$ (resp. from $x_2 = 0$).

3.4 Nonstationary Equilibria

Let σ_G be the no-shoot strategy with grim retaliation. We will show that (σ_G, σ_G) is a NE. Suppose both players are using σ_G; then P_1 gets the payoff

$$\widehat{V}_{11}^1 = V_{11}^1(0,0) = 0.$$

Now P_1 considers deviating. If he shoots once at some round, then P_2 will keep shooting back in all subsequent rounds and then P_1's best strategy is to also keep shooting. He only needs to consider the deviating strategy in which he shoots at the first round, with probability one. This deviation yields:

$$\begin{aligned}\widetilde{V}_{11}^1 &= p_1V_{11}^1(1,0) + (1 - p_1)\,V_{11}^1(1,1)\\ &= \frac{-ap_1^2p_2 + ((a+b)p_2 + a)\,p_1 - bp_2}{(p_2 - 1)p_1 - p_2}.\end{aligned}$$

Then we have

$$\widehat{V}_{11}^1 - \widetilde{V}_{11}^1 = \frac{-a\,(p_1 + p_1p_2\,(1 - p_2)) + bp_2\,(1 - p_1)}{(1 - p_2)p_1 + p_2}$$

and for a NE we must have

$$bp_2\,(1 - p_1) > a\,(p_1 + p_1p_2\,(1 - p_2))$$

and, from the similar condition for P_2,

$$bp_1\,(1 - p_2) > a\,(p_2 + p_1p_2\,(1 - p_1))$$

Combining these we get that (σ_G, σ_G) is a NE if

$$\frac{b}{a} > \max\left\{\frac{p_1 + p_1p_2\,(1 - p_2)}{p_2\,(1 - p_1)}, \frac{p_2 + p_1p_2\,(1 - p_1)}{p_1\,(1 - p_2)}\right\}.$$

4 The Evolutionary Tournament

4.1 Fitness Based Tournament

This tournament consists of J generations, each of which involves N players and M strategies. In each generation every player plays a duel against every other player; the duel can be either discounted or undiscounted. The idea is that each strategy is evaluated by its *fitness*, i.e., its performance (summed over all players who use the strategy) against all players and the fitness determines the proportion of each strategy in the nextgeneration. The tournament is described by the following pseudocode.

Algorithm 1 Fitness Based Evolutionary Tournament

Input: kill prob. p, discount factor γ, no. of players N, strategy set $\mathcal{S}$, no. of generations J
for $n \in \{1, 2, ..., N\}$ **do**
    P_n selects randomly $\sigma_n \in \mathcal{S}$
end for
for $j \in \{1, 2, ..., J\}$ **do**
    for $(n_1, n_2) \in \{1, ..., N\} \times \{1, ..., N\}$ **do**
        P_{n_1} and P_{n_2} fight a duel
    end for
    for $n \in \{1, ..., N\}$ **do**
        A_n^j is P_n's total payoff from all duels he fought
    end for
    for $\sigma \in \mathcal{S}$ **do**
        $B_\sigma^j = \sum_{n: P_n \text{ uses } \sigma} A_n^j$
    end for
    for $\sigma \in \mathcal{S}$ **do**
        $F_\sigma^j = \frac{B_\sigma^j}{\sum_{\sigma \in \mathcal{S}} B_\sigma^j}$ is the fitness of σ
    end for
    for $(n_1, n_2) \in \{1, ..., N\} \times \{1, ..., N\}$ **do**
        P_n selects $\sigma \in \mathcal{S}$ with prob. F_σ^j
    end for
end for
return $\left(F_\sigma^J\right)_{\sigma \in \mathcal{S}}$

It is also possible to modify the algorithm so that strategies propagate into the next generation acording to individual player performance. Initial experiments indicate that this tournament gives reasonable results when applied to the discounted game, but performs terribly when applied to the undiscounted one.

4.2 Imitation Based Tournament

This tournament also involves N players, M strategies and J generations, and in each generation every player plays a duel against every other player; the duel can be either discounted or undiscounted (so far I have only tested this with the undiscounted game). The idea here is that, in every generation each player has a probability π of switching his strategy to that of the best performing player. This is a simple evolutiomnary dynamic and it should be easy to establish its properties (and also for general tournaments, not just for duels). The tournament is described by the following pseudocode.

Algorithm 2 Imitation Based Evolutionary Tournament

```
Input: kill prob. p, kill gain a, death loss b, no. of players N, strategy set S, mutation prob. π, no. of generations J
for n ∈ {1, 2, ..., N} do
    P_n selects randomly σ_n ∈ S
end for
for j ∈ {1, 2, ..., J} do
    for (n_1, n_2) ∈ {1, ..., N} × {1, ..., N} do
        P_{n_1} and P_{n_2} fight a duel
    end for
end for
for n ∈ {1, ..., N} do
    A_n^j is P_n's total payoff from all duels he fought
end for
n̂ = arg max_n A_n^j
for n ∈ {1, 2, ..., N} do
    P_n switches to P_n̂'s strategy w.p. π.
end for
return The players' strategies
```

5 Additional Remarks

1. We must perform more experiments.
   (a) Apply imitation based tournament to the undiscounted game.
   (b) Use more strategies.
   (c) Compare performance when using theoretical payoff formulas vs. simulation-obtained payoffs.
   (d) Try playing *only one match* (two players) per generation. Or, more generally, $K \leq N^2$ matches.
   (e) Compare performance of player evaluation vs. strategy evaluation. Why don' these give identical results?
2. Fitness functions:
   (a) Fitness functions apprently cannot handle well *negative scores*. Can we improve this?
3. Extensions to tournaments on *graphs*.
   (a) When the graph is a clique, each player chooses the best overall strategy.
   (b) In graph, players should choose the best player / strategy in their neighborhood.
   (c) For general graph, we could compute a local fitness function, using payoffs of each vertex neighborhood.
4. Need to study more evolutionary game theory.
   (a) What are the properties of evolutionary stable equilibria?
   (b) What are alternative evolutionary dynamics?
5. *Mathematical* extensions.
   (a) Model the change of strategy proportions across generations (ODE? MC?).

(b) In the fitness tournament we essentially have an update

$$z_{m,t+1} = \frac{F\left(\sum_{k\neq m} z_{k,t} Q_{mk}\right)}{\sum_{m=1}^{M} F\left(\sum_{k\neq m} z_{k,t} Q_{mk}\right)} \tag{5.1}$$

We need to study the *convergence* of (5.1). E.g., for $F : \mathbb{R} \to \mathbb{R}^+$, increacing function of s, is it a contraction? Look at fixed point theorems (Nash?), derivatives etc. If we cannot get theoretical results, maybe study numerically.

(c) Can we replace `randsample` with a deterministic operation? Then we will have a deterministioc system, perhaps easier to study.

(d) Further (and more general, not just for duel tournaments) of the imitation based dynamic.

i. This appears to be a very simple rule and it also looks like its analysis will not be too complicated.

ii. Also consider the case when at most $K \in \{1, ..., N\}$ changes can take place per generation (or iteration).

iii. Note: in this case, instead of evolutionary tournament and generations, we can think of a "process" and its iterations.

iv. Variation of the imitation protocol: P_n imitates the strategy of a randomly chosen player (neighbor) P_m with probability π which depends on the extended fitness values of involved partners:

$$\pi = \frac{1}{1 + e^{-\frac{\text{Fit}(n) - \text{Fit}(m)}{K}}}$$

(e) Lainiotis strategy propagation:

$$z_{m,t+1} = \frac{z_{m,t} F\left(\sum_{k\neq m} z_{k,t} Q_{mk}\right)}{\sum_{m=1}^{M} z_{m,t} F\left(\sum_{k\neq m} z_{k,t} Q_{mk}\right)}$$

6. **More generally**: what can evolutionary tournament (or evolutionary Game Theory) tell us about NE selection?

asas