Teoria de Jogos Evolutivos – O Dilema do Prisioneiro

Dentre as diversas áreas que compartilham problemas e métodos com a Física Estatística está a teoria de jogos e a biologia. Em especial, um problema com interesse que transcende estas áreas é o surgimento do comportamento cooperativo entre indivíduos ou entes de uma comunidade. A cooperação em uma sociedade implica em um custo individual para beneficiar outros. Do ponto de vista evolutivo, fica a questão que sendo as espécies moldadas pela competição, onde apenas os mais aptos sobrevivem, como a cooperação pode surgir e se manter estável? Claramente a cooperação mútua leva a um benefício coletivo, no entanto, a traição leva a um benefício individual maior. Essa é uma das grandes questões em aberto nas ciências [1,2] e seus impactos envolvem desde uma melhor compreensão dos fenômenos biológicos a um aprofundamento no entendimento de diversas questões sociais e econômicas da nossa sociedade.

A formulação matemática deste problema surgiu na área da matemática chamada de teoria de jogos clássica, que procura, entre outras coisas, trazer um foco quantitativo à resolução de problemas de conflito estratégico entre seres racionais. Posteriormente, a junção destes conceitos com um processo evolutivo, deu origem à teoria de jogos evolucionários, onde os impactos dos conflitos estratégicos são utilizados na determinação da evolução da população, de forma que as melhores estratégias se propagariam mais facilmente. Mas, em geral, as análises não indicavam a possibilidade de um estado estável com presença de cooperação entre os indivíduos ao final da evolução.

Na década de 90 Nowak e May [3] introduziram modelos e técnicas normalmente utilizadas na mecânica estatística para modelagem da teoria de jogos evolucionários. Em especial, deram início ao uso de redes e grafos para modelar a interação entre indivíduos, de forma semelhante ao que vimos ao estudar o modelo de Ising e o processo de propagação de doenças numa população. Esta nova abordagem gerou resultados muito importantes, uma vez que mostrou que a cooperação pode emergir de forma estável mesmo em cenários altamente competitivos em estruturas que forma grupos (ilhas) de indivíduos que cooperam num mar de traidores.

Nesta tarefa iremos procurar entender como a cooperação surge num modelo simplificado definido numa rede quadrada.

O dilema do prisioneiro.

Trataremos aqui do dilema do prisioneiro, que talvez seja o problema mais famoso em teoria de jogos. De acordo com [1], esse jogo consegue captar, de maneira matematicamente elegante, a essência do conflito entre ser egoísta e obter benefícios próprios ou ser altruísta e tentar ajudar seu parceiro. Na história típica desse problema temos dois bandidos, digamos Paulo e João, que foram presos sob suspeita de um roubo. Porém as provas são insuficientes para prendê-los. Ambos estão isolados um do outro, e o promotor dá a eles as seguintes alternativas: trair seu parceiro, confessando quem cometeu o crime, ou ficar calado. Caso Paulo e João fiquem ambos calados, a investigação poderá prendê-los por apenas 2 anos, mas eventualmente precisará soltá-los por falta de provas. Caso Paulo resolva ficar calado, mas João o traia, confessando o crime de ambos, Paulo ficará na prisão por 10 anos, enquanto João será solto imediatamente. Porém caso ambos se traiam mutuamente, os dois confessando o crime, ambos pegarão 5 anos de cadeia. Considerando que eles não podem se comunicar, e não sabem o que o outro fará, qual a melhor atitude?

O dilema surge do paradoxo ao notar que a melhor opção para João e Paulo é cooperarem mutuamente, ficando somente 2 anos presos. Porém, ao pensar em cooperar, João imediatamente nota que se Paulo tiver certeza de sua cooperação, irá traí-lo para sair livre imediatamente. Assim João decide trair Paulo. O raciocínio é simétrico, levando ambos a se traírem devido à desconfiança, causando o pior cenário possível para ambos. Esse tipo de dilema é justamente o utilizado no cálculo quanto a delações premiadas, ou outros esquemas em que a polícia concede benefícios a criminosos que delatem seus comparsas. Nesse caso, a cooperação equivale a ficar em silêncio perante o promotor, e a deserção equivale a confessar o que eles fizeram. Matematicamente, podemos criar uma matriz identificando quantos anos de cadeia João (ou Paulo) terá que cumprir dependendo de sua ação e a ação de seu parceiro:

João\Paulo	Coopera	Deserta
Coopera	2 \ 2	10 \ 0
Deserta	0 \ 10	5 \ 5

Essa matriz nos diz quantos anos de prisão cada estratégia recebe contra cada outra estratégia possível (o primeiro número é relativo a João e o segundo a Paulo) e recebe o nome de matriz de ganho (mais comumente chamada matriz de payoff na literatura estrangeira e nacional). Aqui os números apresentados equivalem aos anos de cadeia de cada um, porém o mais comum é representar nessa matriz números positivos, indicando benefícios ganhos pelo jogador. Note que a matriz de payoff é idêntica tanto para João, quanto para Paulo. Jogos desse tipo são chamados simétricos, já que ambos jogadores possuirão os mesmos ganhos, dado que usem as mesmas estratégias.

Modelo a ser considerado

Trataremos aqui de uma versão simplificada do dilema do prisioneiro entre os jogadores A e B descrito pela seguinte matriz de ganhos (matriz de payoff) [4]:

A \ B	Coopera	Deserta
Coopera	1	0
Deserta	b	0

Neste caso, se ambos cooperam, ambos recebem a mesma recompensa, 1. Caso ambos desertem, não há recompensa para nenhum dos dois. No caso de A desertar e B cooperar, A receberá a recompensa b, fator que indica a tentação à deserção. Nesta versão simplificada, A não terá nenhum ganho caso B deserte. Para o jogo capturar adequadamente a essência do dilema do prisioneiro que foi descrita acima, devemos escolher b>1, ou seja, o ganho para um desertor caso o oponente coopere é maior que o ganho que ele obteria com a cooperação mútua.

Para levar em conta o fator evolucionário no modelo, assim como o papel desempenhado pela estrutura espacial das interações entre indivíduos, consideraremos uma rede quadrada onde cada nó representa um indivíduo que poderá adotar uma de duas estratégias, ser um cooperador (C) ou ser um desertor (D). Inicialmente cada indivíduo será classificado como cooperador ou desertor e ao longo da evolução do sistema poderá trocar sua estratégia de acordo com uma regra a ser estabelecida que envolve suas interações com sua vizinhança e consigo mesmo. Consideraremos aqui o modelo para evolução considerado por Szabó e Töke [4]. Neste modelo, um indivíduo escolhido ao acaso, X, tem o seu ganho total apurado ao considerar sua interação com os quatro vizinhos mais próximos da rede quadrada, sendo E_X seu ganho total. Caso este indivíduo esteja cercado apenas por desertores, seu ganho total será nulo se ele for um desertor e 1 caso seja um cooperador. No caso de ser um desertor cercado por cooperadores, seu ganho total será 4b, enquanto um cooperador cercado por cooperadores terá ganho total igual a 5. Dentre seus quatro vizinhos, ele escolherá aleatoriamente um deles, Y, cujo ganho total é E_Y. O jogador X irá adotar a estratégia do jogador Y com probabilidade:

$$W = \frac{1}{1 + \exp\left[-\frac{E_Y - E_X}{K}\right]},$$

onde K é um fator que introduz um ruído relacionado à irracionalidade das escolhas. Para $K = 0$ o jogador X adotará a estratégia de Y caso $E_Y > E_X$. Para valores positivos de K, haverá um intervalo de diferenças no ganho total no qual decisões irracionais poderão ser tomadas, ou seja, X poderá não adotar a estratégia de Y mesmo ela sendo melhor que a sua. Vamos nos restringir apenas a casos onde $K < 1$.

Perceba, neste ponto, as semelhanças que este modelo apresenta com os modelos de Ising e de propagação de doenças, mostrando como os métodos e modelos da Física Estatística podem contribuir no entendimento de fenômenos biológicos, sociais e econômicos. Assim, utilizaremos uma simulação de Monte Carlo para analisar esse sistema. Aconselho se basearem nas simulações do Modelo de Ising usando o algoritmo de Metropolis.

A Tarefa

Considerando uma configuração inicial aleatória de cooperadores e desertores na rede evoluiremos o sistema de acordo com o modelo descrito acima. O modelo possui duas fases absorventes, que são fases onde a dinâmica cessa após todos indivíduos assumirem o mesmo estado, seja ele cooperativo ou desertor. Para $b < b_1$ a densidade de cooperadores, c, será 1, enquanto para $b > b_2$ teremos $c = 0$. Para $b_1 < b < b_2$, após um tempo suficientemente longo, o sistema irá tender a um estado estacionário onde a concentração de cooperadores oscilará em torno de um valor constante. Escolha **um** valor de K no intervalo de 0,02 a 0,5 e simule redes de tamanho $L = 200$ para valores de b entre 1 e 2. Observe a evolução temporal do sistema e a distribuição espacial de cooperadores e desertores. Estime os valores de b_1 e b_2 para o valor de K escolhido. Não se esqueça de indicar claramente o valor de K escolhido e de apresentar os dados que suportam as estimativas de b_1 e b_2 assim como a distribuição de cooperadores e desertores para valores intermediários de b.

Referências:

[1] M.A. Amaral, "Teoria dos jogos evolucionários e o surgimento da cooperação: dinâmicas inovativas e jogos mistos" Tese de Doutorado, Departamento de Física, UFMG (2017). http://lilith.fisica.ufmg.br/posgrad/Teses_Doutorado/decada2010/marco-amaral/MarcoAntonioAmaral-tese.pdf

[2] E. Pennisi, Science 309, 93 (2005).

[3] M. A. Nowak e R. M. May, Nature 359, 826 (1992).

[4] G. Szabo e C. Toke, Physical Review E, 58, 69 (1998).