Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 7, n. 1, 2020.

Trabalho apresentado no XXXIX CNMAC, Uberlândia - MG, 2019.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

O Futebol Como Uma Cadeia de Markov

Moiseis S. Cecconello [1]
Departamento de Matemática - UFMT
Greison Araújo de Oliveira[2]
Universidade Federal de Mato Grosso - UFMT

Resumo. Neste trabalho, descrevemos a dinâmica da troca de passes entre os jogadores de um time de futebol como um processo estocástico modelado por uma cadeia de Markov. Vamos analisar jogos de futebol a partir da perspectiva de cadeias de Markov, procurando dar significado para termos como *melhor jogador do time na partida* através de uma medida objetiva.

Palavras-chave. Cadeia de Markov, Futebol, Estados Estacionários, Modelagem Matemática.

1 Introdução

As cadeias de Markov são largamente usadas como modelos matemáticos para eventos de natureza estocástica ou aleatórias em muitas áreas das ciências, tais como em física, biologia, engenharias, economia e ciências sociais [3]. Quase que certamente, uma das aplicações mais relevante de cadeia de Markov da atualidade, é o uso que o Google faz das mesmas para determinar a importância de páginas da internet através do seu mecanismo de buscas [2].

Como é de conhecimento comum, um dos principais elementos que promove a dinâmica em um jogo de futebol é a troca de bolas entre os jogadores das equipes envolvidas na partida. Quando um determinado jogador possui a posse da bola, o mesmo se movimenta procurando uma melhor opção para passá-la adiante, para outro jogador, ou buscar uma finalização ao gol adversário. Tal dinâmica possui caráter estocástico, cujas probabilidades de trocas de passes entre os jogadores depende de diversos fatores, como o posicionamento tático da equipe, a importância do jogador para a equipe, a distância entre os jogadores e etc.

Dessa forma, na Seção 3 desse trabalho, propomos que a dinâmica proveniente da troca de passes de uma equipe em um jogo de futebol pode ser modelada por uma cadeia de Markov. Além disso, utilizamos alguns jogos realizados pela Seleção Brasileira para estimar a probabilidade de transição entre os estados da cadeia de Markov e analisar os

[1]moiseis@gmail.com.
[2]greisonaraujo16@gmail.com

seus estados estacionários, visando identificar elementos que possam caracterizar a Seleção Brasileira enquanto uma equipe de futebol.

Até o presente momento, a proposta de modelagem aqui apresentada é inédita de acordo com os principais mecanismos de buscas da internet. Em [1], os autores fazem uso de cadeias de Markov para avaliar o desempenho de times de beisebol, bem como a influência de determinados jogadores no desempenho de tais times.

2 Cadeias de Markov

Para atender os propósitos desse trabalho, vamos considerar apenas cadeias de Markov em tempo discreto, com um número finito de estados.

Consideremos, então, um conjunto de estados $S = \{s_1, s_2, \ldots, s_k\}$ e um processo estocástico $(X_n)_{n\in\mathbb{N}}$ tomando valores sobre o conjunto de estados S. Isto é, em cada $n \in \mathbb{N}$, X_n assume um dos possíveis valores de S, de acordo com algum probabilidade P. Tal processo estocástico é dito ser uma *cadeia de Markov* quando a distribuição de probabilidade de X_{n+1} depende apenas da distribuição de probabilidade de X_n. Isto é,

$$P(Z_{n+1} = i | Z_1 = j_1, Z_2 = j_2, \ldots, Z_n = j) = P(Z_{n+1} = i |, Z_n = j). \tag{1}$$

Sendo assim, uma cadeia de Markov sobre S é determinada por uma *matriz de transição* $P = [p_{ij}]_{1\leqslant i,j\leqslant k}$ em que

$$p_{ij} = P(Z_{n+1} = i |, Z_n = j). \tag{2}$$

Em consequência da definição acima, temos que

$$\sum_{i=1}^{k} p_{ij} = 1. \tag{3}$$

Como definido acima, a matriz de transição descreve a dinâmica entre os estados possíveis de uma variável, definidos pelo conjunto S, em cada passo do tempo. Assim, se $x^n = (x_1^n, x_2^n, \ldots, x_k^n)^T$ é um vetor que descreve a probabilidade X_n assumir um dos estados de S, então o vetor de probabilidades x^{n+1} de X_{n+1} é dado pela recorrência

$$x^{n+1} = Px^n. \tag{4}$$

Como descrevem probabilidades, as coordenadas de x^m são não negativas e somam 1, para todo $m \in \mathbb{N}$.

Pela recorrência anterior, o processo estocástico X_n possui um *estado estacionário* $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k)^T$ caso a equação

$$P\bar{x} = \bar{x}$$

tenha solução. A existência de estado estacionário $\bar{x}$, bem como a convergência de x^n para $\bar{x}$, para cadeias de Markov, são garantidas pelo teorema de Perron-Frobenius [2].

Seja μ_i o número médio de passos necessários para o processo estocástico, iniciando no estado s_i, retornar ao estado s_i. Dada uma distribuição estacionária $\bar{x}$, com $\bar{x}_i > 0$, temos então que

$$\mu_i = \frac{1}{\bar{x}_i}. \tag{5}$$

Isto é, o tempo médio de retorno ao estado s_i é inversamente proporcional ao valor $\bar{x}_i$ [3].

Portanto, neste contexto, podemos interpretar as coordenadas de um estado estacionário de uma cadeia de Markov como a importância que o estado associado a esta coordenada tem sobre os demais estados na dinâmica do processo estocástico.

3 O futebol como uma cadeia de Markov

Um dos principais elementos que determina a dinâmica de uma partida de futebol é a troca de passes entre os jogadores de um time. Quando um determinado jogador possui a posse de bola, o mesmo se movimenta a fim de encontrar uma melhor opção para passar a bola para outro jogador do time, ou buscar uma finalização ao gol adversário. Naturalmente, a troca de passes entre os jogadores depende de diversos fatores, tais como, o posicionamento tático da equipe, a importância do jogador para a equipe, a distância entre os jogadores, a posicionamento de jogadores adversários, e, claro, da capacidade técnica dos jogadores envolvidos no jogo entre outros.

Os fatores mencionados anteriormente, portanto, fazem a dinâmica de um jogo de futebol possuir característica estocástica, no que diz respeito ao passe ou a troca de bola entre jogadores. Isto é, quando um jogar detém a posse da bola, diversas opções podem aparecer para o passe, cada uma dessas opções com suas respectivas chances de sucesso, determinando a natureza estocástica do jogo.

Diante disso, nesse trabalho propomos descrever a dinâmica da troca de passes entre os jogadores de um time como uma cadeia de Markov. Para efeitos de simplificação, não levaremos em consideração aspectos como posicionamento do jogador em campo e, tão pouco, fatores relacionados ao tempo que cada jogador detém a posse de bola, entre outros fatores.

Isto é, aqui nesse trabalho, vamos supor que um jogo de futebol é uma processo dinâmico estocástico X_n, agindo sobre um espaço com 11 estados possíveis, digamos $S = \{j_1, j_2, \ldots, j_{11}\}$, que descreve os eventos *o jogador i está com a bola* ou *a bola está com o jogador i*. O índice n do processo estocástico X_n se refere ao número de passes trocados entre o jogadores. Isto é, X_n descreve o estado do sistema após n trocas de passes.

Por ser um processo estocástico com um número finito de estados e tempo discreto, a dinâmica da troca de passes em um jogo de futebol pode, então, ser representada por uma matriz de transição $P = [p_{ij}]_{1 \leqslant i,j \leqslant k}$. Assim, no contexto da aplicação dessa seção, o elemento p_{ij} da matriz de transição representa a probabilidade do jogador j passar a bola para o jogador i, ou equivalentemente, a probabilidade do jogador i receber a bola do jogador j quando j detém a posse da bola.

Nessa representação ideal do jogo de futebol por cadeias de Markov, a equação de recorrência

$$x^{n+1} = Px^n,$$

descreve a transição da probabilidade de um jogador i estar com a posse da bola após um passe efetuado. Podemos, então, estar interessados em determinar o estado estacionário do processo estocástico associado.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics. v. 7, n. 1, 2020.

No contexto da aplicação aqui apresentada, a coordenada $\bar{x}_i$ de um estado estacionário $\bar{x}$ é inversamente proporcional ao número de médio passes necessários para a bola retornar ao jogador i. Isto é, quanto maior é o número médio de passes necessários para a bola retornar para um jogador, menor é o valor da coordenada do estado estacionário associada ao referido jogador.

Dessa forma, podemos interpretar o valor de $\bar{x}_i$ como a *importância* do jogador i para a equipe. Quanto maior é o valor de $\bar{x}_i$, menor é o número médio de passes necessários para a bola retornar ao jogador i, de modo que, em um número fixo de passos, mais vezes i detém a posse da bola. Assim, a relação expressa em (5) pode ser usado como um critério objetivo para determinar a importância de um atleta de um time em uma partida de futebol.

4 Aplicações e análises de jogos

Nesta seção, utilizaremos jogos da Seleção Brasileira de futebol para obtermos os dados necessários e construirmos as matrizes de transição e, em seguida, prosseguimos com as análises dos jogos através dos respectivos estados estacionários.

Para cada um dos jogos a seguir, os dados foram coletados via transmissão de vídeo disponíveis em sites da internet. Para cada passe dado corretamente pela equipe brasileira, são anotados o jogador de origem e o jogador de destino do passe e, tais dados, são catalogados em uma matriz $A = [a_{ij}]_{1\leqslant i,j\leqslant 11}$, cuja coluna j representa o jogador de origem e a linha i representa o jogador de destino do passe, como por exemplo a matriz (7). A matriz de transição $P = [p_{ij}]_{1\leqslant i,j\leqslant 11}$ é então obtida dividindo os elementos de cada coluna pela sua respectiva soma, obtendo assim as probabilidades de passes entre os jogadores, que define a matriz de transição. Isto é, p_{ij} é dado por:

$$p_{ij} = \frac{a_{ij}}{\sum_{i=1}^{n} a_{ij}}. \tag{6}$$

A matriz (7) sintetiza os dados obtidos do jogo Brasil $\times$ Alemanha, realizado no Japão, no dia 30 de junho de 2002, jogo este válido pela final da Copa do Mundo de 2002.

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 3 & 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 2 & 0 & 3 & 1 & 3 & 0 \\ 2 & 0 & 1 & 2 & 0 & 2 & 5 & 1 & 8 & 0 & 0 \\ 0 & 2 & 2 & 4 & 3 & 0 & 1 & 0 & 4 & 3 & 2 \\ 0 & 1 & 1 & 1 & 7 & 2 & 0 & 3 & 7 & 2 & 0 \\ 0 & 2 & 1 & 2 & 0 & 1 & 3 & 0 & 5 & 2 & 2 \\ 4 & 0 & 2 & 3 & 13 & 2 & 4 & 7 & 0 & 8 & 2 \\ 4 & 2 & 2 & 4 & 5 & 3 & 3 & 3 & 8 & 0 & 5 \\ 0 & 3 & 0 & 4 & 1 & 3 & 4 & 5 & 5 & 4 & 0 \end{pmatrix} \tag{7}$$

Na matriz acima, as linhas, e também as colunas, são definidas pela seguinte ordem dos jogadores: Marcos (1), Lúcio (3), Edmilsom (5), Roque Junior (4), Cafu (2), Gilberto

DOI: 10.5540/03.2020.007.01.0410

(8), Kleberson (15), Roberto Carlos (6), Ronaldinho (11), Rivaldo (10) e Ronaldo (9). Os números entre parenteses indicam a numeração do uniforme usado pelo jogador. Para efeitos de ilustração, nesse jogo, o atleta Ronaldo lançou a bola corretamente para o Rivaldo em cinco ocasiões de um total de 11 passes completos realizados pelo Ronaldo, conforme indicado pela última coluna de A, indicando assim uma probabilidade de 5/11 do atleta Rivaldo receber um passe do Ronaldo.

Já a matriz apresentada em (8), sintetiza os passes completados dados pela equipe brasileira nos cinco jogos realizados na Copa do Mundo de 2018, com linhas, e também as coluna, definidas pela seguinte ordem dos jogadores: Allisson (1), Marcelo (12), Miranda (3), Thiago Silva (4), Fagner (2), Casemiro (5), Neymar (10), Philippe Coutinho (11), Paulinho (15), Willian (19), Gabriel Jesus (9). Para efeitos de simplificação, mantemos a contagem de passes completados mesmo quando um jogador é substituído.

$$A = \begin{pmatrix} 0 & 1 & 8 & 7 & 2 & 2 & 1 & 0 & 0 & 3 & 0 \\ 6 & 0 & 65 & 36 & 4 & 33 & 60 & 73 & 4 & 12 & 1 \\ 17 & 38 & 0 & 57 & 10 & 14 & 10 & 15 & 10 & 5 & 0 \\ 11 & 8 & 45 & 0 & 55 & 24 & 6 & 14 & 18 & 18 & 0 \\ 9 & 4 & 10 & 49 & 0 & 22 & 6 & 10 & 28 & 29 & 4 \\ 1 & 28 & 24 & 29 & 15 & 0 & 15 & 33 & 14 & 21 & 5 \\ 4 & 94 & 11 & 12 & 7 & 21 & 0 & 84 & 8 & 13 & 19 \\ 2 & 83 & 36 & 25 & 9 & 41 & 66 & 0 & 17 & 13 & 15 \\ 1 & 6 & 3 & 10 & 28 & 25 & 8 & 21 & 0 & 19 & 6 \\ 6 & 6 & 4 & 18 & 59 & 17 & 6 & 12 & 20 & 0 & 10 \\ 4 & 18 & 0 & 3 & 7 & 8 & 26 & 7 & 7 & 9 & 0 \end{pmatrix} \tag{8}$$

Embora a cadeia de Markov definida pela matriz A não seja regular, podemos restringir nossa análise à cadeia de Markov obtida por desconsiderar a primeira linha e primeira coluna da matriz (8). Nesse caso, não é difícil verificar que os dados resultantes definem uma matriz P tal que P^2 tem todas as entradas estritamente positivas [3]. A Tabela 1 mostra o estado estacionário $\bar{x}$, em ordem decrescente dos valores de suas entradas, da cadeia de Markov definida pela matriz de transição gerada a partir dos dados da matriz (7). Assim, de acordo com o critério adotado na seção anterior, a análise do estado estacionário nos permite concluir que o jogador número 11, Ronaldinho Gaúcho, é o atleta que tem o menor número médio de passes necessários para a bola retornar até o mesmo, seguido pelos atletas Rivaldo e Ronaldo.

Não é difícil verificar que a matriz P, obtida de (8), define uma cadeia de Markov regular, pois P^2 tem todas entradas estritamente positivas [3]. Portanto, como anteriormente, de acordo com a relação expressa em (5), a análise do estado estacionário nos permite concluir que o jogador Philippe Coutinho é o atleta que tem o menor número médio de passes necessários para a bola retornar até o mesmo, seguido pelos atletas Marcelo e Neymar, conforme indica os dados na Tabela 2.

Na Figura 1, vemos os grafos orientados gerados a partir das matrizes de incidências 7 e 8. Em cada grafo, os nós indicam os jogadores e as arestas são os passes orientados de acordo com a origem dos mesmos. As origens dos passes também são indicadas pelas cores das arestas. Já as espessuras das arestas são proporcionais ao número de passes

Tabela 1: Estado estacionário, em ordem decrescente, da matriz de transição em (7)

Número	Jogador	Probabilidade
11	Ronaldinho	0.192
10	Rivaldo	0.180
9	Ronaldo	0.136
15	Kleberson	0.101
8	Gilberto	0.100
6	Roberto Carlos	0.098
2	Cafu	0.087
4	Roque Júnior	0.057
3	Lúcio	0.031
5	Edmilson	0.016

Tabela 2: Estado estacionário, em ordem decrescente, da matriz de transição em (8).

Número	Jogador	Probabilidade
11	Philippe Coutinho	0.165
12	Marcelo	0.153
10	Neymar	0.149
5	Casemiro	0.094
4	Thiago Silva	0.087
3	Miranda	0.078
2	Fagner	0.078
19	Willian	0.073
15	Paulinho	0.064
9	Gabriel Jesus	0.049
1	Allisson	0.010

efetuados entre dois jogadores conectados. O mapa de cores é gerado a partir dos valores do estado estacionário das tabelas 1 e 2, respectivamente. Conforme podemos notar no grafo da esquerda, a equipe brasileira mostra um equilíbrio na parte ofensiva, tendo uma simetria, com relação ao centro do campo na distribuição dos jogadores com os melhores índices, definidos a partir do estado estacionário. Já o grafo da direita nos mostra que a equipe brasileira apresenta uma maior dinâmica pelo lado esquerdo de seu ataque. Isto é, a equipe apresenta uma assimetria na distribuição dos jogadores com melhores índices, com relação a faixa central do campo de jogo.

O mapa de cores, gerado a partir dos valores das entradas do estado estacionário conforme indicado, nos mostra que a equipe brasileira apresenta uma maior dinâmica pelo lado esquerdo de seu ataque. Isto é, a equipe apresenta uma assimetria na distribuição dos jogadores com melhores índices, com relação a faixa central do campo de jogo.

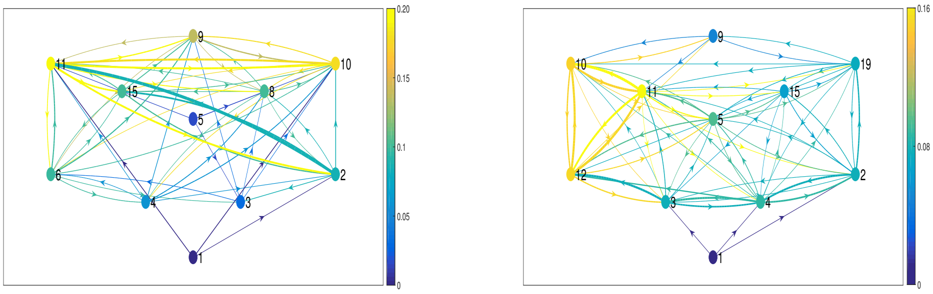

Figura 1: Grafos gerados a partir das matrizes 7 (esquerda) e 8 (direita). Os nós representam os atletas, identificados pelo número do uniforme, enquanto que as cores indicam o valor da respectiva coordenada dos estados estacionários das tabelas 1 e 2.

5 Conclusões

Neste trabalho, propomos um modelo para descrever a dinâmica de um jogo de futebol por meio de uma cadeia de Markov. A análise do estado estacionário da cadeia de Markov de um jogo de futebol pode, por exemplo, conferir um critério objetivo para a identificação dos atletas mais relevantes, que compõe a equipe, no que se refere à dinâmica da partida em termos da troca de passes. Mais ainda, uma análise nos valores das entradas de um estado estacionário pode também revelar graus de simetria com relação as posições de atuação dos jogadores de uma equipe. Isto é, a análise dos estados estacionários pode permitir identificar estratégias e táticas da dinâmica de um determinado time. Dessa forma, o modelo proposto aqui pode ser de utilidade prática para equipes de futebol, auxiliando tanto para identificar informações relevantes sobre equipes adversárias, quanto para análise da própria equipe.

Agradecimentos

Expressamos nosso agradecimento ao Laboratório de Matemática Aplicada - LAMAP, do Departamento de Matemática da UFMT, pelo acesso aos recursos computacionais que permitiram a análise dos dados aqui apresentados.

Referências

[1] B. Bukiet,E. R. Harold and J. L. Palacios. A Markov Chain Approach to Baseball, Operations Research, v. 1, n. 45, p. 14– 23, 1997.

[2] C. Rousseau, Y. Saint - Aubin. Matemática e atualidade, Volume 1, SBM, Rio de Janeiro, 2015.

[3] N. Privault. Undesrtanding Markov Chains: examples and Applications, Springer, Singapore, 2013.

