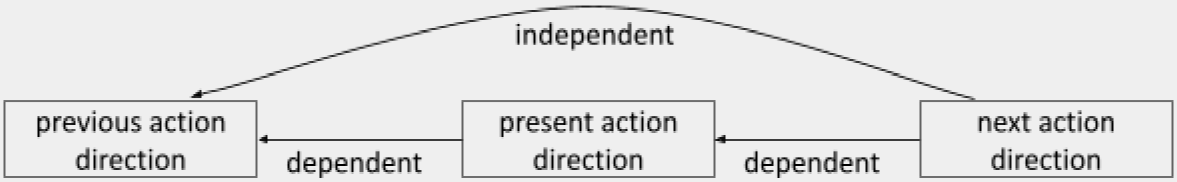
Moving the ball forward is a general goal in a couple of the most popular sports. However, if we look into a game moment by moment, the strategic decisions on where to move the ball are actually diverse. In soccer, a successful offense sequence always consists of manipulating the ball back and forth at the team's own rhythm. This kind of rhythm lies in the transition probabilities between each movement. The introduction of the Markov Chain can provide a gateway to understanding those transition probabilities and how they converge into a state influence on the game. The purpose of this module is to demonstrate how to use Markov chain theory to predict the probability of ball movement in a game, with known play patterns and the starting action.

Introduction of Markov Chain Theory

Basic Concept

Markov chain (or Markov process) is a stochastic process that has the "**memoryless**" property. A process is memoryless if the conditional probability distribution of the next state of the process depends only upon the present state, not any previous states.

Suppose we assume that in a soccer game, the probability of the ball movement direction in each action (i.e., state) is dependent on the last action and is independent of any action before the last action. In that case, we are claiming this sequence of movement direction can be described by a Markov chain.



Q: What are other events that have the memoryless property?

State and state space

A Markov chain contains a sequence of random variables X_1, X_2, ... X_t, each one indicates the present state at time t. The set of possible states that each X_t can take is the **state space**, S.

In the soccer game case, if we categorize the ball movement directions into three states: forward, horizontal, and backward, we create a discrete state space:

$$S = \{\text{'forward', 'horizontal', 'backward'}\}$$

Q: If the state space is continuous instead of discrete, is it still a Markov chain?

Transition matrix

The changes in the state are called **transitions**. The conditional probabilities associated with state changes are **transition probabilities**, which can be expressed as:

$$P(X_{t+1}=s|X_t=s_t)$$

The memoryless property mentioned above can be expressed as:

$$P(X_{t+1}=s|X_t=s_t, X_{t-1}=s_{t-1}, ..., X_o=s_o) = P(X_{t+1}=s|X_t=s_t)$$

The conditional probabilities $P(X_{t+1}=s|X_t=s_t)$ vary in transitions between various states. With the probability of transition from i to j as $Pr(j|i) = P_{i,j}$, and the number of possible states in S as α, a **transition matrix** of the Markov chain can be formed as:

$$P = (p_{i,j}),\ 1 \leq i, j \leq \alpha, \text{ or}$$

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,j} & \cdots & P_{1,\alpha} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,j} & \cdots & P_{2,\alpha} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{i,1} & P_{i,2} & \cdots & P_{i,j} & \cdots & P_{i,\alpha} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ P_{\alpha,1} & P_{\alpha,2} & \cdots & P_{\alpha,j} & \cdots & P_{\alpha,\alpha} \end{bmatrix}$$

In this matrix, every row contains the transition probabilities from state i to all possible states, thus the rows should each sum to 1.

Suppose we had access to data from the FA Women's Super League and analyzed the game Aston Villa Women vs Arsenal WFC on 2021-02-28. We can generate a Markov chain as below:

S = {'forward', 'horizontal', 'backward'}
X_n = movement state at each action

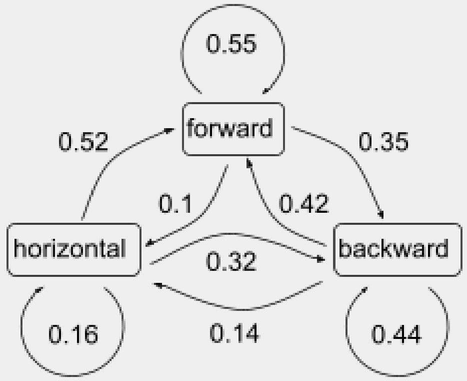
Transition matrix:

$$P = \begin{pmatrix} P_{\text{forward, forward}} & P_{\text{forward, horizontal}} & P_{\text{forward, backward}} \\ P_{\text{horizontal, forward}} & P_{\text{horizontal, horizontal}} & P_{\text{horizontal, backward}} \\ P_{\text{backward, forward}} & P_{\text{backward, horizontal}} & P_{\text{backward, backward}} \end{pmatrix}$$

Transition matrix with calculated values:

$$P = \begin{pmatrix} 0.55 & 0.1 & 0.35 \\ 0.52 & 0.16 & 0.32 \\ 0.42 & 0.14 & 0.44 \end{pmatrix}$$

Transition probability graph:



Q: When non-positive transition probability exists, it indicates that at time t, X_t cannot transition to all possible states defined in S. Is it a situation potentially existing in the soccer game data?

Stationary probability vector

At time t, the distribution among the states is represented by a **state vector** V_t, which is a row matrix that has one column for each state:

$$\begin{array}{cc} & \begin{array}{cccc} 1 & 2 & \ldots & \alpha \end{array} \\ V_t = & [\begin{array}{cccc} P_1 & P_2 & \ldots & P_\alpha \end{array}] \end{array}$$

For each time t in the sequence, the next state vector V_{t+1} can be calculated by multiplying the present state vector by the transition matrix P:

$$V_{t+1} = V_t P$$

Given the starting state vector V_0, we have $V_1 = V_0 P,\ V_2 = V_0 P^2, \ldots$ that is:

$$V_{t=}V_0 P^t$$

If the transition matrix P is constant throughout the sequence, the Markov chain is **time-homogeneous**. A time-homogeneous Markov chain with a finite number of states is a **finite Markov chain**. A finite Markov chain can have a **stationary probability vector** π, which represents the long-term, **steady-state distribution** of the states, where the probabilities do not change further as time goes to infinity. π is invariant by the matrix P:

$$\pi P = \pi$$

The stationary probability vector π is fundamental in understanding the long-term behavior of systems modeled as Markov chains and is used in numerous fields such as finance, economics, network analysis, and sports.

Now assuming we have the transition matrix at the start of a game and would like to predict the ball movements in this game, we can calculate the stationary probability vector.
If the first movement direction in the data is 'horizontal', we have the starting state vector V_0 = [0, 1, 0]. Given transition matrix P and V_0, we have:

$V_1 = V_0 P$ = [0.52, 0.16, 0.32]

$V_2 = V_0 P^2$= [0.50, 0.12, 0.37]

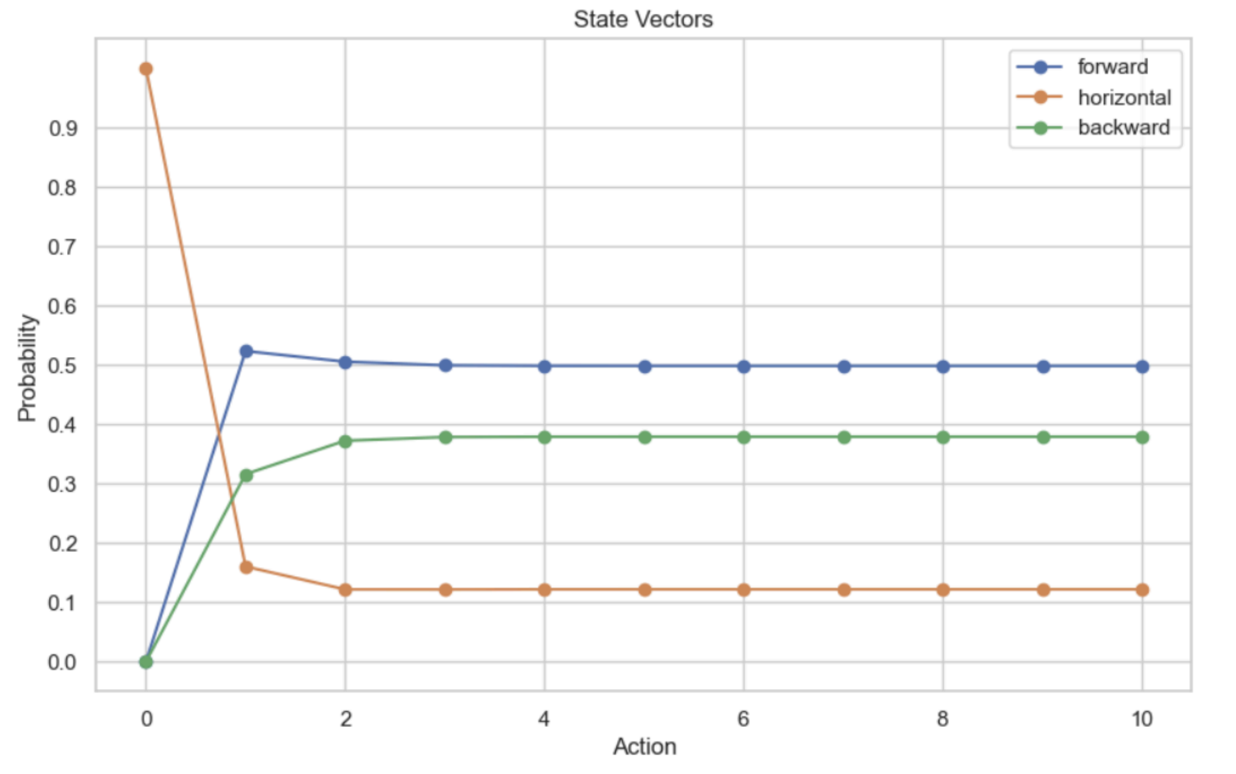
$V_3 = V_0 P^3$= [0.50, 0.12, 0.37]

…

$V_t = V_0 P^t$= [0.50, 0.12, 0.37]

As seen, V_t converges into a stable state – the stationary probability vector π [0.50, 0.12, 0.37]

If we plot the sequence of state vectors, we'll have:



Q: What are the potential usages of the stationary probability vector π?

Sample data source: https://github.com/statsbomb/open-data

Sample code:
https://github.com/lorihe/SCORE_Network/blob/main/soccer_action_direction_markov_chain.ipynb