Análisis Futbol

2024-02-14

Tabla de contenidos

Introducción **3**

1 Formulación **4**

2 Dinámica del modelo **6**

3 Descripción y Justificación de la recompensa **13**

4 Justificación de las acciones **14**

References **15**

Introducción

Para la construcción del Proceso de Decisión de Markov (MDP) del planteado en este trabajo, primero debemos recordar algunas reglas durante un partido de fútbol para aquellos que no las conozcan y también haceremos algunas puntualizaciones respecto a la naturaleza de un partido.

Además se darán algunas hipótesis para simplificar el modelo sin pérder la escencia del juego.

Un partido de fútbol se componen de 2 equipos con 11 jugadores cada uno, un balón y un arbitro (o un grupo de arbitros). Los equipos se posicionan dentro de un terreno de juego que consta de 2 porterias (una para cada equipo), el objetivo de cada equipo es anotar un gol (que el balón ingrese por completo dentro de la porteria) en la portería del equipo contrario. Por cada gol un equipo recibe lo que sería 1 punto, al final del partido el equipo que anote más goles gana y en caso de anotar la misma cantidad se considera empate.

Para anotar un gol cada jugador puede tocar el balón con cualquier parte del cuerpo salvo el brazo, sin embargo, de los 11 jugadores de cada equipo hay un futbolista que puede tomar el balón con las manos dentro de las áreas marcadas alrededor de las porterias.

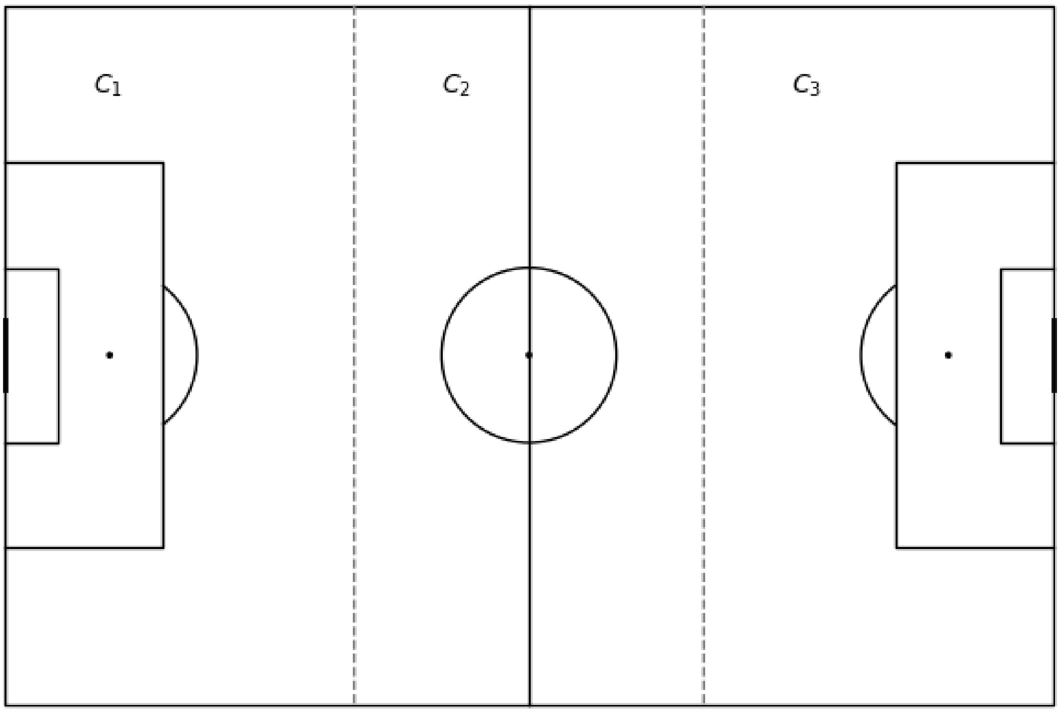
Van Roy, Robberechts, et al. (2021)

(2021)

1 Formulación

El Proceso de Decisión de Markov se compone de los siguientes elementos:

- El conjunto de estados estará conformado por las siguientes tres divisiones del campo



Donde C_1, C_2, C_3 representa que el balón se encuentre en alguna de las tres divisiones.

Además se agregan tres estados absorbentes:

- L_p = pérdida de posesión del balón.
- nG = realizar un tiro y que no termine en gol.
- G = realizar un tiro y que termine en gol.

De esta forma el conjunto de estados $\mathcal{S}$ queda como

$$\mathcal{S} = \{C_1, C_2, C_3, L_p, nG, G\}$$

- El conjunto de acciones admisibles se considerarán 3 acciones que serán

  - t = tiro
  - p = pase
  - r = regate

  De esta forma el conjunto de acciones queda como

$$\mathcal{A} = \{t, p, r\}$$

- Para las de transiciones haremos uso de las probabilidades de transición definidas de la siguiente forma:

$$P : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$$

  Que se interpreta como la probabilidad de estar en un estado S_i realizar una acción a_k y terminar en un estado S_j. Notemos que se aceptan los casos cuando $i = j$ y más adelante se mostrará que algunas probabilidades serán 0.

- La función de recompensa $R : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$, será

$$R(S_i, a_k, S_j) = \begin{cases} 1 & \text{si} \quad S_j = G \\ 0 & o.c. \end{cases}$$

2 Dinámica del modelo

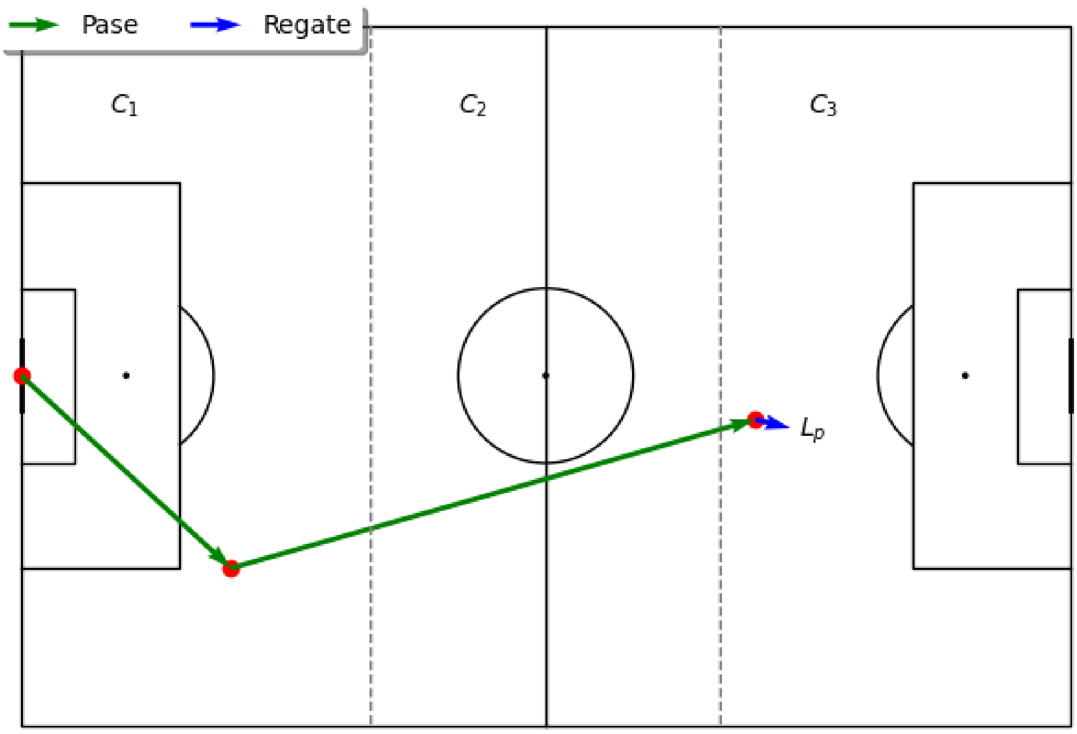
En el contexto del fútbol llamamos *jugada* a una sucesión de acciones donde el balón se traslada desde un punto inicial donde el equipo A tiene el balón hasta un punto final que puede ser: perder el balón, tirar a puerta y no anotar gol o tirar y anotar gol.

Ejemplo: El balón comienza en el saque de meta del portero, el portero da un pase a un defensa que se encuentra en el primer tercio, que esté da un pase a un delantero que se encuentra en el tercer tercio y al intentar un regate pierde el balón.

En nuestro contexto se verá como el hecho de iniciar la sucesión de acciones desde alguna sección C_i y terminar en alguno de los 3 estados absorbentes.

Ejemplo

$$C_1 \xrightarrow{p} C_1 \xrightarrow{p} C_3 \xrightarrow{r} L_p.$$



Para movernos de un estado S_i a un estado S_j mediante una acción a_k haremos uso de las probabilidades de transición, estas probabilidades las estimaremos utilizando datos extraídos de FBREF para 4 clubes: Chivas, América, Cruz Azul y Pumas de la temporada 2023-2024

Primero vamos a interpretar las probabilidades de transición con la finalidad de descartar aquellas transciones que no serán posibles con nuestro modelo y con la naturaleza de un partido.

- Fijamos el estado C_1.
  - Consideramos la acción p:

    $P(C_1, p, C_1)$ = La probabilidad de estar en la zona C_1 dar un *pase* y terminar en la zona C_1.

    $P(C_1, p, C_2)$ = La probabilidad de estar en la zona C_1 dar un *pase* y terminar en la zona C_2.

    $P(C_1, p, C_3)$ = La probabilidad de estar en la zona C_1 dar un *pase* y terminar en la zona C_3.

    $P(C_1, p, L_p)$ = La probabilidad de estar en la zona C_1 dar un *pase* y perder el balón.

    $P(C_1, p, nG)$ = La probabilidad de estar en la zona C_1 dar un *pase* y no anotar gol.

    $P(C_1, p, G)$ = La probabilidad de estar en la zona C_1 dar un *pase* y anotar gol.

    De esta lista de probabilidades para la acción p notemos que $P(C_1, p, nG) = P(C_1, p, G) = 0$ esto pues la única acción admisible para terminar en los estados $\{nG, G\}$ es t.
  - Consideramos la acción r:

    $P(C_1, r, C_1)$ = La probabilidad de estar en la zona C_1 hacer un *regate* y terminar en la zona C_1.

    $P(C_1, r, C_2)$ = La probabilidad de estar en la zona C_1 hacer un *regate* y terminar en la zona C_2.

    $P(C_1, r, C_3)$ = La probabilidad de estar en la zona C_1 hacer un *regate* y terminar en la zona C_3.

    $P(C_1, r, L_p)$ = La probabilidad de estar en la zona C_1 hacer un *regate* y perder el balón.

    $P(C_1, r, nG)$ = La probabilidad de estar en la zona C_1 hacer un *regate* y no anotar gol.

    $P(C_1, r, G)$ = La probabilidad de estar en la zona C_1 hacer un *regate* y anotar gol.

De esta lista de probabilidades para la acción r notemos que $P(C_1, r, nG) = P(C_1, r, nG) = 0$ esto pues la única acción admisible para terminar en los estados nG, G es t.

Además $P(C_1, r, C_3) = 0$, pues al realizar un *regate* solo tenemos dos opciones:

nos mantenemos en la zona C_1 o avanzamos a la zona siguiente C_2.

- Consideramos la acción t:

  En este caso tendremos que $P(C_1, t, C_1) = P(C_1, t, C_2) = P(C_1, t, C_3) = P(C_1, t, L_p) = 0$ pues después de realizar un *tiro* solo tendremos dos estados posibles $\{nG, G\}$.

  $P(C_1, t, nG) =$ La probabilidad de estar en la zona C_1 hacer un *tiro* y no anotar gol.

  $P(C_1, t, G) =$ La probabilidad de estar en la zona C_1 hacer un *tiro* y anotar gol.

  Sin embargo, como suponemos que las acciones que toman las futbolistas son razonables, no tiene sentido realizar un *tiro* desde la zona C_1, pues la distancia hacia la porteria contaría es muy lejana y la probabilidad de anotar un gol es prácticamente nula. Por tanto

$$P(C_1, t, S) = 0, \quad \forall S \in \mathcal{S}.$$

- Fijamos el estado C_2.

  - Consideramos la acción p:

    $P(C_2, p, C_1) =$ La probabilidad de estar en la zona C_2 dar un *pase* y terminar en la zona C_1.

    $P(C_2, p, C_2) =$ La probabilidad de estar en la zona C_2 dar un *pase* y terminar en la zona C_2.

    $P(C_2, p, C_3) =$ La probabilidad de estar en la zona C_2 dar un *pase* y terminar en la zona C_3.

    $P(C_2, p, L_p) =$ La probabilidad de estar en la zona C_2 dar un *pase* y perder el balón.

    $P(C_2, p, nG) =$ La probabilidad de estar en la zona C_2 dar un *pase* y no tirar a gol.

    $P(C_2, p, G) =$ La probabilidad de estar en la zona C_2 dar un *pase* y terminar en gol.

    De esta lista de probabilidades para la acción p notemos que $P(C_2, p, nG) = P(C_2, p, nG) = 0$ esto pues la única acción admisible para terminar en los estados $\{nG, G\}$ es t.

- Consideramos la acción r:

  $P(C_2, r, C_1) =$ La probabilidad de estar en la zona C_2 hacer un *regate* y terminar en la zona C_1.

  $P(C_2, r, C_2) =$ La probabilidad de estar en la zona C_2 hacer un *regate* y terminar en la zona C_2.

  $P(C_2, r, C_3) =$ La probabilidad de estar en la zona C_2 hacer un *regate* y terminar en la zona C_3.

  $P(C_2, r, L_p) =$ La probabilidad de estar en la zona C_2 hacer un *regate* y perder el balón.

  $P(C_2, r, nG) =$ La probabilidad de estar en la zona C_2 hacer un *regate* y no anotar gol.

  $P(C_2, r, G) =$ La probabilidad de estar en la zona C_2 hacer un *regate* y anotar gol.

  De esta lista de probabilidades para la acción r notemos que $P(C_2, r, nG) = P(C_2, r, G) = 0$ esto pues la única acción admisible para terminar en los estados $\{nG, G\}$ es t.

- Consideramos la acción t:

  En este caso tendremos que $P(C_2, t, C_1) = P(C_2, t, C_2) = P(C_2, t, C_3) = P(C_2, t, L_p) = 0$ pues después de realizar un *tiro* solo tendremos dos estados posibles $\{nG, G\}$.

  $P(C_2, t, nG) =$ La probabilidad de estar en la zona C_1 hacer un *tiro* y no anotar gol.

  $P(C_2, t, G) =$ La probabilidad de estar en la zona C_1 hacer un *tiro* y anotar gol.

• Fijamos el estado C_3

- Consideramos la acción p:

  $P(C_3, p, C_1) =$ La probabilidad de estar en la zona C_3 dar un *pase* y terminar en la zona C_1.

  $P(C_3, p, C_2) =$ La probabilidad de estar en la zona C_3 dar un *pase* y terminar en la zona C_2.

  $P(C_3, p, C_3) =$ La probabilidad de estar en la zona C_3 dar un *pase* y terminar en la zona C_3.

  $P(C_3, p, L_p) =$ La probabilidad de estar en la zona C_3 dar un *pase* y perder el balón.

  $P(C_3, p, nG) =$ La probabilidad de estar en la zona C_3 dar un *pase* y no tirar a gol.

$P(C_3, p, G) =$ La probabilidad de estar en la zona C_3 dar un *pase* y terminar en gol.

De esta lista de probabilidades para la acción p notemos que $P(C_3, p, nG) = P(C_2, p, G) = 0$ esto pues la única acción admisible para terminar en los estados $\{nG, G\}$ es t.

– Consideramos la acción r:

$P(C_3, r, C_1) =$ La probabilidad de estar en la zona C_3 hacer un *regate* y terminar en la zona C_1.

$P(C_3, r, C_2) =$ La probabilidad de estar en la zona C_3 hacer un *regate* y terminar en la zona C_2.

$P(C_3, r, C_3) =$ La probabilidad de estar en la zona C_3 hacer un *regate* y terminar en la zona C_3.

$P(C_3, r, L_p) =$ La probabilidad de estar en la zona C_3 hacer un *regate* y perder el balón.

$P(C_3, r, nG) =$ La probabilidad de estar en la zona C_3 hacer un *regate* y no anotar gol.

$P(C_3, r, G) =$ La probabilidad de estar en la zona C_3 hacer un *regate* y anotar gol.

De esta lista de probabilidades para la acción r notemos que $P(C_3, r, nG) = P(C_3, r, G) = 0$ esto pues la única acción admisible para terminar en los estados $\{nG, G\}$ es t.

Además $P(C_3, r, C_1) = 0$, pues al realizar un *regate* solo tenemos dos opciones o nos mantenemos en la zona C_3 o retrocedemos a la zona anterior C_2.

– Consideramos la acción t:

En este caso tendremos que $P(C_3, t, C_1) = P(C_3, t, C_2) = P(C_3, t, C_3) = P(C_3, t, L_p) = 0$ pues después de realizar un *tiro* solo tendremos dos estados posibles $\{nG, G\}$.

$P(C_3, t, nG) =$ La probabilidad de estar en la zona C_1 hacer un *tiro* y no anotar gol.

$P(C_3, t, G) =$ La probabilidad de estar en la zona C_1 hacer un *tiro* y anotar gol.

- Por último como $\{L_p, nG, G\}$ son estados absorbentes, entonces $\forall a \in \mathcal{A}$.

$$P(L_p, a, S) = \begin{cases} 1 & \text{si} \quad S = L_p \\ 0 & o.c. \end{cases}$$

$$P(nG, a, S) = \begin{cases} 1 & \text{si} \quad S = nG \\ 0 & o.c. \end{cases}$$

$$P(G,a,S) = \begin{cases} 1 & \text{si} \quad S = G \\ 0 & o.c. \end{cases}$$

Para facilitar las estimaciones se renombrarán cada probabiidad con los parámetros que se muestran en las siguientes tablas

Tabla 2.1: Parámetros para las probabilidades de C_1

Probabilidades	Parámetros
$P(C_1, p, C_1)$	α_{1p}
$P(C_1, p, C_2)$	α_{2p}
$P(C_1, p, C_3)$	α_{3p}
$P(C_1, p, L_p)$	α_{4p}
$P(C_1, r, C_1)$	α_{1r}
$P(C_1, r, C_2)$	α_{2r}
$P(C_1, r, L_p)$	α_{3r}

Tabla 2.2: Parámetros para las probabilidades de C_2

Probabilidades	Parámetros
$P(C_2, p, C_1)$	β_{1p}
$P(C_2, p, C_2)$	β_{2p}
$P(C_2, p, C_3)$	β_{3p}
$P(C_2, p, L_p)$	β_{4p}
$P(C_2, r, C_1)$	β_{1r}
$P(C_2, r, C_2)$	β_{2r}
$P(C_2, r, C_3)$	β_{3r}
$P(C_2, r, L_p)$	β_{4r}
$P(C_2, t, nG)$	β_{1t}
$P(C_2, t, G)$	β_{1t}

Tabla 2.3: Parámetros para las probabilidades de C_3

Probabilidades	Parámetros
$P(C_3, p, C_1)$	γ_{1p}
$P(C_3, p, C_2)$	γ_{2p}
$P(C_3, p, C_3)$	γ_{3p}
$P(C_3, p, L_p)$	γ_{4p}
$P(C_3, r, C_2)$	γ_{1r}
$P(C_3, r, C_3)$	γ_{2r}
$P(C_3, r, L_p)$	γ_{3r}
$P(C_3, t, nG)$	γ_{1t}

Probabilidades	Parámetros
$P(C_3, t, G)$	γ_{2t}

3 Descripción y Justificación de la recompensa

En un partido de fútbol gana el equipo que anota más goles, en caso de anotar los mismos goles se considera empate y no existe desempate de ningún tipo.[1] Por lo que la recompensa será la de anotar un gol $R = 1$, pues es lo único que puede hacer que un equipo gane un partido. No existe penalización porque los goles válidos anotados no pueden ser descontandos.

[1] No se consideran los partidos de eliminación directa donde existe el desempate por penales.

4 Justificación de las acciones

Las acciones que puede realizar un equipo durante un partido son limitadas y se pueden enlistar. Sin embargo para nuestro modelo vamos a seleccionar las 3 más importantes que son el *pase, tiro* y *regate*.

- *tiro*: Es la acción que nos permite anotar goles.
- *pase*: Ayuda a un equipo a mover el balón por el campo sin necesidad de desplazarse o dejar rivales atrás.
- *regate*: Permite que podamos trasladar el balón de un lugar a otro y dejar a rivales atrás.

References

Van Roy, Maaike, Pieter Robberechts, Wen-Chi Yang, Luc De Raedt, y Jesse Davis. 2021. “Leaving goals on the pitch: Evaluating decision making in soccer”. *arXiv preprint arXiv:2104.03252*.

Van Roy, Maaike, Wen-Chi Yang, Luc De Raedt, y Jesse Davis. 2021. “Analyzing learned markov decision processes using model checking for providing tactical advice in professional soccer”. En *AI for Sports Analytics (AISA) Workshop at IJCAI 2021*.