DEPARTAMENTO DE COMPUTACION
Facultad de Ciencias Exactas y Naturales - UBA

Sistemas de ecuaciones lineales

24 de septiembre de 2020 — Métodos numéricos

Integrante	LU	Correo electrónico
Florencia Zanollo	934/11	florenciazanollo@gmail.com
Luis García Gómez	675/13	garciagomezluis.94@gmail.com

Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Ciudad Universitaria - Pabellón I
Intendente Güiraldes 2160 - C1428EGA
Ciudad Autónoma de Buenos Aires - Argentina
Tel/Fax: (54 11) 4576-3359
http://exactas.uba.ar

Sistemas de ecuaciones lineales

En la actualidad, debido a la importancia y magnitud de los eventos deportivos, a veces es necesario un sistema de rankeo que funcione a pesar de tener pocos encuentros, teniendo en cuenta la dificultad del adversario, pero sin bias previo. Este trabajo trata de la resolución de sistemas de ecuaciones lineales en el contexto específico del *Colley Matrix Method* para ranking de equipos. Se busca analizar la utilidad de este método, su estabilidad respecto a los límites de la aritmética finita de las computadoras y su justicia a la hora de rankear equipos. Utilizando como punto de comparación otros métodos como *porcentaje de victorias* (WP) y *Elo rating*, sobre los resultados de partidos de ATP 2015 y NBA 2016. Se muestra que las características de este método lo hacen justo bajo ciertas normas y el error aritmético generado es chico.

Palabras claves: Eliminación Gaussiana, Ranking, Colley Matrix Method, Ecuaciones Lineales.

Índice

1. Introducción

El objeto de estudio de este trabajo es Colley Matrix Method (CMM) para generar rankings en competiciones deportivas. En virtud de entenderlo mejor, empezaremos introduciéndolo para luego plantear algunas preguntas sobre cuestiones aritméticas.

Vamos a presentar una implementación en C++ que nos va a permitir en base a un dataset propuesto por la cátedra elaborar una tabla de posiciones que podría utilizarse luego para determinar que equipos ganan una competencia y cuales de ellos están en condiciones de clasificar a otras instancias.

Desde el punto de vista del análisis cuantitativo presentaremos el error absoluto del método utilizando casos específicos y tomando como referencia implementaciones ya desarrolladas más estables numéricamente.

Desde el punto de vista del análisis cualitativo presentaremos un estudio por planteo de casos de los cuales podremos derivar algunas características del método que explicaremos cuando hagamos la interpretación del sistema matricial que lo determina. Consecuentemente podremos plantear una estrategia para explotarlo y determinar en qué casos podría ser injusto un evento que lo utilice.

Finalmente haremos una comparación de CMM contra otros sistemas de rankeo, como en este caso son WP (porcentaje de victorias) y Elo para encontrar similitudes y diferencias.

1.1. Problema

Una forma intuitiva de establecer un sistema de puntaje para un evento es utilizar el estimador que relaciona la cantidad de partidos ganados con la cantidad de partidos totales: estamos hablando del *porcentaje de victorias* (WP)

$$r_i = \frac{w_i}{n_i} \tag{1}$$

Donde w_i denota la cantidad de partidos ganados, n_i la cantidad de partidos jugados y r_i es el ratio buscado.

Por un lado, puede pasar que para un equipo, $r_i = 1$ habiendo jugado y ganado un partido mientras que para otro $r_j = 0{,}7$ habiendo jugado diez y ganado siete. Por otra parte, WP, no contempla en su desarrollo características de los oponentes o cuantas veces jugó contra ellos, dado que no es lo mismo jugar una vez contra el número 1 que hacerlo cuatro veces.

Este es el problema que intenta resolver Colley Matrix Method.

1.2. Colley Matrix Method (CMM)

El desarrollo de CMM se basa en la *Regla de Laplace de sucesos*, es decir, en cómo $\frac{1+w_i}{2+n_i}$ es mejor estimador para situaciones de las que se sabe poco que $\frac{w_i}{n_i}$.

Observemos que

$$w_i = \frac{(w_i - l_i)}{2} + \frac{n_i}{2} \tag{2}$$

Por lo que si tomamos el nuevo estimador, entonces

$$r_i = \frac{1+w_i}{2+n_i} = \frac{1 + \frac{w_i - l_i}{2} + \frac{n_i}{2}}{2+n_i} = \frac{1 + \frac{w_i - l_i}{2} + \sum_{i=1}^{n_i} r_j^i}{2+n_i} \tag{3}$$

Si asumimos que inicialmente todos los equipos tiene un $r_i = \frac{1}{2}$, entendiendo a l_i como la cantidad de partidos perdidos por el equipo i y a r_j^i como la cantidad de partidos jugados entre el equipo i y el equipo j.

Dado que en un evento hay n equipos, cada uno de los cuales con su propio *rating* r_i, notemos que:

$$(2+n_i) * r_i - \sum_{i=1}^{n_i} r_j^i = 1 + \frac{w_i - l_i}{2} \tag{4}$$

Obtenemos un sistema lineal de ecuaciones $Cr = b$ donde r es el vector incógnita con los ratings y

$$b = 1 + \frac{w_i - l_i}{2} \tag{5}$$

$$C_{i,j} = \begin{cases} -n_{i,j} \text{ si } i \neq j \\ 2 + n_i \text{ si } i = j \end{cases} \tag{6}$$

¿Pero este sistema siempre tiene solución? claramente, en una situación en la que es necesario rankear equipos siempre debería poder lograrse. ¿Habrán situaciones para las cuales no se podrá hacer? ¿Elaborará un ranking razonable?

2. Desarrollo

Existen múltiples formas en las que se puede resolver algorítmicamente un sistema de ecuaciones lineales como el nuestro, del tipo $Cr = b$. Con distintas complejidades al implementar y distintas características, entre ellas se encuentran:

- **Eliminación Gaussiana sin pivoteo**: Este es uno de los métodos métodos más sencillos desde el punto de vista teórico. Consiste en derivar un sistema como el nuestro en un sistema con una matriz triangular para que sea fácil de resolver. Como ventaja: es de fácil implementación; como desventaja: no todos los sistemas pueden resolverse mediante Eliminación Gaussiana sin pivoteo y es susceptible a errores numéricos. Entre otras cosas, consta con una complejidad de $O(n^3)$.
- **Factorización LU**: Esencialmente se apoya en el método anterior. Consiste en obtener una factorización $C = LU$ donde L es triangular inferior con unos en la diagonal y U es triangular superior. Como ventaja respecto a la Eliminación Gaussiana sin pivoteo es que mediante la resolución de dos sistemas de ecuaciones simples se puede lograr una complejidad de $O(n^2)$. Como desventaja: aún no todos los sistemas pueden resolverse utilizando este método.
- **Factorización de Cholesky**: Consiste en encontrar una factorización $C = LL^t$ donde L es una matriz triangular inferior. Una vez se obtiene la factorización se puede resolver el sistema en cuestión utilizando, al igual que antes, dos sistemas intermedios. La diferencia con los métodos previos es que hay un algoritmo que permite hallar las matrices de estos sistemas sin Eliminación Gaussiana. Como ventaja: es más eficiente para representar sistemas de ecuaciones en memoria y cuenta con una complejidad de $O(n^2)$ aunque con mejores constantes que el Método de Factorización LU. Por otra parte, tiene una mayor estabilidad numérica. Como desventaja: no todos los sistemas pueden resolverse mediante este método, sólo aquellos cuya matriz principal sea *simétrica definida positiva.*

Todos los métodos se complementan con el algoritmo de *backward substitution* ($O(n^2)$) para resolver el sistema.

2.1. Elección del algoritmo

Nos fue requerido resolver el sistema mediante Eliminación Gaussiana sin pivoteo (EG). Observemos que como la matriz C de nuestro sistema es simétrica definida positiva [1] cumple con las condiciones para resolverse con cualquiera de los métodos mencionados.

2.2. Algunas características de nuestro sistema

Para ello, analicemos primero cómo es la matriz C de nuestro sistema y por qué cumple con las condiciones para resolverse mediante el método mencionado.

- **De tener solución, es única**: dado que contamos con $n = \#\{equipo\}$ ecuaciones y $n = dim(r)$ incógnitas.
- **Es simétrica**: ya que la cantidad de partidos jugados entre el equipo i y el equipo j es la cantidad de partidos jugados por el equipo j e i.
- **Es estrictamente diagonal dominante**: Por definición $n_i = \sum_{j \neq i} n_{i,j}$, por lo que es claro que $|C_{i,i}| > \sum_{j \neq i} |C_{i,j}|$ ya que $C_{i,i} = 2 + n_i$.

Al ser C estrictamente diagonal dominante tiene, entre otras, las siguientes propiedades:

- Es no singular
- Sus submatrices principales también son estrictamente diagonal dominantes

Por otro lado, sabemos que si las submatrices principales de una matriz son no singulares, entonces la misma tiene factorización LU, lo que implica la correcta aplicación de la Eliminación Gaussiana sin pivoteo.

Se implica directamente que a C **se le puede aplicar la Eliminación Gaussiana sin pivoteo**.

2.3. Sobre la estabilidad numérica de la Eliminación Gaussiana sin pivoteo

La aritmética finita de las computadoras trae aparejados problemas para representar nuestro sistema numérico. En el caso de la Eliminación Gaussiana sin pivoteo podría presentarse al realizar una división por algún elemento de la diagonal de la matriz en cuestión que esté muy cercano a cero. En tal caso cualquier error se vería amplificado. No obstante el ser C una matriz estrictamente diagonal dominante y simétrica nos asegura tener en la diagonal el elemento más grande en módulo tanto por fila como por columna, por lo que el pivoteo parcial en un caso como este sería redundante y se logra una cierta estabilidad numérica.

No obstante, las pruebas de control sobre la implementación acotan el error por 10^{-4}.

2.4. Sobre WP y Elo

Como comentamos en la introducción, WP es una de las razones por las que CMM existe. No obstante, Elo fue creado de forma independiente y hoy es utilizado considerablemente en el ajedrez competitivo y se ha extendido a otras ramas de actividades.

Aunque a nivel implementativo no son muy interesantes en este trabajo, si vamos a ahondar en la cuestión cualitativa de ambos para compararlos con CMM 3.7.

3. Experimentación

En esta sección vamos a presentar, explicar y mostrar resultados de cada uno de los experimentos considerados.

Desde el punto de vista cuantitativo, vamos a analizar el error absoluto de CMM.

3.1. Error absoluto

Como se explicó en 2.3 debido a la aritmética finita de la computadora, las operaciones pueden llegar a ser muy inestables [2].

Un pequeño ejemplo:

$$\begin{pmatrix} 5 & -1 & -1 & -1 \\ -1 & 4 & 0 & -1 \\ -1 & 0 & 5 & -2 \\ -1 & -1 & -2 & 6 \end{pmatrix}$$

Luego de correr EG debería quedar:

$$\begin{pmatrix} 5 & -1 & -1 & -1 \\ 0 & \frac{19}{5} & -\frac{1}{5} & -\frac{6}{5} \\ 0 & 0 & \frac{91}{19} & -\frac{43}{19} \\ 0 & 0 & 0 & \frac{396}{91} \end{pmatrix}$$

Pero debido a la falta de precisión de los double queda:

$$\begin{pmatrix} 5 & -1 & -1 & -1 \\ 0 & 3{,}8 & -0{,}2 & -1{,}2 \\ 0 & 0 & 4{,}78947 & -2{,}26316 \\ 0 & 0 & 0 & 4{,}35165 \end{pmatrix}$$

Se puede ver que $\frac{396}{91} \cong 4{,}351648352$ teniendo una diferencia de 0,00000165 con 4,35165. Difieren poco pero es un caso chico, este error se podría repetir y acarrear terminando en un valor en esa posición muy diferente al esperado, afectando así el ranking final.

Para experimentar sobre esto construimos una serie de partidos ficticios con el propósito de que la matriz quede con *valores diferentes en su diagonal*, posiblemente este no sea el peor caso de error pero suponemos que es suficientemente malo. Para calcular el error absoluto, debido a que no contamos con el valor de control, utilizamos la factorización de Cholesky ya que es más estable. Cabe destacar que este método se puede aplicar dado que la matriz resulta simétrica y definida positiva[1].

Yendo en detalle hicimos que cada equipo (1 a total-1) juegue *id* cantidad de partidos, es decir, el equipo 1 juega un partido, el 2 dos, etc. Salvo el último equipo que va a jugar la cantidad que sea necesaria para cumplir con los partidos de los demás.

Como se puede observar en la figura 1 aumenta el error a medida que aumenta la cantidad de equipos, lo cuál es esperable ya que se realizan más despejes, cuentas y se acarrea más el error.

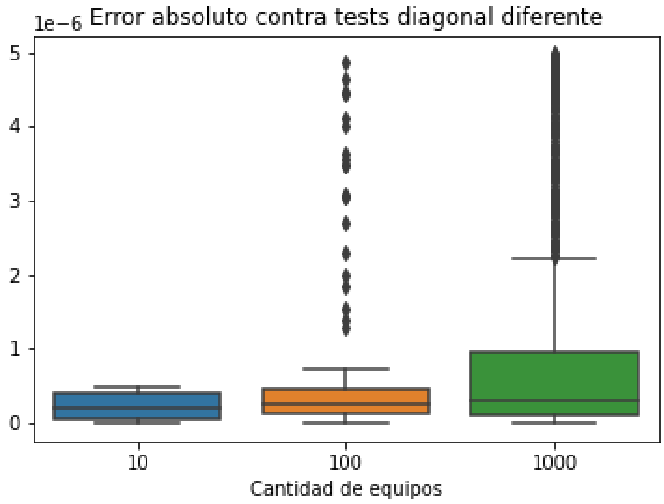

Figura 1: Diferencia de error vs test diagonal diferente

Luego, simplemente por probar otros casos decidimos utilizar el batch de tests generado por la cátedra. Al obtener los resultados notamos que los *tests completos* generaban un error absoluto promedio mayor al que obtuvimos con nuestra *diagonal diferente*, ver figura 3.

Lo que suponíamos que podía estar pasando es que al haber muchos -1 al comienzo y no tantos 0s el error generado termina siendo mayor. Para probar esto hicimos nuestros propios tests de *todos vs todos*, para diferentes cantidades de equipos. Pero al final, como se puede apreciar en la figura 2, no pudimos llegar a un error tan grande como los de la cátedra. Suponemos que esto se debe a que nuestro oráculo no es tan estable como pensábamos y queda abierta la pregunta de cómo fueron calculados los resultados esperados para los casos de la cátedra.

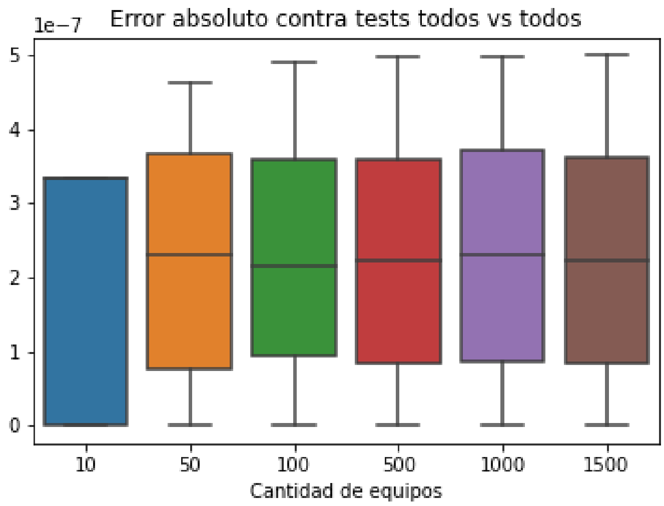

Figura 2: Diferencia de error vs test todos vs todos

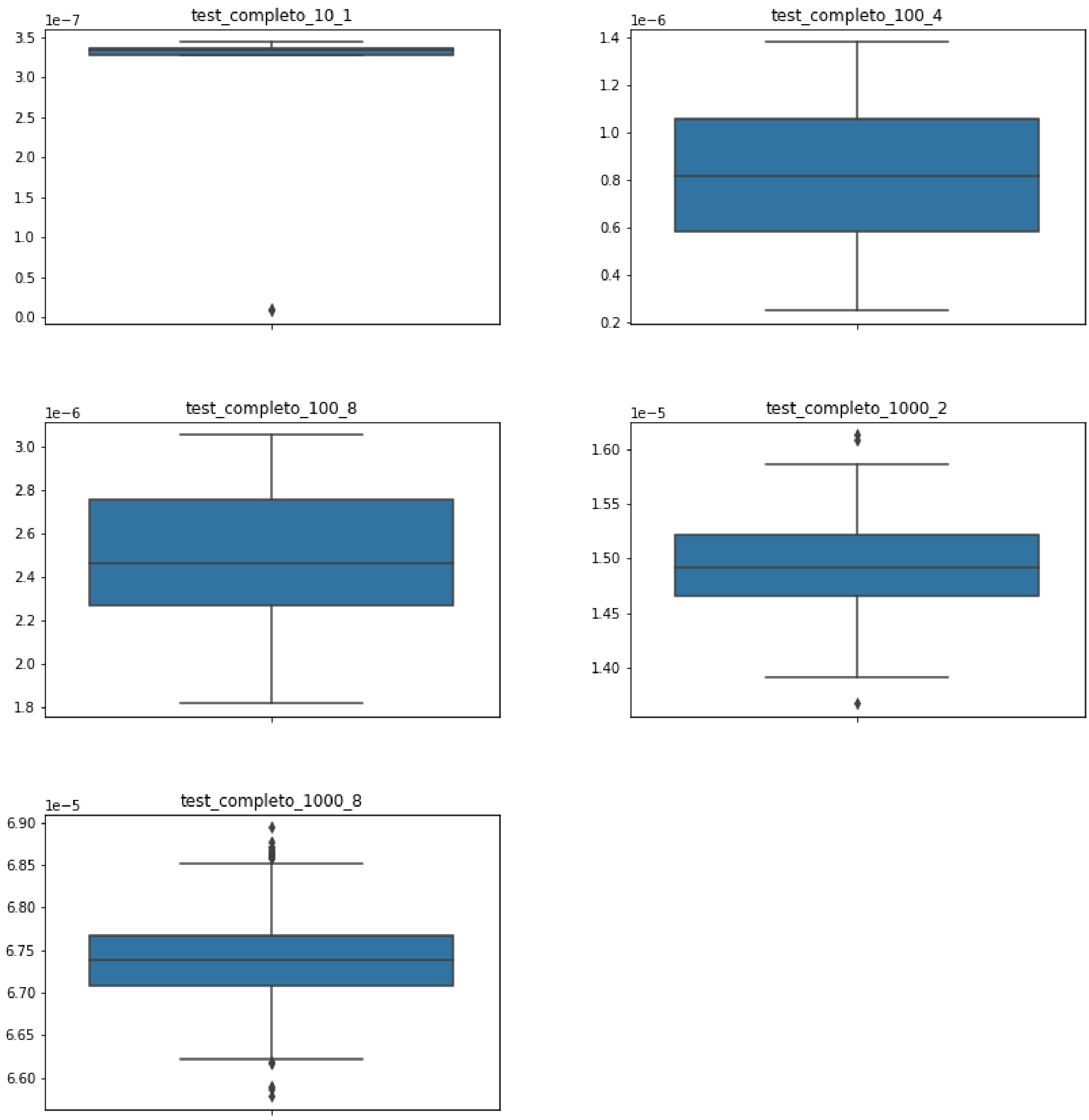

Figura 3: Diferencia de error vs test completos de la catedra

Desde la perspectiva cualitativa, vamos a plantear y responder preguntas que nos lleven a entender mejor CMM y a encontrar ventajas y desventajas del mismo. Para ello primeramente propondremos casos pequeños, que no necesariamente representan casos de competencias reales.

3.2. Planteo de casos

Caso 1

Se inicia un sistema con dos equipos: 1, y 2, con un único partido entre ellos: 1 vs 2 a favor de 1.
Pregunta: ¿Qué pasará con los ratings? ¿Hay alguna relación entre ellos?

Posición	Equipo	Rating
1	1	0.625
2	2	0.375

Cuadro 1: Ranking CMM luego de iniciar un partido 1 vs 2, a favor de 1

Observación: Acorde a la *Regla de Laplace de sucesos* utilizar el estimador planteado por la misma nos permite que no hayan cambios abruptos en la resolución de los ratings de los equipos involucrados.

Caso 2

Sobre el sistema anterior: se agregan cinco partidos a favor de 1.
Pregunta: ¿Qué pasará con los ratings? ¿Hay alguna relación entre ellos?

Posición	Equipo	Rating
1	1	0.714286
2	2	0.285714

Cuadro 2: Ranking CMM luego de cinco partidos 1 vs 2, a favor de 1

Observación: Los ratings se ajustan acorde a los resultados de los partidos. Como vimos hasta ahora, los mismos dependen de la cantidad de partidos que hayan entre los equipos competidores.

Caso 3

Sobre el sistema anterior: se suma un equipo, 3. 3 vs 2 a favor de 3.
Pregunta: ¿Será mejor el rating de 3 que el de 2?

Posición	Equipo	Rating
1	1	0.68
2	3	0.58
3	2	0.24

Cuadro 3: Ranking CMM luego de introducir a 3 y agregar un partido a favor de 3

Observación: Después de 3 vs 2 a favor de 3, 3 supera fácilmente a 2 modificando sus posiciones. El rating de 2 desciende esperadamente, pero el de 1 también aunque sigue manteniendo su posición. Al no estar 1 involucrado en este partido, vemos que CMM afecta el rating (y tal vez el ranking) de equipos que no formaron parte del partido.

Caso 4

Sobre el sistema anterior: se agregan diez partidos, 3 vs 2 a favor de 3.

Pregunta: ¿Se acercará el rating de 3 al rating de 1?

Posición	Equipo	Rating
1	3	0.662963
2	1	0.644444
3	2	0.192593

Cuadro 4: Ranking CMM luego de agregar diez partidos a favor de 3

Observación: En este caso se ve que CMM permite alterar no solo el rating, si no la posición de los equipos, ya que 3 vs 1 nunca tuvo lugar. Esto de alguna forma nos habla de "la justicia del método": ¿es justo que 1 descienda si nunca se enfrentó con 3?. Claramente la relación entre ambos equipos viene dada por los partidos que tuvieron de forma separada con 2. 1 vs 2 se dió hasta ahora seis veces mientras que 3 vs 2 se dió hasta ahora once veces.

Caso 5

Sobre el sistema anterior: se agregan cinco partidos, 1 vs 2 a favor de 1.

Pregunta: Si la relación entre 1 y 3 es intermediada por la cantidad de partidos que ambos ganaron a 2, ¿será cierto que ambos tendrán el mismo rating si igualan la cantidad de partidos con 2?

Posición	Equipo	Rating
1	1	0.657143
1	3	0.657143
2	2	0.185714

Cuadro 5: Ranking CMM luego de que 1 y 3 obtengan la misma configuración de partidos

Observación: En este caso se ve que a igual cantidad de partidos ganados de 1 y 3 vs 2, se obtiene el mismo rating.

Caso 6

Sobre el sistema anterior: se suma un equipo, 4. 4 vs 3 a favor de 4.

Pregunta: ¿Qué pasará si un nuevo equipo tiene un partido contra el equipo que está en la primera posición?

Posición	Equipo	Rating
1	4	0.692159
2	1	0.606041
3	3	0.576478
4	2	0.125321

Cuadro 6: Ranking CMM luego de agregar al sistema anterior un nuevo equipo con un partido a favor

Observación: En un solo partido, 4 llegó a la primera posición. Se puede explicar que 1 haya quedado por arriba de 3 ya que la configuración de partidos jugados es un factor que está presente en CMM: 4 vs 3 deja a 3 con un partido más, pero perdido, a diferencia de 1 que no sumó partidos.

Caso 7

Sobre el sistema anterior: se agregan cien partidos, 2 vs 3 a favor de 2.

Pregunta: ¿Puede alcanzar 2 las primeras posiciones si gana partidos contra los equipos de las últimas posiciones?

Posición	Equipo	Rating
1	1	0.905919
2	4	0.52859
3	2	0.479722
4	3	0.0857696

Cuadro 7: Ranking CMM luego de agregar cien partidos a favor de 2

Observación: Se observa que por un lado 2 y 3 intercambiaron posiciones, lo que muestra que, aunque los ajustes de rating se hagan con cada partido, los máximos valores que pueden tener se ven limitados a los valores de los equipos que se enfrentaron. Por otro lado 1 y 4 intercambiaron lugares también: esto puede deberse a que hasta este momento 3 y 4 se enfrentaron en una ocasión a favor de 4, es probable que su caída en el puesto sea debido a la caída de 3 (su victoria se desprestigia).

Caso 8

Sobre el sistema anterior: se agregan tantos partidos como sea necesario para desplazar a 1 de la primera posición, 3 vs 1 a favor de 3.

Pregunta: Estando 3 en la última posición, ¿se podrá desplazar a 1 de la primera posición en pocos partidos?

Posición	Equipo	Rating
1	2	0.597924
2	4	0.583805
3	1	0.566854
4	3	0.251416

Cuadro 8: Ranking CMM luego de siete partidos 3 vs 1 a favor de 3

Observación: En siete partidos, 3 fue capaz mediante CMM de que 2 reemplazara a 1 en el primer lugar. ¿Por qué 2 y no 4? Es probable que sea debido a la cantidad de partidos 1 vs 2 que se hayan dado.

Caso 9

Sobre el sistema del caso 7 3.2: se agregan tantos partidos como sea necesario para desplazar a 1 de la primera posición, 3 vs 2 a favor de 3.

Pregunta: Estando 3 en la última posición, ¿se podrá desplazar a 1 de la primera posición en pocos partidos?

Posición	Equipo	Rating
1	1	0.858055
2	4	0.554697
3	2	0.423156
4	3	0.164091

Cuadro 9: Ranking CMM luego de veinte partidos 3 vs 2 a favor de 3

Observación: En veinte partidos, 3 aún no fue capaz mediante CMM de que un equipo reemplazara a 1 en el primer lugar.

3.3. ¿Qué podemos decir sobre CMM en base a las observaciones?

Estos ejemplos sencillos nos permitieron entender un poco más CMM a la hora de rankear equipos.

Desde el caso 1 al caso 5, probamos introducir un equipo y hacerlo avanzar compitiendo con el equipo de la última posición. Vimos que puede alcanzar un mejor puesto en el ranking tratando de obtener una configuración similar a la de los equipos que están en la posición deseada, lo cual es esperado para un ranking en general. Vimos también que el resultado de un partido influye sobre otros equipos mediante una "transitividad"(caso 4).

Desde el caso 6 al caso 9, quisimos hacer foco en como un equipo puede avanzar en la tabla. Ingresamos un participante más a nuestro sistema, dejando un total de cuatro y volviéndolo un poco más complejo, corroboramos que los equipos son susceptibles a cambios en su rating si algún equipo con el que tuvo relación tiene un partido. Por otra parte, notamos que un equipo en la tabla avanza más lento conforme se enfrenta a equipos que están en las últimas posiciones y más rápido conforme se enfrenta a otros que está en las primeras:

- en el caso 7: 3 vs 2, favor a 3. Después de cien partidos 2 había avanzado una posición.
- en el caso 8: 3 vs 1, favor a 3. En siete partidos se logra que 2 desplace a 1.
- en el caso 9: 3 vs 2, favor a 3. En veinte partidos no se logró mover a 1 del primer lugar.

Tiene sentido esto último ya que tal vez pueda considerarse que no hay suficiente mérito en vencer a equipos que están en las últimas posiciones.

3.4. Interpretación del sistema $Cr = b$ y una explicación a las observaciones

Como se desarrolla en [1], la solución al sistema $Cr = b$ son los ratings de los equipos involucrados en la competencia. Saliendo de su expresión matricial, obtenemos que para un equipo $e_i \in \{e_1, \ldots, e_n\}$:

$$r_i = \frac{1 + \frac{w_i - l_i}{2} + \sum_{j=1, j\neq i}^{n} n_{i,j} * r_j}{2 + n_i} \tag{7}$$

donde para e_k:

- r_k es el rating buscado
- w_k es la cantidad de partidos ganados
- l_k es la cantidad de partidos perdidos
- n_k es la cantidad de partidos jugados
- $n_{k,l}$ es la cantidad de partidos jugados entre e_k y e_l

La primera observación que podemos hacer es que conforme aumente la cantidad de partidos, **el denominador principal aumenta**, con lo cual r_i **se vuelve más pequeño**.

En segundo lugar, hay dos términos en el numerador principal que contribuyen a aumentar o disminuir r_i:

- $\frac{w_i - l_i}{2}$: Alcanza su valor más alto cuando se ganan todos los partidos, contrariamente, es muy desventajoso si se pierden más partidos de los que se ganan dado que se vuelve negativo. En general, del numerador principal, representa la mayor parte del valor de r_i. Observemos también que a medida que este valor cambia, el denominador principal aumenta, por lo que la **ventaja máxima se encuentra ganando partidos**.
- $\sum_{j=1, j\neq i}^{n} n_{i,j} * r_j$: Es una suma en la que cada en cada término se encuentra la cantidad de partidos disputados entre e_i y e_j ponderado por r_j. Cada término de la suma **aporta tanto como** $n_{i,j}$ **cuando** r_j **tiene un valor cercano a** 1 y aporta menos cuando el mismo tiene un valor cercano a 0.

Esto nos permite explicar el carácter “transitivo” de CMM: partidos de equipos con los que e_i haya jugado impactarán en uno de los términos de la última sumatoria vista de forma proporcional a la cantidad de partidos disputados entre e_i y el equipo en cuestión.

3.5. Una estrategia para avanzar en el ranking minimizando la cantidad de partidos

En base a la interpretación anterior, podemos sacar ventaja de CMM si:

- Se compensa la relación entre los términos $(2 + n_i)$ y $\frac{w_i - l_i}{2}$ para evitar que r_i disminuya. Esto se logra manteniendo al máximo la cantidad de partidos ganados.
- De ser posible, jugar menos partidos que los demás equipos.
- Teniendo en cuenta el carácter transitivo mencionado previamente, jugar uniformemente con equipos que tengan bajo rating (dado que la probabilidad de ganarles es mayor) una cantidad considerable de partidos, mientras más, mejor. Eventualmente en la intención de llegar a mejores posiciones de estos equipos, alguno (o ninguno) empezará a subir en la tabla impactando proporcionalmente a la cantidad de partidos que haya jugado con el equipo que aplique la estrategia. También se puede jugar siempre con un mismo equipo de rating bajo, pero la posibilidad de encontrar una ventaja se limita a qué tan bien le vaya a ese único equipo.

3.6. ¿Es CMM un método justo?

Puede pasar que el rating de un equipo se vea afectado por una partida en la que no estuvo involucrado.

Sabemos que de poder elegir contra quien y cuantos partidos jugar, un equipo puede plantear una estrategia para subir su rating sin haber jugado como los otros.

En el caso 9 3.2 presentado en la sección anterior. se observa que si un equipo que está en las últimas posiciones quiere llegar al primer lugar, es difícil que lo haga ganando partidas con equipos que están en posiciones más bajas o bien que no tienen relación alguna con los equipos que están por encima del que quiere subir. Aún después de haber jugado, como en ese caso más de veinte partidos. Por eso también aplicar la estrategia mencionada previamente puede ser arriesgada.

En base a estas cosas, podría calificarse a CMM como un método **no justo, de no tener la competencia reglas claras** que determinen la forma de juego, asegurando que no se puedan incurrir en tales estrategias o asegurando que todos jueguen contra todos. Cabe destacar que en este último caso CMM pierde un poco de gracia y un rankeo por medio de puntos (como en fútbol por ejemplo) sería más efectivo.

3.7. CMM y otros algoritmos de rankeo

En la tabla 10 se puede ver los nombres de los tenistas en el top 10 de CMM, a modo informal de comparación agregamos su posición en WP y Elo así como también su ranking “real”[1]. CMM hace un gran trabajo “descubriendo” cuales son los jugadores top 10, faltando solo uno del ranking oficial (Jo-Wilfried Tsonga) aunque mezcla un poco las posiciones.

Más allá de estos datos curiosos vamos a ver la relación lineal entre CMM y los demás rankings, usando scatterplots, coeficiente de Pearson y *Kendall's weighted* τ. Este último lo descubrimos buscando métodos para comparar rankings, según entendemos es un algoritmo bastante usado y es parte del conjunto de funciones de Scipy.

Según se explica tanto en la documentación oficial de Scipy como en el paper que presenta la idea [3] el algoritmo compara los rankings dando más importancia a los valores de mayor magnitud, pero teniendo en cuenta aún los más chicos. La explicación matemática de cómo se realiza queda fuera de nuestro alcance por el momento aunque nos gustaría verla en detalle de ser posible.

Analizaremos estos algoritmos de rankeo utilizando el conjunto de datos de partidos de ATP 2015[4].

[1] Según `https://www.atptour.com/en/rankings/singles?rankDate=2015-12-28&rankRange=0-100`

id	first_name	last_name	CMM_rank	Elo_rank	WP_rank	Ranking_real
104925	Novak	Djokovic	1	1	27	1
103819	Roger	Federer	2	2	28	3
104918	Andy	Murray	3	3	31	2
104527	Stanislas	Wawrinka	4	6	41	4
105453	Kei	Nishikori	5	5	34	8
104745	Rafael	Nadal	6	4	35	5
104607	Tomas	Berdych	7	9	42	6
103970	David	Ferrer	8	7	33	7
104755	Richard	Gasquet	9	8	43	9
105683	Milos	Raonic	10	12	45	14

Cuadro 10: Posiciones de top 10 de CMM en otros rankings.

3.7.1. Porcentaje de victorias

En la figura 4 podemos ver que la correlación lineal entre CMM y WP es bastante buena salvo por unos grupos de outliers. El coeficiente de Pearson es 0.805, mostrando que esos grupos de outliers no afectan *tanto* y la relación es cercana. Por otro lado el coeficiente de Kendall es 0.763, un poco más bajo lo que indicaría, si mal no entendemos, que la relación no es tan buena para los primeros puestos. Esto concuerda en principio con la tabla 10 presentada al comienzo.

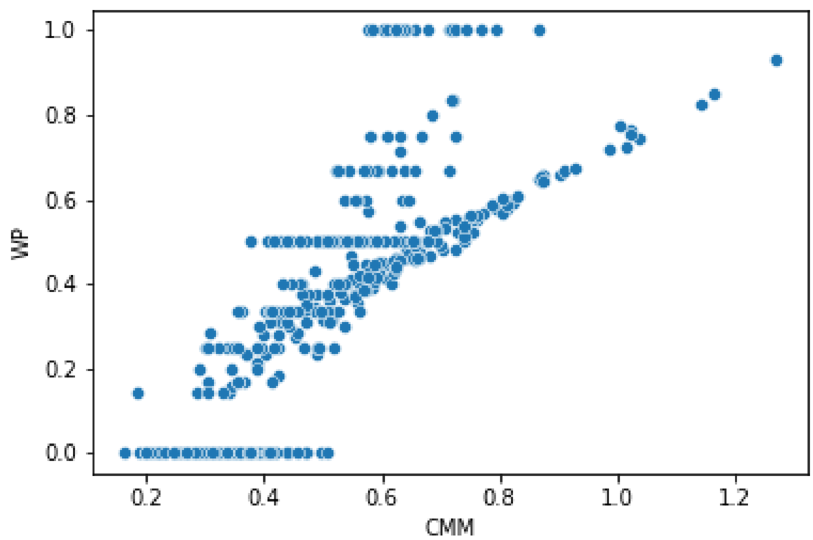

Figura 4: Correlación lineal entre CMM y WP

En cuanto a los tres grupos principales de outliers podemos ver que se tratan de los que tienen WP = 0, 0.5 y 1. Al no tener el ajuste de la regla de Laplace si no se ganó/perdió algún partido los valores van a parar a uno de los extremos; además si se tiene la misma cantidad de ganadas que perdidas el valor de WP será 0.5 ya que no considera “fuerza” del adversario.

Previamente analizamos la composición de r_i para CMM, ahora vamos a analizarla para WP y asi entender las similitudes que se observan.

Para el algoritmo WP:

$$r_i = \frac{w_i}{n_i} \tag{8}$$

donde w_i y n_i tienen el mismo significado que para la definición de CMM 3.4.

Es importante recordar que la definición de CMM tiene su origen en una variación de WP, utilizando la Regla de Laplace de sucesos. Por lo que:

$$r_i = \frac{1 + \frac{w_i - l_i}{2} + \sum_{j=1, j\neq i}^{n} n_{i,j} * r_j}{2 + n_i} = \frac{1 + w_i}{2 + n_i} \quad (9)$$

si se toma la igualdad $w_i = \frac{w_i - l_i}{2} + \frac{n_i}{2}$, asumiendo que todos los equipos inician el sistema con $r_i = \frac{1}{2}$.

Observemos que al final, la expresión de r_i para CMM es muy parecida a la expresión de r_i para WP, por lo que es esperable un comportamiento similar, siempre teniendo en cuenta que el estimador que es objeto de estudio en este trabajo, tiene como objetivo aportar coherencia al asignar ratings y en eso difieren.

3.7.2. Elo

Elo y CMM se parecen menos linealmente, como se puede ver en la figura 5. Su coeficiente de Pearson es 0.789, un poco menos que con WP. Mientras que su coeficiente de Kendall es 0.868, esto nos indica que, si bien no son muy parecidos linealmente tienden a rankear de manera similar los jugadores con mayor cantidad de "puntaje". Esto también concuerda con la tabla 10 de tops.

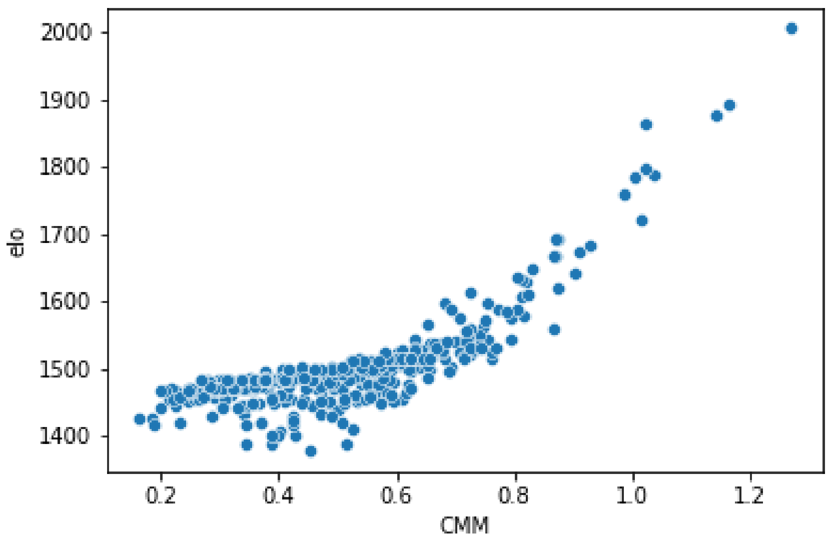

Figura 5: Correlación lineal entre CMM y Elo

De la misma forma que en la sección anterior, pasamos a analizar el cálculo de Elo para entenderlo mejor. Elo se compone de dos partes:

- La probabilidad de que un equipo e_i gane sobre un equipo e_j: $p(elo_i, elo_j) = \frac{1}{1+10^{\frac{elo_j - elo_i}{400}}}$
- Actualización del rating Elo de un equipo e_i que se enfrentó a un equipo e_j: $u(elo_i, elo_j, p_i) = elo_i + k(p_i - p(elo_i, elo_j))$ donde p_i es el puntaje que hizo el equipo e_i en el partido y k es un coeficiente de actualización, fija en tiempo de implementación, que entre otras cosas determina cuantos puntos se ganan o se pierden como máximo.

Algunas diferencias respecto a CMM que se pueden deducir en base a lo mencionado:

- La actualización de elo_i sólo depende del equipo contrincante y de su rating elo (elo_j). Por lo que no hay influencia sobre otros equipos.
- Hasta k puntos puede avanzar o retroceder elo_i dependiendo del resultado del partido y de qué tan probable sea que un equipo gane a otro. Los equipos con resultados más inesperados (a favor o en contra) son los que reciben la mayor diferencia.
- En base a lo anterior, es poco ventajoso que equipos fuertes se enfrenten a equipos débiles dado que en esos partidos la ganancia de puntos en el ranking es mínima.
- Se entiende a Elo como un sistema que permite medir el progreso, dado que recompensa o penaliza las sorpresas.

3.7.3. Otro set de datos

Hicimos las mismas comparaciones utilizando el set de datos de NBA 2016 proporcionado por la cátedra. En principio lo hicimos más por completitud suponiendo que no íbamos a encontrar nada nuevo en cuanto a los algoritmos y sus equivalencias. Además el set de datos al ser bastante más chico, con sólo 30 equipos y 1002 partidos, quizás no proporciona información suficiente para sacar conclusiones fuertes.

Pero al ver la correlación de CMM y WP nos sorprendimos un poco. Con un coeficiente de Pearson de 0.999 y Kendall de 0.989 la correlación lineal es casi perfecta, como se puede ver en la figura 6. Esto implicaría que las "ganancias" de CMM vs WP se pierden.

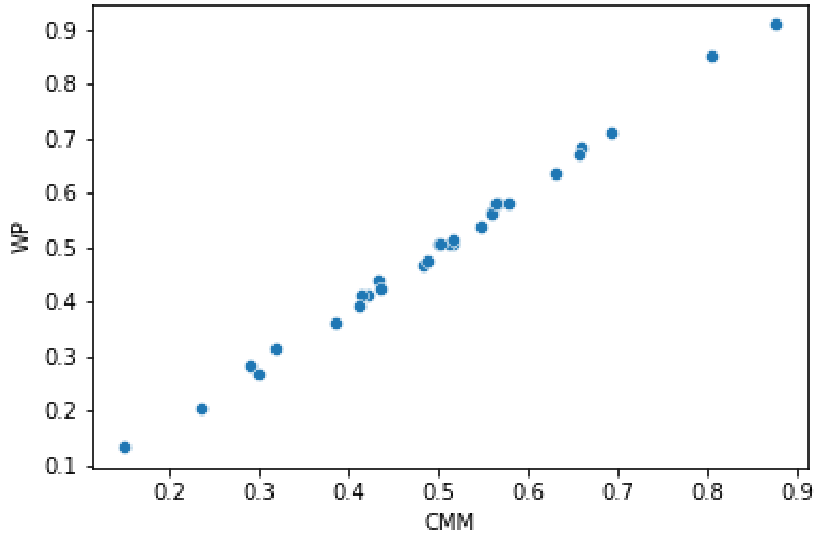

Figura 6: Correlación lineal entre CMM y WP para NBA

Por otro lado los coeficientes de Pearson y Kendall para con Elo dan 0.987 y 0.935 respectivamente, se puede observar en la figura 7 que están bastante ordenados. Como último queríamos comentar la relación entre WP y Elo, 0.985 Pearson y 0.939 Kendall, es decir, los tres son casi equivalentes.

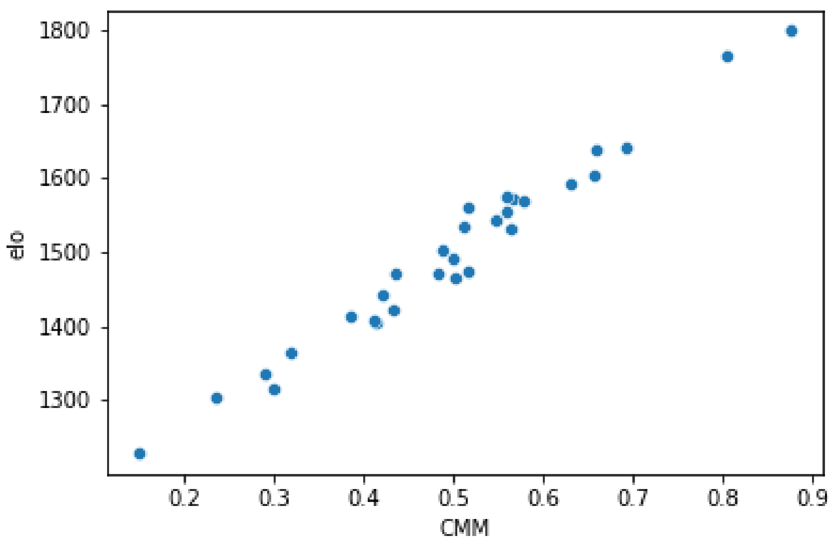

Figura 7: Correlación lineal entre CMM y Elo para NBA

Al analizar el conjunto de datos observamos que todos los equipos juegan mínimo una vez contra los demás y todos juegan una cantidad de partidos parecida (con una diferencia de a lo sumo 3 partidos). Esto transforma este caso en uno super particular, ya que las mejoras de CMM para tener en cuenta contrincantes no son muy necesarias, todos tienen casi el mismo fixture; y la regla de Laplace no aplica tampoco ya que tenemos la misma cantidad de datos para todos los equipos (entre 65 y 68 partidos c/u).

Por ende algo tan simple como WP nos da tan buenos resultados y las “mejoras” de CMM y Elo quedan obsoletas.

4. Conclusiones

Hasta ahora, analizamos CMM desde varias perspectivas.

Desde la perspectiva cuantitativa, comparamos los resultados del método utilizando diferentes matrices C, que representan distintas configuraciones de competencias. Llegamos a que comparando los resultados contra otra implementación más estable numéricamente no obtuvimos grandes errores absolutos. Esto puede deberse principalmente a como es una matriz *C de Colley* y que impacto tiene en la estabilidad numérica resolver el sistema mediante Eliminación Gaussiana sin pivoteo, la cual se detalla en la sección 2.3.

Desde la perspectiva cualitativa, partimos desde el planteo de casos hechos a mano, descubrimos el carácter ”transitivo”del método en la sección 3.3 y planteamos una estrategia para aprovecharlo cuando la competencia se da con ciertas características en 3.5, entre otras cosas. Comparamos a CMM con Elo y WP mediante el análisis de los resultados de ATP 2015 y llegamos a que los tres devuelven un ranking similar al real y lo utilizamos como punto de partida para estudiar la composición de Elo y WP y contrastarlos con CMM: aunque WP es similar a CMM, sin el sesgo que mencionamos, Elo no y presentamos algunas de sus características. Complementamos esta información mirando el set de datos de NBA 2016 de la cátedra.

¿En qué ámbitos más allá de los que todo el mundo tiene acceso en el día a día podemos utilizar CMM? ¿Podemos utilizarlo para definir una nueva política de reemplazo de caché? Aunque el error absoluto en general es pequeño, ¿es suficiente una implementación como la nuestra para pedir errores aún más chicos? ¿Hubieramos obtenido resultados diferentes si las comparaciones las hubiéramos hecho con otro valor de k para el algoritmo de Elo? Son algunas de las preguntas que quedan abiertas en este trabajo.

Referencias

[1] Wesley N. Colley. Colley's Bias Free College Football Ranking Method: The Colley Matrix Explained.

[2] David Goldberg. What Every Computer Scientist Should Know About Floating-Point Arithmetic. 1991.

[3] Grace S. Shieh. A weighted Kendall's tau statistic. 1998.

[4] Jeff Sackmann. Tennis ATP.