Eigenvalores y eigenvectores aplicados para predecir el comportamiento de la migración por medio del método de Markov

Miguel A. Lopez-Olvera[1], Susana Pineda-Solis[1] y Rodrigo Luna-Esquivel[1]

[1] *ENES Juriquilla, Universidad Nacional Autónoma de México, Querétaro, Querétaro, México*

Introducción

Eigenvalores y eigenvectores de una matriz

En álgebra lineal, los vectores propios, autovectores o eigenvectores de un operador lineal son los vectores no nulos que, cuando son transformados por el operador, dan lugar a un múltiplo escalar de sí mismos, con lo que no cambian su dirección. A menudo, una transformación queda completamente determinada por sus vectores propios y valores propios. [1].

Si A es una matriz de nxn hay vectores x diferentes de 0 dentro de los números reales tal que Ax se un multiplo del escalar de x. El escalar denotado por λ se llama valor propio de la matriz A y el vector $x \neq 0$ se conoce como vector propio de A correspondiente a λ, tal que $Ax = \lambda x$ para algun escalar λ, denotado que λ es un eigenvalor de A si y solo si la siguiente ecuación se complue $(A - \lambda I)x = 0$ siendo el polinomio caracteristico de A mientras que su ecuación caracerística esta descrita por $|A - \lambda I_n| = 0$.

Dicho lo anterior, se conoce que existen diferentes métodos que ayudan a conocer los eigenvalores y los eigenvectores de una matriz A de nxn, de los cuales solo se tomarán dos para el desarrollo de este proyecto.

Los métodos utilizados para éste documento son:

- *Método de potencias:* En análisis numérico, el método de las potencias es un método iterativo que calcula sucesivas aproximaciones a los autovectores y autovalores de una matriz. El método se usa principalmente para calcular el mayor y menor eigenvalor con sus correspondientes autovectores de mayor.Para aplicar el método de las potencias se supone que la matriz A de nxn tiene n eigenvalores con un conjunto asociado de eigenvectores linealmente independientes, el método converge lentamente.[2].

Datos de contacto: Miguel A. Lopez-Olvera, Susana Pineda-Solis & Rodrigo Luna-Esquivel, alejandrolo829@gmail.com, susanaps23@gmail.com & rockdrigoluna92@gmail.com

- *Método de QR por ortogonalización de Gram-Schmidt:* El método de ortogonalización de Gram. Schidt es un algoritmo para construir, a partir de un conjunto de vectores de un espacio vectorial con producto interno, otro conjunto ortonormal de vectores que genere el mismo subespacio vectorial.
  El proceso se basa en un resultado de la geometría euclídea, el cual establece que la diferencia entre un vector v y su proyección sobre otro vector u, es perpendicular al vector u. Dicho resultado constituye una herramienta para construir, a partir de un conjunto de dos vectores no paralelos, otro conjunto, conformado por dos vectores perpendiculares.
  La descomposición o factorización QR de una matriz es una descomposición de la misma como producto de una matriz ortogonal por una triangular superior. La descomposición QR es la base del algoritmo QR utilizado para el cálculo de los vectores y valores propios de una matriz.[3, 4]

Procesos de Markov

Un proceso de Markov es un proceso aleatorio indexado por tiempo, y con la propiedad de que el futuro es independiente del pasado, dado el presente. Los procesos de Markov, llamados así por Andrei Markov , se encuentran entre los más importantes de todos los procesos aleatorios. En cierto sentido, son los análogos estocásticos de las ecuaciones diferenciales y las relaciones de recurrencia , que, por supuesto, se encuentran entre los procesos deterministas más importantes.

La complejidad de la teoría de los procesos de Markov depende en gran medida de si el espacio de tiempo es (tiempo discreto) o (tiempo continuo) y si el espacio de estados es discreto (contable, con todos los subconjuntos medibles) o un espacio topológico más general . Cuando T = $[0,\infty)$ o cuando el espacio de estados es un espacio general, generalmente es necesario imponer supuestos de continuidad para descartar varios tipos de comportamiento extraño que de otra manera complicarían la teoría. Los procesos de Markov se utilizan ampliamente en ingeniería, ciencia y modelado empresarial. Se utilizan para modelar sistemas que tienen una memoria limitada de su pasado. En los estudios de crecimiento poblacional, la población de la próxima generación depende principalmente de la población actual y posiblemente de las últimas generaciones. Cuando el espacio de estados es discreto, los procesos de Markov se conocen como cadenas de Markov . La teoría general de las cadenas de Markov es matemáticamente rica y relativamente simple.

Cuando T = N y el espacio de estados es discreto, los procesos de Markov se conocen como cadenas de Markov de tiempo discreto. La teoría de tales procesos es matemáticamente elegante y completa, y es comprensible con una confianza mínima en la teoría de la medida. De hecho, las herramientas principales son la probabilidad básica y el álgebra lineal. Cuando T = [0,∞) y el espacio de estados es discreto, los procesos de Markov se conocen como cadenas de Markov de tiempo continuo.Si evitamos algunas dificultades técnicas (creadas, como siempre, por el espacio de tiempo continuo), la teoría de estos procesos también es razonablemente simple y matemáticamente muy agradable. La propiedad de Markov implica que el proceso, muestreado en momentos aleatorios cuando cambia el estado, forma una cadena de Markov incrustada en tiempo discreto, por lo que podemos aplicar la teoría que ya habremos aprendido.La propiedad de Markov también implica que el tiempo de retención en un estado tiene la propiedad "sin memoriaz, por lo tanto, debe tener una distribución exponencial, una distribución que conocemos bien. Probabilidades de estado límite(estacionaria). La probabilidad de transición de estado de n pasos pij (n) es la probabilidad condicional de que el sistema esté en el estado j después de exactamente n transiciones, dado que actualmente se encuentra en el estado i. Las probabilidades de transición de n pasos se pueden obtener multiplicando la matriz de probabilidad de transición por sí misma n veces.

MIGRACIÓN Y LOS MODELOS DE MARKOV

El método markoviano usualmente supone homogeneidad temporal del fenómeno que intenta modelar, por lo que estima las probabilidades de transición (i.e. de cambio de una situación a otra: o de moverse de un lugar a otro) con técnicas de tipo estático . Sin embargo, la migración es un fenómeno socioespacial altamente dinámico que está influido por una serie de factores cuyo efecto no siempre es bien conocido y resulta muy complicado de anticipar. Quizá por estas razones, las Cadenas de Markov, utilizadas desde los años ochenta en México, han perdido cierta fuerza en años recientes como instrumento de modelación y proyección de flujos migratorios, pero no como instrumentos de proyección de población. Las Cadenas de Markov representan uno de los procesos estocásticos más útiles para modelar fenómenos a partir de una evolución probabilística en el futuro, conociendo solamente la situación presente. Un proceso estocástico registra un comportamiento no-determinista. Es decir, la trayectoria o evolución del proceso depende tanto de variables inherentes o causales del proceso, como de variables aleatorias o estocásticas. En matemáticas,un proceso estocástico es un concepto útil para develar el comportamiento de variables aleatorias que cambian (evolucionan) en función del tiempo(aunque en ocasiones podría tratarse de otra variable: el espacio, por ejemplo). Los flujos migratorios (y muchos otros procesos) observados en el tiempo son a menudo modelados mediante procesos estocásticos, entendidos como cualquier colección de variables aleatorias, dependientes del tiempo $X(t)$, donde el tiempo puede ser discreto $t = 0,1,2,3...$ o continuo, $t \geqslant 0$. En cualquier momento, describe la observación de una variable aleatoria que denotaremos X_t o $X(t)$. Sea X_n un proceso estocástico discreto con espacio de estados contable $E = i, j, k, ...$ Si para todos los enteros $n \geqslant 0$ y todos los estados $i_1,...,i_{n-1}$, entonces:
$P(X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1},...,X_0 = i_0) = P(X_{n+1} = j|X_n = i)$
Donde ambos lados de la ecuación están bien definidos. Entonces este proceso estocástico es llamado una Cadena de Markov o un proceso de Markov, y se le llama cadena homogénea si el lado derecho de la ecuación es independiente de n. La ecuación siguiente son las propiedades de Markov para todos los estados i, j.
$p_{ij} \geqslant 0, \sum p_{ij} = 1.$
El espacio de situaciones puede ser infinito, y por lo tanto la matriz no es en general estudiada mediante álgebra lineal. Sin embargo, satisfacen las operaciones básicas de suma y multiplicación.Así, $(p_{ij})_j$ es un vector de probabilidades para cada i.

OBJETIVOS

Objetivo general

Demostrar que los eigenvalores y los eigenvectores de una matriz A pueden usarse para la predicción del comportamiento de la inmigración con la seleccion previa de los países a trabajar y la afluencia de clientes en empresas, encontrando el eigenvector con el método de potencias y el método QR con ortogonalización de Gram-Schmidt, para aplicar los conocimientos adquiridos en la materia de Computación III impartida en el V Semestre de la Licenciatura en Tecnología.

Objetivos particulares

1. Implementar el método de potencias y el método QR con ortogonalización de Gram-Schmidt, para obtener los eigenvectores que ayudarán a predecir el número de personas inmigrantes que habrá o el número visitas que recibirá una empresa, para así demostrar que el uso de los eigenvalores y eigenvectores nos ayudan a la predicción de estos datos.

2. Discutir los resultados obtenidos de la inmigración así como de la afluencia esperada por empresas, por medio de la obtención de las gráficas correspondientes a cada uno, para ver qué es lo que pasa en cada uno de los casos y así poder sugerir la dependencia de estos procesos.

Planteamiento del problema

Es necesario que el estudiante disponga de bases teóricas durante su formación, pero también es indispensable aplicar sus conocimientos en las diferentes áreas del conocimiento, ese es el poder que tiene la computación, ya que puede ser aplicada prácticamente a cualquier área que el estudiante conozca o tenga interés por comprender.
La experimentación interactiva con fenómenos del mundo real actúa como complemento a la enseñanza de programación, fomentando el interés de los estudiantes, y desarrollando aptitudes en ellos para resolver problemas

físicos que por lo general, son tratados con muchas restricciones y alejados de los problemas reales. La migración es un problema social real que se ha incrementado en los últimos años, siendo ocasionado por las condiciones socioeconómicas, la falta de oportunidades de trabajo críticas que llevan a la extrema pobreza pobreza y que hacen que las personas salgan hacia otros países con la esperanza de buscar una mejor calidad de vida, entre muchas otras razones, en este documento se aplican los conocimientos teóricos aprendidos en la materia de computación III sobre eigenvalores y eigenvectores, así como los diferentes métodos que existen para encontrarlos, todo esto a un problema real, que es la relación que existe de emigración entre distintos países seleccionados.

Hipótesis

El método de potencias y el método QR con ortogonalización de Gram-Schmidt, ayudarán a conocer los eigenvalores y los eigenvectores, que a este último le corresponde la predicción del comportamiento, siendo el número de iteraciones el número de años transcurridos desde el año 0 o inicial.

EXPLICACIÓN DEL CÓDIGO

Preparación de datos

Para poder predecir la evolución de una cadena de Markov, se debe tener un set de datos que relacione a múltiples individuos entre sí. Para el caso de esta implementación, se requiere de un set de datos que relacione la cantidad de emigrantes que cierto país aporta a otro. Estas relaciones van a formar una matriz de adyacencia de la siguiente forma :

Países	EUA	Canadá	España	México	UK
EUA	0	825’040	120’608	11’489’684	716’260
Canadá	270’217	0	11’395	85’825	531’584
España	48’225	7’172	0	53’158	302’020
México	762’290	9’914	23’589	0	3’699
UK	215’915	108’074	98’372	10’457	0

TABLA 1: RELACIONES DE APORTACIÓN DE EMIGRANTES ENTRE PAÍSES EN 2019. DATOS OBTENIDOS DE [11]

Cada elemento en la posición [i][j] de la matriz de la Tabla 1 representa la cantidad de emigrantes que el país de la columna [j] otorga al país de la fila [i]. Por ejemplo, si tomamos en cuenta a la columna de EUA y a la fila de México, se ve que el número de emigrantes que EUA otorga a México es de 762’290. Contrariamente, si se toma en cuenta a la columna de México y a la fila de EUA, se ve que México otorga 11’489’684 de emigrantes a EUA. Esto explica por qué en la diagonal principal de la matriz hay puros ceros, ya que es contradictorio decir que un país puede aportarse a si mismo emigrantes.
El vector inicial, denominado V_0, que se va a tener en cuenta para la cadena de Markov va a tener la forma del que se muestra en la Tabla 2.
Cada elemento de este vector representa la cantidad total de emigrantes que el país correspondiente a esa posición brindó al resto de países en el año 0 de la iteración, es decir, en el 2019. Ahora bien, a pesar de que la matriz de la Tabla 1 puede ser tomada en cuenta para un proceso/cadena de Markov, es mejor crear una matriz de probabilidad a partir de los datos que ya se tienen. Para esto, se debe dividir el valor de cada elemento de la Tabla 1, que se encuentra en una cierta columna [j], entre los emigrantes totales del país correspondiente a la columna [j], datos que se encuentran en la Tabla 2. Es decir, se va a calcular el porcentaje de emigrantes que un país X aporta a otro país Y del total de emigrantes del país X. De este modo nos aseguramos que la matriz cumpla que la suma de probabilidades de las columnas individuales de la matriz dan siempre 1, lo que es un aspecto importante para facilitar los algoritmos de búsqueda de eigenvalores y eigenvectores. La matriz de probabilidad creada, que ahora se va a denominar P, tiene la forma de la Tabla 3.

Países	Emigrantes totales
EUA	1’296’647
Canadá	950’200
España	254’964
México	11’639’124
UK	1’553’563

TABLA 2: APORTACIÓN TOTAL DE EMIGRANTES DE CADA PAÍS EN 2019. DATOS OBTENIDOS DE [11]

Países	EUA	Canadá	España	México	UK
EUA	0	0.8682	0.4749	0.9871	0.4610
Canadá	0.2084	0	0.0449	0.0074	0.3422
España	0.0372	0.0075	0	0.0046	0.1944
México	0.5879	0.0104	0.0929	0	0.0024
UK	0.1665	0.1137	0.3873	0.0009	0

TABLA 3: MATRIZ DE PROBABILIDAD P.

Se puede comprobar que si, por ejemplo, se toma la columna de EUA y se suma todas sus probabilidades, da aproximadamente 1, lo mismo para la columna de México, Canadá, España y UK. Gracias a esta forma, se puede asegurar que el vector final V_F al cual el sistema va a converger, va a ser aquel eigenvector de la matriz de probabilidad P asociado específicamente al eigevalor de 1, y que por lo tanto cumple la ecuación de $P * V_f = V_f$. Es precisamente esta ecuación la que nos dice que, a pesar de aplicar la matriz de probabilidad al vector V_f, este ya no cambia, por lo que se puede decir que el sistema se mantiene en el mismo valor, es decir, se alcanzó la convergencia.
Otra forma de conocer los valores de convergencia es por medio de la fórmula $P^n * V_0 = V_n$. Entre mayor sea el valor de n, mayor va a ser la similitud entre el vector resultante V_n y el vector final V_f.
Entonces, lo que se debe hacer es comprobar que la evolución del vector V_n para n iteraciones converge, efectivamente, en el vector V_f. Además, al graficar los distintos valores que V_n va a ir tomando, se va a observar la evolución del vector inicial V_0 hasta convertirse en V_f. Con relación a los datos migratorios, la gráfica de evolución de V_n va a representar la forma en que cambia la cantidad total de emigrantes que cada país provee a lo largo del tiempo (un año por cada iteración de n). A continuación se explica la implementación.

Implementación

Para la implementación, se utilizaron los datos de la Tabla 3 como matriz de probabilidad P, y los datos de la Tabla 2 como vector inicial V_0. Se usaron estos datos por tener una

relación clara entre los países y para ser fácilmente representados en una misma gráfica, aunque los métodos que a continuación se explicarán funcionarían para cualquier cantidad de países que se quiera tomar en cuenta, siempre y cuando la matriz P y vector V_0 cumplan con sus características ya descritas.
Primero, para encontrar el eigenvector de la matriz P asociado al eigenvalor de 1, se hace lo siguiente:

1. Aplicar el método de potencias a la matriz P para encontrar el eigenvalor mayor de la matriz y su vector asociado. Gracias a las características de la matriz de probabilidad, es seguro asegurar que el eigenvalor mayor de la matriz siempre va a ser 1, y su eigenvector asociado es nuestro vector de interés. El único problema de este método es que necesita un vector de entrada que en realidad puede tener cualquier valor, pero se recomienda que sea un vector de puros 1's para facilitar el cómputo.

2. Aplicar el método de descomposiciones recursivas QR de Gram S. a la matriz P. Recordar que como resultado, este método arroja dos matrices, una matriz R_n, en cuya diagonal principal se encuentran los eigenvalores de P, y una matriz Q_n, en cuyas columnas se encuentran los eigenvectores de P asociados a los eigenvalores de R_n. Ya que sólo nos interesa el eigenvalor de 1 de la matriz P, se debe realizar una búsqueda sobre la diagonal principal de R_n buscando aquel valor que sea más similar a 1. Una vez encontrado, se guarda el índice de la columna donde se encontró al eigenvalor de 1 para que entonces, aquella la columna con ese mismo índice, pero de la matriz Q_n, sea nuestro eigenvector de interés.

3. Ahora, tanto el eigenvector obtenido del método de potencias como el obtenido del método de descomposición QR deben ser normalizados para demostrar que ambos eigenvectores son aproximadamente los mismos. Esto se puede hacer ya que la parte importante de los eigenvectores son las proporciones que los elementos del eigenvector tienen entre sí, no tanto sus valores. Normalizar un eigenvector no afecta en nada a la proporción del mismo, y para realizar la normalización, se debe hacer lo operación $V_norm = V/|V|$, donde V es el vector original, $|V|$ es la norma del vector y V_norm es el vector normalizado.

4. Ahora, tanto el eigenvector obtenido del método de potencias como el obtenido del método de descomposición QR deben ser normalizados para demostrar que ambos eigenvectores son aproximadamente los mismos. Esto se puede hacer ya que la parte importante de los eigenvectores son las proporciones que los elementos del eigenvector tienen entre sí, no tanto sus valores. Normalizar un eigenvector no afecta en nada a la proporción del mismo, y para realizar la normalización, se debe hacer lo operación $V_norm = V/|V|$, donde V es el vector original, $|V|$ es la norma Manhattan del vector y V_norm es el vector normalizado.

5. Teniéndolos normalizados, ambos eigenvectores ya deberían tener valores similares entre sí. Sin embargo, debido a que la matriz de probabilidad está en unidades de probabilidades, los elementos de los eigenvectores también van a representar probabilidades, por lo que es necesario conocer a cuanto equivale en personas emigrantes cada probabilidad presente en los eigenvectores. Para esto, únicamente hay que multiplica los dos eigenvectores normalizados (el de descomposición QR y el de potencias) por el número total de emigrantes que se están considerando en el estudio. En este caso, el número total de emigrantes es la suma del total de emigrantes de todos los países de la Tabla 3.

6. Luego de haber normalizado los eigenvectores resultantes del método de potencias y descomposición QR, y haberlos multiplicado por el número total de emigrantes, los vectores resultantes van a ser aproximadamente iguales entre sí, y ambos van a representar al vector final V_f, que es el vector al cual el sistema va a tender cuantos más años pasen.Notar que el vector V_f no va a depender de ninguna forma del vector inicial V_0, sino únicamente de la matriz de probabilidades P. Esto quiere decir que a pesar de las condiciones iniciales, el sistema siempre va a converger hacia un mismo punto, aunque la forma en que lo haga sea distinta.

Para crear la gráfica que muestre la evolución del vector inicial V_0 con el tiempo, y para comprobar que los valores encontrados del vector final V_f son correctos, se utilizará la ecuación $P^n * V_0 = V_n$. La teoría dice que para un valor de n infinito, el vector V_n será igual a V_f. Para comprobar esto, se realiza lo siguiente:

1. Iniciar un ciclo for desde $n = 0$ hasta un valor máximo arbitrario $n = n_{max}$

2. Calcular el valor de $P^n = P * P * \ldots$. Esto se puede hacer a través del uso de las funciones de multiplicación matricial que vienen en Numpy. Otra forma de calcular P^n es usar la fórmula $P^n = Q_n * R_n^n * Q_n^-$, donde Q_n y R_n son las matrices obtenidas de la descomposición QR de Gram S. en pasos anteriores de la implementación. En teoría, esta última forma de calcular P^n es más sencilla y por ello fue la utilizada en esta implementación, aunque puede llevar a imprecisiones si no se obtienen matrices Q_n y R_n correctas.

3. Con el valor de P^n calcular $P^n * V_0 = V_n$ con uso de instrucciones de Numpy.

4. El vector V_n de cada ciclo será concatenado en forma de columna dentro de una matriz llama $Historic\_Values$

5. Repetir los pasos anteriores hasta terminar el ciclo for.

6. Al final del ciclo for, la matriz $Historic\_Values$, que tiene una dimensión de $(len(V_n), n_{max})$, contendrá a la evolución del vector V_n a lo largo del tiempo. Por lo tanto, cada columna de esta matriz va a ser un valor de V_n en un cierto año n, mientras que cada fila de la matriz va a ser la evolución del número de emigrantes del país correspondiente a lo largo de n años. Entonces, cada fila de la matriz $Historic\_Values$ va a ser graficada contra el número de años transcurridos, de modo que en la gráfica final se tendrán 5 evoluciones distintas, una por cada país que se tomó en cuenta en este proyecto.
Además, el vector V_n de la última columna de la matriz

Historic_Values, que correspondería al año n_{max}, debería tener valores similares al vector final V_f, demostrando que conforme el valor de n aumenta, el vector V_n converge en V_f.

Resultados y Discusión

En las figuras 1 y 2 se observa que el eigenvector de la matriz P (asociado al eigenvalor de 1) resultante del método tanto de las potencias como de la descomposición QR, es aproximadamente el mismo, con diferencias únicamente en los dígitos menos significativos de los elementos. Esto es gracias a que el eigenvector resultante de su respectivo método ya fue normalizado, lo que nos permite visualizar las mismas proporciones de los elementos en ambos vectores.

```
[0.45054265 0.13386831 0.0395731  0.27019533 0.10582062]
```

Fig. 1: Eigenvector de la matriz P asociado al eigenvalor de 1, obtenido con el método de potencias

```
[0.45054266 0.13386831 0.0395731  0.27019532 0.10582061]
```

Fig. 2: Eigenvector de la matriz P asociado al eigenvalor de 1, obtenido con el método de descomposición QR de Gram S.

En las figuras 3 y 4 se observan los eigenvectores correspondientes de cada método, multiplicados por el número total de emigrantes de todos los países considerados en esta implementación. Como ya se explicó, esto se realiza para poder poder decir que estos eigenvectores son iguales al vector final V_f, que en el caso de esta implementación representa al número de emigrantes al que cada país va a converger cuando pase una gran cantidad de años.
Por ejemplo, si consideramos a EUA, se puede decir que en el 2019 (año 0) brindaba 1'296'647 de emigrantes al resto de países (Tabla 2), mientras que luego de haber pasado una cantidad extensa de años, se puede decir que EUA va a aportar 7'070'590 emigrantes aproximadamente (figuras 3 y 4). El motivo de que EUA aumente en su cantidad de emigrantes es por que recibe una gran cantidad de los mismos, provenientes especialmente de México que le brindó a EUA más de 11 millones de personas en el 2019 (Tabla 2). Entonces, el sistema interpreta que al haber mayor número de emigrantes provenientes de otros países en EUA, el número total de emigrantes de este mismo país va a aumentar, y como la probabilidad de que los emigrantes de EUA salgan a otros países no cambia, entonces se va a encontrar en un ciclo donde ingresan más emigrantes de los que salen del país, haciendo que el número de emigrantes totales de EUA suba hasta llegar a un punto de estabilidad/convergencia (ingresan el mismo número de emigrantes al país de los que salen).
Algo contrario sucede con México, país el cual en el 2019 (año 0) brindaba 11'639'124 de emigrantes al resto de países (Tabla 2), mientras que luego de haber pasado una cantidad extensa de años, pasó a aportar 4'240'390 emigrantes aproximadamente (figuras 3 y 4). El motivo de la disminución es que México brinda más emigrantes al resto de países de los que recibe, lo que ocasiona que se encuentre en un sistema donde el número total de emigrantes del país disminuya hasta encontrar un equilibrio/convergencia (ingresan el mismo número de emigrantes a México de los que salen).
Comparando los resultados de las figuras 3 y 4 se nota que los valores de convergencia de cada país son muy similares, con diferencias en algunas décimas o centésimas en algunos casos. Esto es gracias a la ya preexistente similitud entre los eigenvectores de las figuras 1 y 2.

```
EEUA  =  7070590.1583051635
Canada  =  2100862.0505541516
España  =  621040.3792080746
Mexico  =  4240309.799238392
UK  =  1660695.6126942178
```

Fig. 3: Vector final V_f calculado con el método de las potencias

```
EEUA  =  7070590.352786805
Canada  =  2100862.018187144
España  =  621040.3791113383
Mexico  =  4240309.669689429
UK  =  1660695.580225284
```

Fig. 4: Vector final V_f calculado con el método de descomposición QR de Gram S.

Utilizando la ecuación $P^n * V_0 = V_n$ desde $n = 0$ hasta $n = 40$, se obtuvieron 41 valores de V_n que corresponden a la evolución del vector inicial V_0 hasta aproximarse al vector final V_f, o en otras palabras, corresponden a la evolución del número de emigrantes de cada país por 40 años desde el 2019. Considerando al país de EUA, la evolución de su número de emigrantes se muestra en la figura 5. Como se puede ver, el primer valor en la evolución de EUA corresponde al valor inicial declarado en la Tabla 2, mientras que el último valor de la evolución de EUA corresponde a un valor muy similar al que se ve en las figuras 3 y 4 (7 millones aprox.). La evolución del resto de países considerados en esta implementación no es añadida puesto que son arreglos extensos y son más fáciles de interpretar a través de una gráfica (figura 7), pero tienen una forma similar a la evolución de EUA.

```
EEUA  =  [ 1296647.          13151592.00000001  1964359.99973322 11617503.59412332
  3085200.03441791 10577992.10455243  3987239.29520039  9781906.68045431
  4686602.80021727  9166804.03802717  5227422.08778893  8691260.39832789
  5645559.01091618  8323598.323597    5968838.68151451  8039343.34231513
  6218779.80200982  7819573.87961164  6412019.77802229  7649660.88141672
  6561421.71603552  7518294.00945844  6676930.6262608   7416728.76593846
  6766235.41383921  7338204.41244948  6835280.69414345  7277493.93830416
  6888662.49887681  7230556.12210858  6929934.21309149  7194266.52595199
  6961843.10794295  7166209.51572125  6986513.2140923   7144517.46327483
  7005586.70881073  7127746.42689326  7020333.22828542  7114780.03461632
  7031734.38155217]
```

Fig. 5: Evolución de la cantidad de emigrante de EUA a lo largo de 40 años (evolución de V_0)

En la figura 6 se muestra el vector V_{40}, es decir, el vector V_n resultante de la operación $P^{40} * V_0 = V_n$. Este vector representa a la evolución máxima que el número de emigrantes de cada país alcanzó en el transcurso de 40 años. Los valores correspondientes a cada país en el vector V_{40} son similares a los valores de cada país mostrados en las figuras 3 y 4, lo que es un resultado esperable puesto que entre mayor sea la cantidad de años transcurridos, mayor será la similitud entre el vector V_n y el vector final V_f. Para esta cantidad de años, la similitud ya es apreciable, aunque no total, ya que aún hay valores cuyas unidades de las decenas, centenas o incluso millares no coinciden.

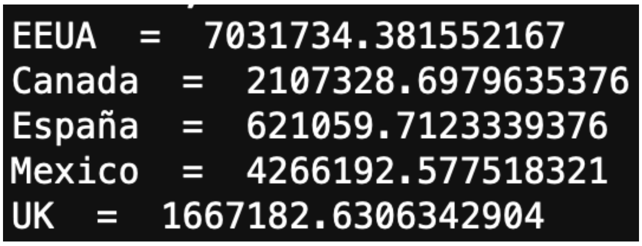

Fig. 6: Vector V_n luego de 40 iteraciones con la ecuación $P^n * V_0 = V_n$

En la figura 7 se encuentra la gráfica de evolución de la cantidad de emigrantes que cada país posee a lo largo de 40 años. En el eje X se dispone el tiempo en unidad de años, mientras que en el eje Y se dispone la cantidad de emigrantes en unidades de 10^7 personas.

Es interesante analizar que todos los países, con excepción de España, oscilan sus número de emigrantes convergiendo poco a poco a un punto que, si se revisa, es aproximadamente el mismo punto que las figuras 3,4 y 6 predicen para cada país. El motivo por el que la cantidad de emigrantes en España no varía mucho puede ser por que no hay una relación muy fuerte de la emigración Española con la emigración del resto de países considerados. También, se nota que las oscilaciones de emigrantes de países como Canadá y UK parecen ir en fase, inclusive hasta sincronizadas. Esto puede deberse a que, de la Tabla 3 se sabe que hay una relación similar entre los emigrantes que UK brinda a Canadá y los emigrantes que Canadá brinda a UK, provocando que un cambio de migrantes en un país repercuta proporcionalmente en el otro.

Algo contrario ocurre con las oscilaciones de México y EUA que parecen ir desfasadas de tal manera que cuando la cantidad de emigrantes de un país alcanza un mínimo, la cantidad de emigrantes del otro país alcanza un máximo y viceversa. Esto se puede deber a que, de nuevo de la Tabla 3, se sabe que México aporta un gran porcentaje de emigrantes a EUA, pero EUA aporta una baja cantidad de emigrantes a México en comparación, haciendo que la relación entre ambos sea inversamente proporcional.

Nota: Se escogió un transcurso de 40 años de manera arbitraria, únicamente para evitar ejecuciones tardadas y resultados irrelevantes para valores de n muy grandes. Con un transcurso de 40 años se observa perfectamente la evolución de cada población migratoria y es claro el punto de convergencia al cual tienden.

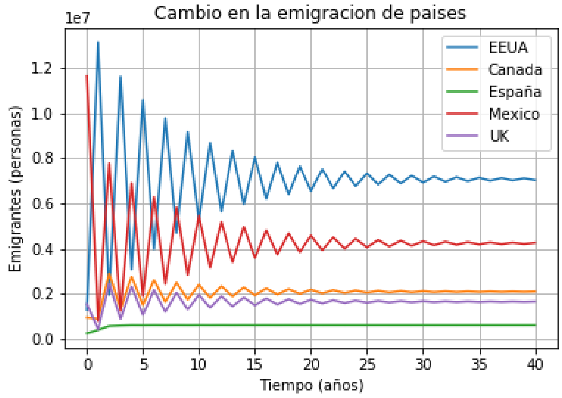

Fig. 7: Gráfica de evolución del vector V_0 a lo largo de 40 años

Conclusiones

Los eigenvalores y los eigenvectores de una matriz A pueden usarse para la predicción del comportamiento de la inmigración y la afluencia de clientes en empresas, esto se sabe obteniendo los eigenvectores con el método de potencias y el método QR con ortogonalización de Gram-Schmidt, siendo que la diferencia entre los resultados de ambos métodos es mínima tomando éstos como correctos, concluyendo que la hipótesis formulada al inicio de este trabajo es verdadera.

Dado que la migración es un fenómeno socioespacial altamente dinámico que está influido por una serie de factores cuyo efecto no siempre es bien conocido y resulta muy complicado de anticipar, las Cadenas de Markov, utilizadas desde los años ochenta en México, no resultan ser una buena forma para la proyección de flujos migratorios y es por eso que han perdido cierta fuerza como instrumento en dicho ámbito. Sin embargo, se puede concluir que el uso de Cadenas de Markov Dinámicas con Medias Móviles puede resultar como una mejor forma para reproducir los flujos migratorios, en lugar de Cadenas de Markov Estáticas como las que se emplearon en este proyecto.

Referencias

[1] Chapra, C. S. (2015). `Numerical Methods For Engineers` (7.a ed.). MCGRAW HILL EDDUCATION.

[2] Colaboradores de Wikipedia. (2019, 22 octubre). Método de las potencias. Wikipedia, la enciclopedia libre. https://es.wikipedia.org/wiki/M%C3%A9todo_de_las_potencias:%7E:text=En%20an%C3%A1lisis%20num%C3%A9rico%2C%20el%20m%C3%A9todo,y%20autovalores%20de%20una%20matriz.text=El%20m%C3%A9todo%20converge%20lentamente%20y,los%20autovectores%20de%20la%20matriz.

[3] Colaboradores de Wikipedia. (2021, 20 enero). Factorización QR. Wikipedia, la enciclopedia libre. https://es.wikipedia.org/wiki/Factorizaci%C3%B3n_QR

[4] Colaboradores de Wikipedia. (2019b, octubre 22). Método de las potencias. Wikipedia, la enciclopedia libre. https://es.wikipedia.org/wiki/M %C3 %A9todo _de_las_potencias: %7E:text=En %20an %C3 %A1lisis %20num %C3 %A9rico %2C %20el %20m %C3 %A9todo,y %20autovalores %20de %20una %20matriz.text=El %20m %C3 %A9to do %20converge %20lentamente %20y,los %20 autovectores %20de %20la %20matriz.

[5] `Introduction to General Markov Processes` https://www.randomservices.org/random/markov/General.html

[6] D.R. Cox, H.D. Miller (1970). “The Theory Stochastic Processes”. Methuen.

[7] A.T. Bharucha-Reid (1960). “Elements Of The Theory of Markov Processes And Their Applications”. McGraw Hill Series in Probability and Statistics.

[8] Gutiérrez C.(2016)PROBLEMAS RESUELTOS DE CADENAS DE MARKOV. https://es.slideshare.net/IsamaraGutirrezCaler /problemas-resueltoscadenasdemarkov-68135356

[9] Garrocho,Jiménez, Álvarez (2016). Modelando la migración interestatal de México: cadenas de Markov estáticas versus Cadenas de Markov dinámicas con medias móviles. https://www.redalyc.org/jatsRepo/112/112498 84005/html/index.html

[10] Ibe, O. C. (2014). Special Random Processes. Fundamentals of Applied Probability and Random Processes, 369–425. doi:10.1016/b978-0-12-800852-2.00012-2

[11] atosmacro. (s. f.). Emigrantes totales 2019. datosmacro.com. Recuperado 29 de enero de 2021, de https://datosmacro.expansion.com/demografia /migracion/emigracion