Aplicaciones de Tensores y Autovalores

María Fernanda Carvajal Guerrero - 2200804*
Angélica María Angarita Leal - 2200801**
Carlos Santiago Rodríguez Sarmiento - 2200799***
Universidad Industrial de Santander
Bucaramanga, Santander

11 de febrero de 2022

Índice

Resumen

Los momentos de una función de variable real continua definida alrededor de un promedio permite establecer relaciones entre datos aleatorios, su media y la relación respecto a otros conjuntos de datos. En el presente trabajo se consideran los primeros tres momentos de una función: El valor total, el promedio pesado y la matriz de covarianza; para ser aplicados sobre un sistema de partículas (Caso 2D y 3D) y una distribución de porcentajes de inversión de GDP en Colombia en defensa, salud, educación y ciencia en el período (2004-2018). Para la primera aplicación, tanto para el caso 2D y 3D, se halla la variable total del sistema, su centro de masa y tensor de inercia, mientras que para la segunda aplicación, se encuentran las matrices de

* e-mail: maria2200804@correo.uis.edu.co
** e-mail: angelica2200801@correo.uis.edu.co
*** e-mail: carlos2200799@correo.uis.edu.co

covarianza (segundo momento de la función) y de correlación y se encuentra alta relación entre la inversión en ciencia y salud. Finalmente, para ambos casos, se encuentran sus autovalores y autovectores que describen los ejes principales de inercia, para el primer caso, y los direcciones de mayor dispersión para el segundo caso.

1. Introducción

Para una función $f(x)$ cuya variable x es real continua y definida alrededor de un valor promedio de la variable, es posible generar sus momentos como:

$$\mu_n = \int_{-\infty}^{\infty} dx f(x)(x-\bar{x})^n \Rightarrow \mu_n = \sum_{i=1}^{N} F\left(|x_i\rangle\right)\left(|x\rangle_i - |\bar{x}\rangle\right)^n \Rightarrow \sum_{i=1}^{N} F\left(|x_i\rangle\right)\left(|x\rangle_i - |\bar{x}\rangle\right)\otimes\underbrace{\otimes\cdots}_{n-1}\left(|x\rangle_i - |\bar{x}\rangle\right), \tag{1}$$

donde, aunque en la segunda relación se particularice el caso discreto, el vector $|x\rangle_i$ puede tener m componentes. Así entonces, los momentos μ_n de una variable, $v_i = \mathcal{F}(|x_i\rangle)$, representan los promedios pesados por potencias de las desviaciones de las componentes aleatorias de un vector $|x\rangle_i$ respecto a su media $|\bar{x}\rangle = \frac{1}{N}\sum_{i=1}^{N}|x\rangle_i$. Siendo los primeros tres momentos:

- **Momento de orden cero**: $\mu_0(v) = \sum_{i=1}^{N} v_i$, representa el 'valor total' de la variable del sistema.
- **Momento de orden uno**: $\mu_1(v) = \sum_{i=1}^{N} v_i\left(|x\rangle_i - |\bar{x}\rangle\right)$, corresponde al promedio pesado de la variable.
- **Momento de orden dos**: $\mu_2(v) = \sum_{i=1}^{N} v_i\left(|x\rangle_i - |\bar{x}\rangle\right)^2$, corresponde a la matriz de covarianza de la variable que generaliza la noción de varianza a dimensiones múltiples.

Es importante resaltar que los vectores $|x\rangle_i$ y $|\bar{x}\rangle$ están expresados en bases no ortogonales con una norma asociada a una matriz densa:

$$\sum_{i=1}^{N} v_i\left(|x\rangle_i - |\bar{x}\rangle\right)^2 = \sum_{i=1}^{N} v_i\left(x_1-\bar{x}_1, \cdots x_k-\bar{x}_k, \cdots x_m-\bar{x}_m\right)_i \begin{pmatrix} x^1-\bar{x}^1 \\ \vdots \\ x^I-\bar{x}^l \\ \vdots \\ x^m-\bar{x}^m \end{pmatrix}_i \tag{2}$$

Ahora bien, a partir de lo anterior es posible describir dos situaciones en particular. En primer lugar un sistema de n partículas con masas distintas, donde cada momento representa su masa total, centro de masa y tensor de inercia respectivamente. Por otro lado, es posible describir las relaciones que hay entre los porcentajes de inversión de GDP que se ha destinado en Colombia para defensa, salud, educación y ciencia para el periodo 2004 - 2018. De esta manera el trabajo se desarrollará en dos partes principales como se presenta a continuación:

- Sistema de partículas (Caso 2D y caso 3D)
  1. Se presentarán los tres primeros momentos del sistema
  2. Se discute la posibilidad de que los vectores cartesianos sean autovectores del sistema
  3. Se encontrarán los ejes principales del sistema
  4. Se encontrará la matriz de transformación entre la base cartesiana y de autovectores
- Inversión del GDP de Colombia (2004 - 2018)
  1. Se encontrará la matriz de covarianza y correlación del sistema
  2. Se hallarán los autovalores y autovectores de la matriz de covarianza y se discute su significado
  3. Se encontrará la matriz de transformación entre la base cartesiana y de autovectores

En la metodología se presentará el procedimiento realizado en un entorno de *Python*, haciendo uso de librerías como *Pandas*[1] y *SciPy*[2], que realizan manipulación de datos y poseen sublibrerías de álgebra lineal. En los resultados se presentará la solución a los ítems enumerados anteriormente y se discute su relación con la aplicación escogida. Finalmente, se presentan los hechos más relevantes de los resultados y su implicación en cada aplicación y se hacen recomendaciones para futuros trabajos con un enfoque más geométrico.

2. Metodología

2.1. Sistema de Partículas

Primeramente, los datos correspondientes a las masas y posición del sistema de partículas inmersas en un volumen fueron extraídos del archivo de datos. Se dividieron respectivamente para el caso 2D y 3D. Cuando fueron cargados los datos a un entorno de programación de *Python*, se calcularon los tres primeros momentos del sistema teniendo en cuenta que $\mathcal{F}(|x_i\rangle) = m_i$, de la siguiente manera:

- **Momento Orden 0** Fue la masa total de todo el sistema.

$$\mu_0 = \sum_{i=1}^{N} m_i. \tag{3}$$

- **Momento Orden 1** Hace referencia al centro de masa del sistema de partículas.

$$\mu_1 = \sum_{i=1}^{N} m_i(|x_i\rangle). \tag{4}$$

[1]https://pandas.pydata.org/
[2]https://scipy.org/

- **Momento Orden 2** A partir de operar la siguiente ecuación se obtuvo el tensor inercia [5].Para el caso 2D se obtuvo una matriz $2x2$, cuyas componentes de la diagonal son los momentos de inercia de x y y.

$$\mu_2 = \sum_{i=1}^{N} m_i(|x_i\rangle)^2.$$

$$\mu_2 = \begin{bmatrix} I_{xx} & -I_{xy} & -I_{xz} \\ -I_{yx} & I_{yy} & -I_{yz} \\ -I_{zx} & -I_{zy} & I_{zz} \end{bmatrix} \tag{5}$$

En general, el cálculo de las ecuaciones de los momentos se obtuvo mediante la librería *Numpy*[3], que permite emplear funciones matemáticas (como la sumatoria) para arreglos multidimensionales como el trabajado en el problema.

Así mismo, usando *Numpy*, se insertaron los vectores cartesianos para cada dimensión y se realizó un producto entre el tensor de inercia correspondiente y los vectores cartesianos como se observa en [6], esto para saber si dichos vectores eran autovectores del tensor momento de inercia.

$$Av = \lambda v \tag{6}$$

Posteriormente, para encontrar los ejes principales de inercia para la distribución de masas, que son los vectores propios del tensor de inercia calculado, se ejecutó la función *scipy.linalg.eig* que brindó los autovalores y autovectores (Ver la figura1), seguidamente se comprobó si eran ortogonales mediante el producto punto, esto es que el producto sea igual a 0 para saber si eran linealmente independientes y en consecuencia formaban una base ortogonal.

```
#Autovalores y autovectores del tensor
evals, evecs = la.eig(I2)
```

Figura 1: Función *Scipy.linalg.eig* que calcula los eigenvalores y eigenvectores de una matriz, en este caso del tensor de inercia 2D

Finalmente, para encontrar la matriz de transformación de la base cartesiana a la base de autovectores se resolvieron sistemas de ecuaciones de igual dimensión al tensor de cada caso, partiendo de:

$$\mu_n \mathbb{B} = {\mu_n}^{\star} \qquad n = 2, 3.$$

Siendo μ_n el tensor para cada caso, $\mathbb{B}$ la matriz de transformación y ${\mu_n}^{\star}$ la representación matricial en la base de autovectores. Así se obtuvo el sistema ecuaciones para cada caso de la forma:

$$Ax = b, \tag{7}$$

donde A son las componentes del tensor en su base original, x el i-ésimo vector columna de la matriz de transformación a buscar y b es la i-ésima columna de la representación matricial de la base de

[3] https://numpy.org/

autovectores. De esta forma se obtuvieron las n tuplas según la dimensión del caso que son las columnas de la matriz de transformación encontrada.

2.2. Covarianza y correlación de la inversión de GDP en Colombia

En primer lugar, los porcentajes de inversión en eduación, salud, ciencia y defensa correspondientes a los años 2004 - 2018 fueron extraídos de la base de datos del Banco Mundial `https://data.worldbank.org/`. Los porcentajes obtenidos fueron reportados sobre el total de PIB de Colombia en cada año. Una vez cargado el conjunto de datos a un entorno de *Python*, se definió el vector cuyas componentes representan la inversión en defensa, salud, educación y ciencia, en ese orden. Así mismo, tomando en cuenta la covarianza de un vector de m componentes es

$$\frac{1}{N}\sum_{i=1}^{N}\left(x_i^j-\bar{x}^j\right)(x_{ki}-\bar{x}_k)\,, \tag{8}$$

se creó la función para calcular la covarianza entre dos componentes arbitrarias de un vector como se muestra en la Fig.2.

```
def cov(x, y):
        xbar, ybar = x.mean(), y.mean()
        return np.sum((x - xbar)*(y - ybar))/(len(x))
```

Figura 2: Función de covarianza para dos componentes arbitrarias. La función *cov*, a pesar de recibir dos argumentos en principio distintos, lleva a la varianza de la componente j-ésima del vector cuando sus entradas son iguales, definida matemáticamente como $\frac{1}{N}\sum_{i=1}^{N}\left(x_i^j-\bar{x}^j\right)^2$.

A continuación, se calculó la covarianza entre las componentes j-ésima y k-ésima del vector, con j,k=1,2,3,4 y se generó la matriz que contiene estos valores; tomando en cuenta que su diagonal representa la varianza pues la entradas a la función *cov* son iguales para j=k.

Por otra parte, haciendo uso de la librería *Pandas*[4], especializada en la manipulación y análisis de datos, se halló la matriz de correlación entre componentes las del vector de inversión, utilizando la función *pandas.DataFrame.corr()*, cuyas entradas se dejaron vacías para que fuera calculada la correlación estándar de las variables.

Retomando la matriz de covarianza, se encontraron sus autovalores y autovectores utilizando la sublibrería de álgebra lineal de *SciPy*, *scipy.linalg*. Este paquete incluye la función *scipy.linalg.eig(a)*, donde a corresponde a la matriz a ingresar y retorna un primer arreglo unidimensional donde se encuentran los autovalores de a y una matriz con los autovectores asociados, en su orden a tales autovalores. A manera de corroboración del método utilizado en la función, se multiplicó la matriz de covarianza por la matriz de eigenvectores por derecha y por su inversa por izquierda. Al realizar

[4] https://pandas.pydata.org/

esto, se obtuvo la matriz diagonal con entradas correspondientes a los autovalores obtenidos en principio y se concluye con esta sección.

Finalmente, se encontró la matriz de transformación que lleva a la matriz de covarianza en su base original a su representación en la base de autovalores. Para ello, se resolvieron cuatro sistemas de ecuaciones que parten, de lo siguiente:

$$\mathbb{C}\mathbb{A} = \mathbb{C}^{\star}, \tag{9}$$

donde $\mathbb{C}$ representa la matriz de covarianza en su base original, $\mathbb{A}$ es la matriz de transformación a buscar y $\mathbb{C}^{\star}$ es la matriz de covarianza en la base de autovectores. Así, los sistemas de ecuaciones vienen dados por |7| donde A son las componentes de la matriz de covarianza en su base original, x es el j-ésimo vector columna de la matriz de transformación a buscar y b es la j-ésima columna de la matriz de covarianza en la base de autovectores.

Finalmente, las cuatro cuádruplas obtenidas son las columnas de la matriz de transformación buscada y se comprueba mediante el producto presentado en 9, pues la multiplicación de la matriz de covarianza con la matriz de transformación formada dio como resultado la representación matricial de la covarianza en la base de autovectores, ya obtenida anteriormente.

3. El experimento y los resultados

3.1. Sistema de Partículas

3.1.1. Sistema 2D

Este sistema va a estar ubicado en el plano xy por tanto, los datos tomados sólo estuvieron en ese plano.

El momento de orden 0 fue obtenido partir de la ecuación [3], donde se obtuvo $\mu_0 = 4627$ siendo este resultado la masa total del sistema.

Para el momento de orden 1 se obtiene un resultado que es necesario analizarlo en varias partes. Primero, en base de la ecuación [4] se obtiene el centro de masa.

$$\mu_1 = [3821047, 3594802]. \tag{10}$$

Segundo, el resultado anterior [10] cumpliendo con la definición de centro de masa es la masa total del sistema multiplicado por el vector posición del centro de masa. Por esta razón se puede utilizar este resultado para corroborar el resultado del momento 0.

El momento de orden 2, tiene más factores que se pueden analizar ya que corresponde al tensor de inercia del sistema [5].

$$\mu_2 = \begin{bmatrix} 3{,}75652840e+09 & -3{,}88039010e+09 \\ -3{,}88039010e+09 & 4{,}11401434e+09 \end{bmatrix}. \tag{11}$$

Como se mencionó anteriormente el tensor de inercia, en este caso eq.13 se pueden analizar a partir de este varios componentes como lo son sus vectores propios tal y como se muestra a continuación.

Primero para saber si los vectores cartesianos $\hat{\imath}$ y $\hat{\jmath}$ son base propia de este sistema, se debió comprobar si estos vectores son autovectores de sistema. Así que usando la eq.6, se obtienen los siguientes resultados:

- $\hat{\imath} = [1, 0]$
$$\mu_2\hat{\imath} = [3{,}7565284e + 09, -3{,}8803901e + 09] \neq \lambda\hat{\imath}.$$
- $\hat{\jmath} = [0, 1]$
$$\mu_2\hat{\jmath} = [-3{,}88039010e + 09, 4{,}11401434e + 09] \neq \lambda\hat{\jmath}.$$

Con base en los resultados anteriores se tiene que ninguno de los dos vectores resultantes se puede escribir como su respectivo vector por un escalar, en este caso que sería también un autovalor del tensor [13]. Así que de esta manera se obtuvo que los vectores cartesianos no corresponden a eigenvectores del tensor de inercia y por tanto tampoco son una base propia para esta distribución de masa.

Entonces, se procedió a hallar los vectores propios del sistema obteniendo como resultado:

$$\lambda_1 = 5{,}07667234e{+}07 \rightarrow |v_1\rangle = \begin{pmatrix} -0{,}72319235 \\ -0{,}69064667 \end{pmatrix} \qquad \lambda_2 = 7{,}81977602e{+}09 \rightarrow |v_2\rangle = \begin{pmatrix} 0{,}69064667 \\ -0{,}72319235 \end{pmatrix}.$$

Después se realiza la prueba para saber si estos autovectores pueden ser base del sistema. Esto se puede comprobar realizando el producto escalar entre esos vectores. De esta manera, si se obtiene que son ortogonales significan que son linealmente independientes y al ser vectores propios se demuestra que son base.

$$|v_1\rangle \cdot |v_2\rangle = -4{,}574599442682926e - 18 \approx 0.$$

A partir del resultado anterior se comprobó que los eigenvectores $|v_1\rangle$ y $|v_2\rangle$ son base ortogonal del sistema. Cabe aclarar que se realiza la aproximación a 0 ya que el $e - 18$ es un factor muy pequeño por tanto es despreciable.

Además, al obtener los vectores propios del tensor y comprobar que son base. Se obtuvo que $|v_1\rangle$ y $|v_2\rangle$ son los ejes principales de inercia para el sistema.

A partir de su base de autovectores se construye la matriz con estos como columnas, siendo esta la matriz constituida por los ejes principales de la distribución.

$${\mu_2}^{\star} = \begin{pmatrix} -0{,}72319235 & 0{,}69064667 \\ -0{,}69064667 & -0{,}72319235 \end{pmatrix},$$

ahora se halla la matriz de transformación de la base cartesiana a la base de eigenvectores, colocando cada tupla como una columna de esta matriz:

$$\mathbb{B} = \begin{pmatrix} -1{,}42454014e - 08 & 8{,}83205182e - 11 \\ -1{,}36043184e - 08 & -9{,}24824892e - 11 \end{pmatrix}.$$

Finalmente, se puede comprobar los resultados hallados anteriormente realizando el producto entre el tensor de inercia [13] y la matriz B, este resultado será la matriz de vectores propios:

$$\mu_2\mathcal{B} = \begin{pmatrix} -0{,}72319235 & 0{,}69064667 \\ -0{,}69064667 & -0{,}72319235 \end{pmatrix} = \mu_2{}^{\star}.$$

3.1.2. Sistema 3D

Para este caso, el momento de orden 0 también fue obtenido partir de la ecuación [3], donde se obtuvo $\mu_0 = 4627$ siendo este resultado la masa total del sistema. Además se puede observar que entre las dos dimensiones no hay cambio de masa, ya que se tienen los mismos valores en esta categoría.

Para el momento de orden 1 se obtiene un resultado que es necesario analizarlo en varias partes. Primero, en base de la ecuación [4] se obtiene el centro de masa, con el cambio que el vector resultante es en 3 dimensiones.

$$\mu_1 = [3821047, 3594802, 71734]. \tag{12}$$

Segundo, el resultado anterior [12] cumpliendo con la definición de centro de masa se puede utilizar este resultado para corroborar el resultado del momento 0, ya que es la masa total del sistema multiplicado por el vector posición del centro de masa.

El momento de orden 2, tiene más factores que se pueden analizar ya que corresponde al tensor de inercia del sistema [5], esta vez en 3 dimensiones.

$$\mu_2 = \begin{bmatrix} 3{,}85948384e+09 & -3{,}88039010e+09 & -5{,}20969800e+07 \\ -3{,}88039010e+09 & 4{,}21696978e+09 & -5{,}38018760e+07 \\ -5{,}20969800e+07 & -5{,}38018760e+07 & 7{,}87054274e+09 \end{bmatrix}. \tag{13}$$

Como se mencionó anteriormente el tensor de inercia, en este caso eq.13 se pueden analizar a partir de este varios componentes como lo son sus vectores propios tal y como se muestra a continuación.

Primero para saber si los vectores cartesianos $\hat{\imath}$, $\hat{\jmath}$ y $\hat{k}$ son base propia de este sistema, se debió comprobar si estos vectores son eigenvectores de sistema. Así que usando la eq.6, se obtienen los siguientes resultados:

- $\hat{\imath} = [1, 0, 0]$

$$\mu_2\hat{\imath} = [3{,}85948384e+09, -3{,}88039010e+09, -5{,}20969800e+07] \neq \lambda\hat{\imath}.$$

- $\hat{\jmath} = [0, 1, 0]$

$$\mu_2\hat{\jmath} = [-3{,}88039010e+09, 4{,}21696978e+09, -5{,}38018760e+07] \neq \lambda\hat{\jmath}.$$

- $\hat{k} = [0, 0, 1]$

$$\mu_2\hat{k} = [-5{,}20969800e+07, -5{,}38018760e+07, 7{,}87054274e+09] \neq \lambda\hat{k}.$$

Con base en los resultados anteriores se tiene que ninguno de los dos vectores resultantes se puede escribir como su respectivo vector por un escalar, en este caso que sería también un autovalor del tensor [13]. Así que de esta manera se obtuvo que los vectores cartesianos no son una base propia para esta distribución de masa ya que no corresponden a vectores propios del tensor de inercia de 3 dimensiones.

Entonces, se procedió a hallar los vectores propios del sistema y así ver si esos autovectores pueden ser una base ortogonal para la distribución.Se obtuvo como resultado:

$$\lambda_1 = 1{,}52996521e{+}08 \rightarrow |v_1\rangle = \begin{pmatrix} -0{,}72315583 \\ -0{,}69061685 \\ -0{,}00969618 \end{pmatrix} \qquad \lambda_2 = 7{,}92289755e{+}09 \rightarrow |v_2\rangle = \begin{pmatrix} -0{,}68914397 \\ 0{,}72240857 \\ -0{,}05662549 \end{pmatrix},$$

$$\lambda_3 = 7{,}87110229e + 09 \rightarrow |v_3\rangle = \begin{pmatrix} -0{,}04611112 \\ 0{,}03426699 \\ 0{,}9983484 \end{pmatrix}.$$

Para comprobar si estos autovectores pueden ser base del sistema se realizó el producto escalar entre esos vectores. De esta manera, si se obtiene que son ortogonales significan que son linealmente independientes y al ser vectores propios se demuestra que son base.

$$|v_1\rangle \cdot |v_2\rangle = -2{,}4383706859198995e - 16 \approx 0.$$

$$|v_1\rangle \cdot |v_3\rangle = -1{,}734723475976807e - 17 \approx 0.$$

$$|v_2\rangle \cdot |v_3\rangle = 1{,}1379786002407855e - 15 \approx 0.$$

A partir del resultado anterior se comprobó que los eigenvectores $|v_1\rangle$, $|v_2\rangle$ y $|v_3\rangle$ son base ortogonal del sistema. Cabe aclarar que se realiza la aproximación a 0 ya que el $e - 16$, $e - 17$ y $e - 15$ son factores muy pequeños por tanto son despreciables.

Además, al obtener los vectores propios del tensor y comprobar que son base. Se obtuvo que $|v_1\rangle$, $|v_2\rangle$ y $|v_3\rangle$ son los ejes principales de inercia para el sistema.

Con base en los autovectores que ya fue comprobado que son base se construye la matriz con estos como columnas, siendo esta la matriz constituida por los ejes principales de la distribución.

$$\mu_2{}^{\star} = \begin{pmatrix} -0{,}72319235 & -0{,}68914397 & -0{,}04611112 \\ -0{,}69061685 & 0{,}72240857 & 0{,}03426699 \\ -0{,}00969618 & -0{,}05662549 & 0{,}9983484 \end{pmatrix},$$

ahora se halla la matriz de transformación de la base cartesiana a la base de vectores propios, colocando cada tupla como una columna de esta matriz:

$$\mathbb{B} = \begin{pmatrix} -4{,}72661617e - 09 & -8{,}69813050e - 11 & -5{,}85828018e - 12 \\ -4{,}51393824e - 09 & 9{,}11798451e - 11 & 4{,}35351871e - 12 \\ -6{,}33751660e - 11 & -7{,}14706861e - 12 & 1{,}26837178e - 10 \end{pmatrix}.$$

Finalmente, se puede comprobar los resultados hallados anteriormente realizando el producto entre el tensor de inercia [13] y la matriz B, este resultado será la matriz de vectores propios:

$$\mu_2\mathbb{B} = \begin{pmatrix} -0{,}72319235 & -0{,}68914397 & -0{,}04611112 \\ -0{,}69061685 & 0{,}72240857 & 0{,}03426699 \\ -0{,}00969618 & -0{,}05662549 & 0{,}9983484 \end{pmatrix} = {\mu_2}^{\star}.$$

3.2. Covarianza y correlación de la inversión de GDP en Colombia

La matriz de covarianza para el vector de inversiones de GDP de Colombia, cuyas componentes en su orden son inversión en defensa, salud, educación y ciencia, se presenta a continuación en la Fig.3.

Figura 3: Matriz de covarianza para el vector de inversiones de GDP de Colombia. La diagonal de la matriz representa la varianza de cada componente, es decir cuán alejado se encuentra una variable de su media. En este caso las componentes (i,i) con i=0,1,2,3 representan la varianza de la inversión en defensa, salud, educación y ciencia respectivamente.

La matriz obtenida es una matriz simétrica, debido a que la covarianza existente entre las componentes no cambia con el orden en que estos se operen. Esto porque la covarianza representa la relación lineal entre la variación de dos componentes respecto a su media aritmética. Entonces, de acuerdo con la matriz de covarianza obtenida, se observa que la componente que menos varió respecto a su promedio entre los años 2004 - 2018 es la inversión en ciencia, contrario a la inversión en salud. Ahora bien, al revisar las covarianzas entre las componentes, se encuentra que existe una dependencia inversa entre la inversión en defensa y las demás inversiones. Es decir, a medida que los valores de inversión en defensa aumentan respecto a su promedio, los valores de las demás componentes se alejan por debajo de su promedio. Análogamente para la inversión en salud, educación y ciencia: Al tener covarianzas positivas, a medida que una componente crece respecto a su promedio las demás lo hacen también. Cabe resaltar, que las variables con la covarianza más cercana a 0 son la inversión en defensa y en eduación; esto implica que tienen la menor relación lineal respecto a las relaciones entre las demás inversiones.

Por otra parte se encontró la matriz de correlación para el vector de inversiones de GDP en Colombia tal y como se presenta en la Fig.4.

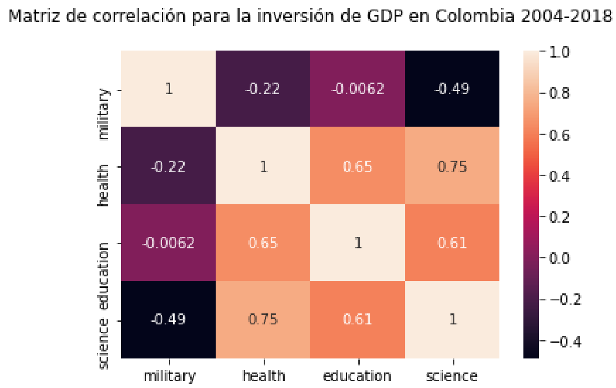

Figura 4: Matriz de correlación para el vector de inversiones de GDP de Colombia. Las componentes de su diagonal son todas 1 porque miden la relación lineal de una componente consigo misma. La matriz de correlaciones solo presenta valores entre -1 a 1, considerándose una alta correlación los valores sobre 0.7.

Tal y como sucedió con la matriz de covarianza, la matriz de correlación es una matriz simétrica, pues la correlación entre dos variables no depende del orden que se operen. Cabe resaltar que la correlación es el cociente entre la covarianza de las variables y el producto de sus desviaciones estándar, razón que explica la simetría de ambas matrices. Ahora bien, al respecto de los valores obtenidos en la matriz, se observa que las inversiones más altamente relacionadas son salud y ciencia, teniendo un coeficiente de correlación de 0.75. Opuesto a la situación entre inversión en educación y defensa que tienen una relación inversa muy cercana a 0; es decir, están muy poco relacionadas aunque a medida que aumente una disminuya la otra. Finalmente, se puede decir que las variables mejor correlacionadas son inversión en ciencia, salud y educación; mientras que la inversión en defensa siempre tiene una relación inversa con las demás inversiones, siendo la más marcada aquella existente entre esta e inversión en ciencia.

Ahora retomando la matriz de covarianza, se obtienen sus autovalores con sus autovectores asociados. Siendo estos:

$$\lambda_1 = 0{,}36 \rightarrow |u_1\rangle = \begin{pmatrix} 0{,}094 \\ -0{,}91 \\ -0{,}39 \\ -0{,}063 \end{pmatrix} \qquad \lambda_2 = 0{,}067 \rightarrow |u_2\rangle = \begin{pmatrix} -0{,}86 \\ 0{,}12 \\ -0{,}5 \\ 0{,}0463 \end{pmatrix}$$

$$\lambda_3 = 0{,}04 \rightarrow |u_3\rangle = \begin{pmatrix} -0{,}5 \\ -0{,}38 \\ 0{,}77 \\ 0{,}057 \end{pmatrix} \qquad \lambda_4 = 0 \rightarrow |u_4\rangle = \begin{pmatrix} 0{,}074 \\ -0{,}041 \\ -0{,}046 \\ 1 \end{pmatrix}.$$

Así, se construye la matriz de covarianza en la base de autovectores:

$$\mathbb{C}^\star = \begin{pmatrix} 0{,}094 & -0{,}86 & -0{,}5 & 0{,}074 \\ -0{,}91 & 0{,}12 & -0{,}38 & -0{,}041 \\ -0{,}39 & -0{,}5 & 0{,}77 & -0{,}046 \\ -0{,}063 & 0{,}046 & 0{,}057 & 1 \end{pmatrix} \tag{14}$$

y se realiza la diagonalización de esta matriz para verificar los autovectores obtenidos multiplicando a derecha e izquierda por la matriz $\mathbb{C}^\star$ y su inversa, respectivamente. Se obtiene que:

$$(\mathbb{C}^\star)^{-1}\mathbb{C}\mathbb{C}^\star = \begin{pmatrix} 0{,}36 & 0 & 0 & 0 \\ 0 & 0{,}067 & 0 & 0 \\ 0 & 0 & 0{,}04 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

donde las componentes en la diagonal de la matriz son los autovalores de la matriz de covarianza. Es importante resaltar que la matriz $\mathbb{C}^\star$ es una matriz unitaria, pues cumple que:

$$(\mathbb{C}^\star)^{-1} = \begin{pmatrix} 0{,}094 & -0{,}91 & -0{,}39 & -0{,}063 \\ -0{,}86 & 0{,}12 & -0{,}5 & 0{,}046 \\ -0{,}5 & -0{,}38 & 0{,}77 & 0{,}057 \\ 0{,}074 & -0{,}041 & -0{,}046 & 1 \end{pmatrix} = \begin{pmatrix} 0{,}094 & -0{,}86 & -0{,}5 & 0{,}074 \\ -0{,}91 & 0{,}12 & -0{,}38 & -0{,}041 \\ -0{,}39 & -0{,}5 & 0{,}77 & -0{,}046 \\ -0{,}063 & 0{,}046 & 0{,}057 & 1 \end{pmatrix}^T = (\mathbb{C}^\star)^T.$$

Ahora bien, el autovector más grande de la matriz de covarianza, en este caso $|u_1\rangle$, representa la dirección en que se da la mayor varianza de los datos y la magnitud es el autovalor asociado al autovector. Luego, el siguiente autovector más grande es ortogonal a este y representa la dirección con la segunda mayor varianza y su magnitud es su autovalor asociado. De manera sucesiva se interpretan los dos autovectores faltantes, que son también ortogonales entre sí y con los anteriores.

Finalmente se encuentra la matriz de transformación de la base original en la que está expresada la matriz de covarianza a la base de autovectores. Para ello, se ubican las tuplas solución como columnas de la matriz $\mathbb{A}$, así

$$\mathbb{A} = \begin{pmatrix} 0{,}26 & -13 & -12 & 1{,}2 * 10^2 \\ -2{,}5 & 1{,}8 & -9{,}3 & -65 \\ -1{,}1 & -7{,}4 & 19 & -72 \\ -0{,}17 & 0{,}68 & 1{,}4 & 1{,}6 * 10^3 \end{pmatrix},$$

y se comprueba que es la matriz de transformación porque al efectuar el producto entre la matriz de covarianza en la base original con A se obtiene la matriz de covarianza en la base de autovectores:

$$\mathbb{A}\mathbb{C} = \begin{pmatrix} 0{,}094 & -0{,}86 & -0{,}5 & 0{,}074 \\ -0{,}91 & 0{,}12 & -0{,}38 & -0{,}041 \\ -0{,}39 & -0{,}5 & 0{,}77 & -0{,}046 \\ -0{,}063 & 0{,}046 & 0{,}057 & 1 \end{pmatrix} = \mathbb{C}^{\star}$$

4. Conclusiones y Recomendaciones

Se consiguió hallar los primeros tres momentos de un sistema de partículas a partir de la definición de momento respecto al origen. Además, se encontraron los autovectores del tensor de inercia que son los ejes principales de inercia para el sistema, logrando así obtener una base ortogonal para organizar de forma más simple la distribución de masas en los casos de 2 y 3 dimensiones. Así mismo se encontró una matriz de transformación que va de la base cartesiana a la base de autovectores. También, se comprobó que los vectores cartesianos no hacían parte del conjunto de autovectores propios del sistema mediante la definición de vector propio, por tanto no son base para la distribución.

Se logró construir la matriz de covarianza y correlación correspondiente a las inversiones de GDP de Colombia en defensa, salud, educación y ciencia para el período 2004-2018. Se encuentra que la inversión con menos varianza en este tiempo es en ciencia, mientras que la inversión en salud ha presentado mayor variación respecto a su media. Al respecto de las covarianzas, la mayor se presenta entre salud y educación, lo que implica que a medida que una tiene valores superiores a su promedio, la otra aumenta igual. Por otra parte, en la matriz de correlación, se encuentra que la correlación más fuerte se da entre la inversión en salud y ciencia.Le siguen a esta la relación entre salud y educación, y ciencia y educación; todas con valores positivos de correlación. Sin embargo, la inversión en defensa siempre está inversamente correlacionada a las demás; es decir, a medida que se aumentó la inversión en defensa, se disminuyó la inversión en ciencia, salud y educación, siendo en el sector de ciencia el más relacionado de esta manera.

En cuanto a la matriz de covarianza, se encontraron los autovectores que describen las direcciones en que aumentan en mayor medida la varianza de los datos, así como sus magnitudes; es decir, sus autovalores. Así mismo, se encontró la matriz de covarianza en la base de autovectores con su correspondiente matriz de transformación que lleva desde la base original a la base de autovectores. Se recomienda revisar la representación gráfica de los datos con sus respectivos autovectores, que permitan identificar geométricamente la varianzas.