

Compression of Images of Handwritten Numbers using PCA.

María P. Rodríguez-Muñoz¹ and Melanie A. Montañó-Ramos²

^{1,2} Tecnológico de Monterrey, Campus Guadalajara, Jalisco, Mexico.

Reception date of the manuscript:

Acceptance date of the manuscript:

Publication date: 17/11/2023

Abstract— Machine Learning models deal with the processing of massive amounts of data, in the case of image classification models processing time can become an issue due to the weight of the data images carry. A solution is to use PCA to compress images as much as possible while preserving most of their information. We apply this on a dataset of 60,000 images whom each is a 28x28 pixel image. We found that 6 principal components out of 28 can be preserved with a total variance of 93.89%. Images can be compressed even more while keeping a visually recognizable digit down to 4 principal components out of 28 with a variance of 74.87%

Keywords— Análisis de componentes principales, reducción de dimensionalidad.

I. INTRODUCCIÓN

En este proyecto, exploramos la aplicación del Análisis de Componentes Principales (PCA) para la compresión de imágenes, centrándonos en dígitos escritos a mano, con un enfoque particular en el número 9. La compresión de imágenes es esencial para reducir la complejidad de los conjuntos de datos, lo que facilita su manejo y análisis. PCA es una técnica fundamental en este contexto, ya que permite identificar las direcciones en las que los datos tienen mayor variabilidad, lo que resulta crucial para preservar la información distintiva de las imágenes mientras se reduce su dimensionalidad.

El objetivo principal es evaluar cómo PCA afecta la representación visual de los dígitos escritos a mano, especialmente el 9, al variar el número de componentes principales seleccionados. Utilizamos métodos gráficos y numéricos para determinar la cantidad óptima de componentes y examinamos la relación entre la proporción de varianza explicada y la calidad de la representación visual. Este estudio no solo contribuye al entendimiento de la aplicación práctica de PCA en la compresión de imágenes, sino que también arroja luz sobre su capacidad para preservar las características esenciales de los dígitos en el contexto específico de la escritura a mano.

II. METODOLOGÍA

a. Media aritmética

La media, promedio o valor esperado (simbolizada por \bar{X} se refiere al valor calculado a partir de la suma de los valores (X_i) de un conjunto de datos, dividido por la cantidad de estos

(N),

$$\bar{X} = \frac{\sum(X_i)}{N} \quad (1)$$

[1] La media es una medida de tendencia central que busca capturar el centro de la distribución, para dar una idea del valor al rededor del cual se agrupan los datos.

b. Varianza, desviación estándar y covarianza

La desviación estándar y la varianza son ambas medidas de dispersión que indican la variabilidad de una distribución. La desviación estándar hace referencia a la medida con la que los datos se alejan de la media de la distribución, y matemáticamente se expresa como la raíz cuadrada de la varianza; como esta distancia puede ser tanto positiva (si los datos se alejan hacia la derecha) como negativa (cuando se alejan hacia el lado izquierdo), se utiliza la varianza como el promedio del cuadrado de estas desviaciones (2), con el propósito de utilizar cifras no negativas que puedan llegar a anular a las cifras positivas, lo que podría arrojar una falsa "poca variabilidad".[2]

$$\sigma^2 = \frac{\sum(x_i - \bar{x})^2}{N} \quad (2)$$

La covarianza, por su parte, es una medida de la relación entre dos variables, es el nivel en el que dos variables van juntas. Dos características pueden tener una covarianza negativa o positiva que indica la dirección en que se mueven, positiva si tienden a aumentar o disminuir juntas, es decir, a moverse en la misma dirección, o negativa si cuando una crece la otra decrece, se mueven no hacen en direcciones opuestas. Esto se puede expresar de la siguiente manera,

$$\text{Cov}(X, Y) = \frac{\sum(X_i - \bar{X})(Y_j - \bar{Y})}{N - 1} \quad (3)$$

donde X_i y Y_i son los elementos de las matrices X y Y , y \bar{X} y \bar{Y} son sus respectivas medias. [3]

c. Estandarización de los datos

La estandarización transforma todas las variables para que tengan medidas de tendencia central iguales que las de una distribución normal, es decir, media de cero y desviación estándar de uno; esto con el objetivo de darle un mismo "formato" (z_i) a los datos que pueden llegar a estar en diferentes escalas, y poder trabajar con ellos. Para esto se toma cada valor de la variable original, se le resta su media (μ) y luego esto se divide entre la desviación estándar (σ^2) de la variable.

$$z_i = \frac{x_i - \bar{x}}{\sigma} \quad (4)$$

d. Ortogonalidad, ortonormalidad y matriz ortogonal

La ortogonalidad se define intuitivamente como algo que es perpendicular; matemáticamente hablando esto se generaliza como la idea de que el producto punto entre dos vectores es cero.

$$v_1 \cdot v_2 = 0 \quad (5)$$

El concepto de ortonormalidad implica que un conjunto de vectores es ortogonal entre sí y cada uno de ellos está normalizado para ser unitario. [4]

Una matriz cuadrada es ortogonal si $A \cdot A^T = I$, lo que implica que $A^{-1} = A^T$, en otro sentido, es necesario que tanto las filas como las columnas de la matriz sean vectores unitarios ortogonales.[5]

e. Combinación lineal

Considérese un conjunto de vectores v_1, v_2, \dots, v_n , la combinación lineal es una operación algebraica en la cual se puede conseguir cualquier vector w a partir del arreglo de los vectores, que se ajustan multiplicándose por escalares y luego sumándose [6], como se observa a continuación,

$$w = c_1 v_1 + c_2 v_2 + \dots + c_n v_n \quad (6)$$

Es la "mezcla" de los vectores base con la finalidad de formar otros vectores.

f. Conjunto generador de un espacio vectorial

Un conjunto de vectores V , es considerado un conjunto generador del espacio vectorial S si cual sea vector dentro de S puede expresarse en términos de una combinación lineal de los componentes de V . [5]

g. Transformación lineal

Una transformación es una operación algebraica que sirve para llevar un vector perteneciente a un espacio vectorial L , a otro espacio vectorial M , $T : L \rightarrow M$, a partir de una regla de transformación.

Para que una transformación sea lineal debe cumplir con las siguientes condiciones

- La transformación de una suma de vectores debe ser igual a la suma de las transformaciones. $T(\mathbf{l} + \mathbf{m}) = T(\mathbf{l}) + T(\mathbf{m})$.

- La transformación de un escalar por un vector debe ser igual al escalar multiplicado por la transformación del vector. $T(c\mathbf{l}) = cT(\mathbf{l})$ [6]

esto asegura la conservación de la estructura lineal.

h. Matriz transpuesta y matriz simétrica

La matriz transpuesta se obtiene al intercambiar los renglones y columnas (o viceversa) de un arreglo matricial. Por ejemplo,

$$A = \begin{pmatrix} 2 & 5 & 0 \\ 7 & 3 & 8 \\ 3 & 0 & 1 \end{pmatrix} \rightarrow A^T = \begin{pmatrix} 2 & 7 & 3 \\ 5 & 3 & 0 \\ 0 & 8 & 1 \end{pmatrix} \quad (7)$$

La matriz simétrica es aquella de dimensiones cuadradas que al transponerse, resulta en la misma matriz original[4]. Es decir, para una matriz A ,

$$A^T = A \quad (8)$$

i. Matriz de covarianza

Una matriz de covarianza es un arreglo matricial que encapsula la covarianza entre pares de variables, es decir, es una descripción del comportamiento correlacional entre las variables del conjunto de datos; para esto, cada celda de la matriz representa la covarianza entre las dos características que intersectan (fila y columna).

Si se tiene una matriz X de dimensiones $n \times p$, para la que n representa el número de observaciones y p es la dimensionalidad de la matriz. Su matriz de covarianza, denotada por Σ , se calcula como: [7]

$$\hat{\Sigma} = \frac{1}{(n-1)} X^T X \quad (9)$$

Como los elementos de la diagonal principal de la matriz corresponden a la relación que existe entre un componente consigo mismo, esto es equivalente a hablar de la varianza del factor pues indica en qué nivel varían sus datos. Como la covarianza cumple con la propiedad de ser conmutativa, es decir $\text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i)$, las celdas en las que los subíndices son los mismos, pero en diferente orden, tienen un mismo valor. Por lo anterior, se puede concluir que la transpuesta de la matriz es igual a la matriz original, por lo tanto, la matriz de covarianza es simétrica. [8]

$$\begin{bmatrix} \text{Var}(X_1) & \text{Cov}(X_1, X_2) & \dots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_2, X_1) & \text{Var}(X_2) & \dots & \text{Cov}(X_2, X_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_n, X_1) & \text{Cov}(X_n, X_2) & \dots & \text{Var}(X_n) \end{bmatrix} \quad (10)$$

j. Eigenvalores y eigenvectores

Sea una matriz cuadrada A de $n \times n$ y un vector columna \mathbf{v} de $n \times 1$, entonces existe un número λ llamado **valor propio** (también conocido como eigenvalor, autovalor o valor característico) de A que, al multiplicarlo por \mathbf{v} , se obtiene

$$A\mathbf{v} = \lambda \mathbf{v} \quad (11)$$

donde \mathbf{v} es conocido como el **vector propio** (llamado también **eigenvector**, **autovector** o **vector característico**) asociado a λ , y $A\mathbf{v}$ es un múltiplo escalado λ veces de \mathbf{v} . Estos eigenvalores se calculan a partir del determinante de la ecuación (11) (reacomodada en $(A - \lambda)\mathbf{v} = 0$, a través de propiedades del álgebra matricial) igualado a cero, pues de esta manera se asegura obtener una solución no trivial para los eigenvalores, lo que resulta en la **ecuación característica** (12), que al resolverse, arroja los valores propios de A . [9]

$$\det(A - \lambda) = 0 \quad (12)$$

Posteriormente, para la obtención de los *eigenvectores* asociados a cada valor λ_i , se resuelve el sistema $(A - \lambda_i I)\mathbf{v} = 0$.

En estadística, los vectores propios de una matriz de covarianza se utilizan para indicar la dirección de la máxima varianza de los datos en un dataset, mientras que los valores propios representan la cantidad de varianza a lo largo de los eigenvectores correspondientes a estos. [10]

k. Análisis de Componentes Principales

El Análisis de Componentes principales, o PCA por sus siglas en inglés, es una técnica de extracción de información, que se refiere al proceso de identificación y recuperación de datos específicos con el objetivo de reducir su dimensionalidad. El PCA transforma un conjunto de datos con atributos correlacionados y los combina para producir nuevas variables, que no están correlacionadas, lo que permite lograr una maximización en la varianza con la mínima cantidad de información posible. El PCA trabaja bajo el objetivo de poder reacomodar los datos de tal manera que se pierda la menor cantidad de información, es decir, organizarla en sus PCs, para esto, proyecta los datos en esta nueva dirección, en la que existe la máxima variación en los datos; al final, el procedimiento resulta en datos proyectados en estas nuevas PCs, que, aunque con distinto significado al original, al trabajar con estos, se estará trabajando con la mayor cantidad de información. [11]

l. Componentes principales

Las componentes principales son herramientas matemáticas que permiten simplificar la complejidad de los conjuntos de datos de alta dimensión o "high-dimensional data" (conjuntos de información con muchas variables), facilitando la visualización, el análisis y el procesamiento posterior. Al aplicar el PCA, el nuevo plano generado tiene como nuevos ejes las **componentes principales** que son ortogonales entre sí. Para lograr esta reducción de dimensionalidad, se toman las componentes más importantes, es decir, aquellas asociadas con los mayores eigenvalores de la matriz de covarianza de los datos, que se consideran las que contienen la mayor varianza y, por lo tanto, la mayor cantidad de información del conjunto de datos original. [12]

m. Estandarización

El paso base para la aplicación del PCA es la estandarización de los datos. Este proceso es crucial porque las diferentes variables en un conjunto de datos generalmente tienen rangos muy distintos, lo que influye en el tamaño de su varianza; al

manejarse números grandes, se habla también de varianzas grandes, las cuales pueden ser confundidas por PCs, o llegar a dominar el resultado del PCA al provocar que se pase por alto la información que llegan a aportar las variables con rangos mucho más pequeños [13], lo que podría llevar a interpretaciones erróneas al no dar la oportunidad de que cada variable influya en la misma medida en el resultado del análisis.

n. Formación de la matriz de covarianza

Tras la estandarización de los datos, el siguiente paso en el PCA es la construcción de la matriz de covarianza, la cual sirve para capturar la variabilidad conjunta entre las variables. Todos los vectores que se pueden generar en el *espacio vectorial* R^2 pueden ser expresados como una combinación lineal de la base generadora estándar de R^2 . De manera más intuitiva es posible visualizar lo anterior como un reacomodo de las magnitudes en las que el vector se expande, y que definen cuanto crece (o decrece) en cada eje, lo que resulta equivalente a que el vector base se gira y escala para obtener cualquier otro vector en R^2 . Matemáticamente esto se ve de la siguiente manera; sea cualquier vector \mathbf{v} en el espacio vectorial planteado, se puede expresar como $\mathbf{v} = a\mathbf{i} + b\mathbf{j}$, donde \mathbf{i} y \mathbf{j} son los vectores unitarios de los ejes x e y , respectivamente, y a y b son escalares. Esencialmente se puede decir que el PCA trabaja bajo este mismo concepto pues "gira" el espacio de datos para alinear los ejes con las direcciones de máxima varianza; es conveniente utilizar el concepto de vector propio (eigenvector) para describir la dirección de la varianza, y la magnitud de esta dispersión (el valor de la varianza) como un coeficiente λ que multiplica a este vector (eigenvalor). Con lo anterior se puede reflexionar que la suma de las varianzas de las PCs es igual a la suma de la varianza de las variables originales, ya que el PCA toma un conjunto de datos y los gira, en lugar de cambiarlos de lugar o alterar su acomodo inicial. Si se imagina que cada cara del cubo Rubik representa una variable en un conjunto de datos, y los colores de los cuadrados en cada cara representan los valores de los datos en su estado original, el cubo está alineado de cierta manera que podría no ser la óptima para ver todos los colores (variables) a la vez; el PCA es un proceso análogo a ajustar la orientación del cubo Rubik para tener una perspectiva en la que se pueda observar la mayor cantidad de colores, es decir, la mayor cantidad de información. En términos de un dataset, esto equivale a reorientar los ejes del espacio de datos de tal manera que los primeros ejes (las primeras PCs) capturan la mayor cantidad de información (varianza) posible.

o. Cálculo de los Eigenvalores y eigenvectores

Se determinan los eigenvalores de la matriz de covarianza, y sus eigenvectores correspondientes, y luego se ordenan de forma descendente según la magnitud de su valor propio. Este ordenamiento es crucial porque los primeros vectores propios (aquellos asociados a los mayores valores propios) son los que capturan la mayor parte de la varianza en el conjunto de datos. [11] Esencialmente, los vectores propios actúan como nuevos ejes en el espacio de datos transformado. Así, cada eigenvector es *n-dimensional*, pues posee tantas

coordenadas como atributos existen en el conjunto de datos original, lo que significa que cada vector propio refleja una combinación específica de las variables originales, reorientadas para revelar las direcciones de mayor variabilidad en el conjunto de datos.

p. Selección de componentes principales

Se eligen las componentes que capturen una cantidad significativa de la varianza total. Comúnmente, se busca que estas expliquen alrededor de un 90% de la varianza; para determinar cuáles elegir se utilizan diferentes métodos:

- **Análisis gráfico "Scree Plot":** esta herramienta es una forma gráfica de evaluar la relevancia de las PCs resultantes. En la *Scree Plot*, el eje x contiene las PCs ordenadas, mientras que el eje y describe los valores propios asociados a cada componente. Esta gráfica tiene una forma muy representativa, en la que se tiene una curva que empieza en la esquina superior izquierda que va cayendo conforme avanza hacia la derecha, de forma no uniforme, hasta llegar a un punto en el que prácticamente se aplanan, al que se le conoce como "codo" ("elbow" en inglés); criterio visual que suele usarse de referencia para saber hasta qué número de componentes es necesario seleccionar.[14] Este comportamiento se explica con que las primeras PCs suelen explicar la mayor parte de la variabilidad de los datos, pero a medida que se avanza, y los valores de los valores propios disminuyen, las componentes contribuyen menos a la varianza total, por lo que se descartan al considerarse de menor importancia.
- **Análisis numérico:** Otra forma de selección de las PCs es determinando numéricamente su importancia. Esto se logra calculando el porcentaje de la varianza total que explica cada componente. Para calcularlo se divide su valor propio (varianza, $Var(Z_i) = \lambda_i$) entre la suma de los valores propios (varianza total, $\sum_{j=1}^p \lambda_j$), obteniendo así porcentaje en que este aporta a la varianza total, como se ilustra a continuación, [15]

$$\frac{\lambda_i}{\sum_{j=1}^p Var(Z_j)} = \frac{\lambda_i}{\sum_{j=1}^p \lambda_j} \quad (13)$$

donde Z_j representa la j -ésima componente principal y p es el total de PCs generadas.

q. Formación de matriz de vectores propios

Una vez que se han seleccionado las PCs se procede a ponerlas dentro de una matriz en la que cada columna corresponde a una componente principal, PC_i , es decir a un eigenvector. Siguiendo la misma lógica que en los pasos anteriores, estos vectores propios (PC_i) se organizan en orden de mayor a menor en términos de su importancia (la cantidad de varianza que contienen), reflejada en los valores propios correspondientes. Esta matriz es mejor conocida como **matriz de "loadings"** pues el valor de cada coordenada corresponde a la influencia o peso que tiene una variable en esa componente en particular.

r. Transformación de los datos originales

En este paso, se proyectan los data points en el nuevo espacio definido por las PCs seleccionadas. Esto se logra a partir de la transformación de cada punto calculando el *producto punto* entre las coordenadas (x,y) del punto original x_i y cada vector columna correspondiente a la PC_j ,

$$T_{ij} = x_i \cdot PC_j \quad (14)$$

donde T_{ij} es un número que indica la nueva posición de x_i en relación con la dirección del PC_j ; es decir, es la proyección del punto x_i en la componente principal j . Se calculan las proyecciones para cada x_i en todas las PCs; las nuevas coordenadas son

$$p_i = [T_{i1}, T_{i2}, \dots, T_{im}], \quad (15)$$

representando p_i la nueva posición de x_i en el espacio reducido.

s. Herramientas computacionales

Los métodos matemáticos responsables de la compresión de imágenes fueron computarizados en Jupyter Notebook en el lenguaje computacional Python. Hicimos uso de las librerías:

1. "pandas" y "numpy": para el manejo y transformación de matrices y datos.
2. "matplotlib": visualización de matrices a imágenes.
3. "math": algunos métodos matemáticos.
4. "PCA" de "sklearn.decomposition": para la reducción de dimensionalidad de las imágenes a través de PCA.

Nuestra base de datos tiene una dimensionalidad (60000, 785), 60 000 filas y 785 columnas. Cada fila contiene la información de la imagen de un solo número, la primera columna de cada fila contiene el dígito asociado con la imagen. Primero transformaremos un solo dígito de una fila aleatoria, para ver de cerca el proceso, y próximamente realizaremos una visualización grupal.

III. RESULTADOS

Se eligió transformar la fila 1459 la cual representa un número 9. Con la función "reshape" del módulo "numpy", cambiamos las dimensiones de la matriz para convertirla de una matriz unidimensional (1, 784) a una matriz bidimensional(28, 28), este paso es esencial ya que las funciones que usaremos a continuación no manejan matrices unidimensionales. Visualizamos la matriz con la función "imshow" de "matplotlib".

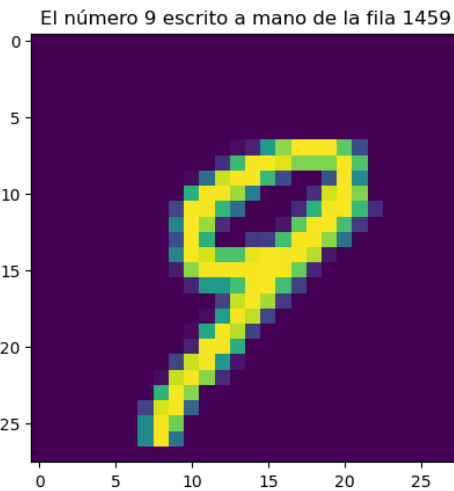


Fig. 1: Imagen original de la fila 1459.



Fig. 3: Imagen después de transformación de la fila 1459.

a. ¿Cuántas dimensiones reducir?

Usando PCA queremos reducir la dimensionalidad de nuestros datos lo mayor posible preservando la información mas relevante. El objetivo es preservar al menos 90% de la varianza total, descompondremos nuestro número-matriz y visualizaremos la diferencia de varianzas entre componentes principales haciendo uso de un "scree plot".

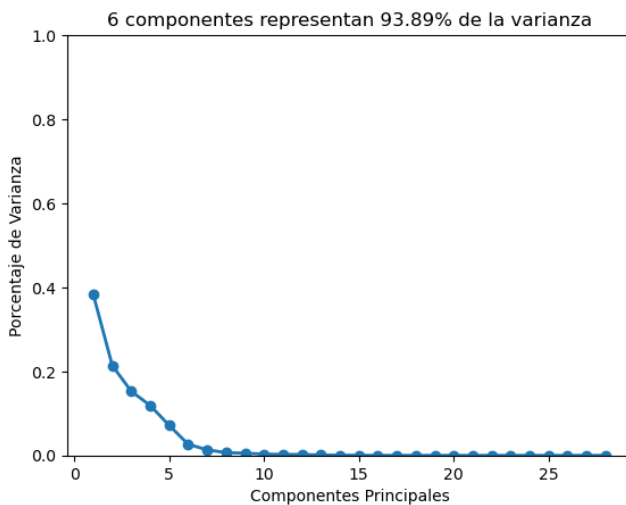


Fig. 2: "Scree plot" de los eigenvectores de la fila 1459.

Los componentes principales que yacen en el eje x están acomodados en orden descendiente a partir de su eigenvalores. Por ello la gráfica asume la figura de una línea descendiente. A partir del "scree plot", podemos con seguridad transformar los datos manteniendo únicamente 6 de los 28 componentes principales, todo esto preservando 93.89% de la información original.

b. Transformando con PCA

Haciendo uso de la función "fit_transform" de "PCA", reducimos los datos a la dimension deseada. Después reconstruimos los datos transformados a la matriz original con la función "inverse_transform" del módulo "PCA".

Se intentaron algunas otras reducciones a continuación:



Fig. 4: Imagen de la fila 1459 después de transformación, donde se preservaron 5 componentes principales con una varianza total de 86.75%

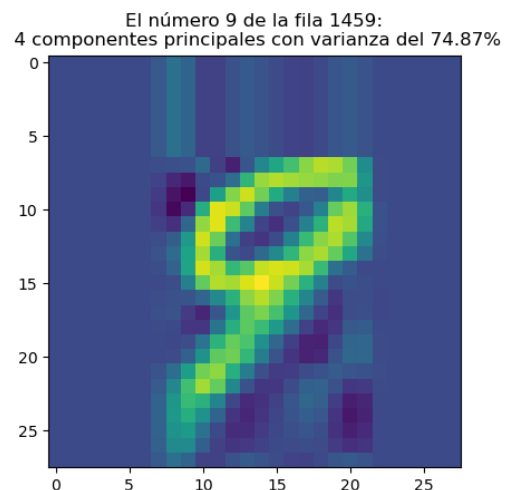


Fig. 5: Imagen de la fila 1459 después de transformación, donde se preservaron 4 componentes principales con una varianza total de 74.87%

La imagen es de considerablemente de menor calidad cuando los componentes principales son 4 y la varianza se redujo a 74.87%. Sin embargo, se puede seguir identificando claramente que se trata de un número 9.

IV. CONCLUSIONES

Este proyecto abordó la compresión de imágenes de dígitos escritos a mano mediante el uso del Análisis de Componentes Principales (PCA). La aplicación de PCA demostró ser eficaz al reducir la dimensionalidad de las imágenes mientras se mantenía una representación fiel del número 9, objeto de estudio. La selección óptima de componentes se basó en métodos gráficos y numéricos, como el "Scree Plot" y la proporción de varianza explicada. Los resultados destacan la utilidad del PCA como una herramienta valiosa para el tratamiento de conjuntos de datos de alta dimensionalidad, ofreciendo una perspectiva efectiva en la compresión de imágenes digitales.

REFERENCES

- [1] J. F. Healey, *Statistics: A Tool for Social Research*, 10th ed. Cengage Learning, 2015. [Online]. Available: <https://www.amazon.com/STATISTICS-RESEARCH-10TH-JOSEPH-HEALEY/dp/1305235290>
- [2] S. Jackson, *Statistics Plain and Simple*. Cengage Learning, 2016. [Online]. Available: <https://books.google.com.mx/books?id=9sEaCgAAQBAJ>
- [3] A. o Entidad Responsable, "Covariance," 2020. [Online]. Available: <https://corporatefinanceinstitute.com/resources/data-science/covariance/>
- [4] D. Poole, *Álgebra lineal*, 4th ed. Cengage Learning, 2017. [Online]. Available: <https://latam.cengage.com/libros/algebra-lineal/>
- [5] J. Barra Escutia and R. Larson, *Matemáticas IV: álgebra Lineal*. Cengage Learning, 2013. [Online]. Available: <https://0-elibro-net.bibliotecails.tec.mx/es/ereader/consorcioitems/40172?page=1>
- [6] F. Guzmán Aguilar, *Álgebra Lineal: Serie Universitaria Patria*. Grupo Editorial Patria, 2014.
- [7] H. R. Frost, "Eigenvectors from eigenvalues sparse principal component analysis," *Journal of Computational and Graphical Statistics*, vol. 31, no. 2, pp. 486–501, 2022. [Online]. Available: <https://doi.org/10.1080/10618600.2021.1987254>
- [8] F. Pozo and N. Parés, *Descomposición en valores singulares: introducción y aplicaciones. Análisis de componentes principales (PCA) y descomposición en valores singulares (SVD)*. Fundació per a la Universitat Oberta de Catalunya, 2019, ch. Álgebra lineal, pp. 1–58. [Online]. Available: <https://upcommons.upc.edu/handle/2117/182401>
- [9] Z. D. G. Cullen Michael R., Wright Warren S., *Matemáticas Avanzadas para Ingeniería*, 4th ed. México: McGRAW-HILL/INTERAMERICANA EDITORES, S.A. DE C.V., 2012.
- [10] H. Abdi, "The eigen-decomposition: Eigenvalues and eigenvectors," *Encyclopedia of measurement and statistics*, pp. 304–308, 2007.
- [11] J. S. Teija Seitola, Visa Mikkola and H. Järvinen, "Random projections in reducing the dimensionality of climate simulation data," *Tellus A: Dynamic Meteorology and Oceanography*, vol. 66, no. 1, p. 25274, 2014. [Online]. Available: <https://doi.org/10.3402/tellusa.v66.25274>
- [12] E. Taskesen, "What are pca loadings and biplots?" *Towards Data Science*, 2022. [Online]. Available: <https://towardsdatascience.com/what-are-pca-loadings-and-biplots-9a7897f2e559>
- [13] R. Bro and A. K. Smilde, "Principal component analysis," *Anal. Methods*, vol. 6, pp. 2812–2831, 2014. [Online]. Available: <http://dx.doi.org/10.1039/C3AY41907J>
- [14] S. Mangale, "Scree plot," 2020. [Online]. Available: <https://sanchitamangale12.medium.com/scree-plot-733ed72c8608>
- [15] C. G. Martínez, "Análisis de componentes principales (pca)," Fecha de último acceso, Última actualización hace aproximadamente 4 años. [Online]. Available: https://rpubs.com/Cristina_Gil/PCA