

## Procesos Estocásticos 86.09

FACULTAD DE INGENIERÍA UNIVERSIDAD DE BUENOS AIRES

# TRABAJO PRÁCTICO N° 1 PRINCIPAL COMPONENT ANALYSIS (PCA)

### Grupo 6

<b>Autor</b>	<b>Padrón</b>	<b>Mail</b>
Delfino Ortega, Gabriel	90689	gdelfino@fi.uba.ar
Hoffmann, Nicolas	103733	nhoffmann@fi.uba.ar
Lützel Schwab, Nahila	100689	nlutzel Schwab@fi.uba.ar
Pareja, Facundo Jose	99719	fpareja@fi.uba.ar

**Nota: Aprobado.**  
**Corrector: Raymundo Albert.**

4 DE ABRIL DEL 2024

## Índice

1. Introduccion	2
2. Ejercicio 1	3
3. Ejercicio 2	5
4. Ejercicio 3	7
5. Conclusiones	8
6. Bibliografia	10

## 1. Introduccion

En muchos ambitos del procesamiento digital de la información es necesario reducir el tamaño de los datos, ya sea para almacenamiento, procesamiento, analisis, transmisión, etc. En lo que respecta al audio, hay muchas técnicas de compresión de datos estandar, algunas de ellas son MP3, ACC, FLAC. Dichas técnicas reducen el tamaño conservando casi intacta la calidad de audio. Estos estandares utilizan técnicas más sofisticadas y eficientes respecto de la que se analizará en el siguiente documento.

En el siguiente trabajo se va a utilizar la técnica estadística PCA(Principal Component Analysis), la cual permite reducir las dimensiones de un conjunto de datos con el fin de disminuir su tamaño y manteniendo intacta la mayor parte de la información. En este caso, se va realizar la compresión y descompresión de los archivos de audio con formato *.wav*.

Se van a realizar los siguientes ejercicios:

### Ejercicio 1: Correlación

En este ejercicio vamos a analizar gráficamente el concepto de correlación entre componentes de un vector aleatorio generado desde la señal de audio.

a) Abrir la señal de audio *audio\_01\_2024a.wav*, normalizarla dividiéndola por su valor eficaz  $x(n) = \text{audio}(n)/\text{RMS}(\text{audio}(n))$ . Luego segmentar la señal para guardar la colección de vectores  $X_m$  con  $L = 2$ .

Nota: definimos vectores de dimensión 2 con el propósito de visualizar la correlación entre las dos componentes del vector. Luego este análisis se podrá generalizar a mayores dimensiones.

b) Hacer un gráfico de dispersión para ver gráficamente la correlación entre las dos componentes del vector  $X_m = [X_0^{(m)} X_1^{(m)}]^T$  (muestras vecinas en el dominio del tiempo). Graficar también los histogramas de ambas componentes  $X_0^{(m)}$  y  $X_1^{(m)}$ .

Sugerencia: para una mejor comprensión de los resultados, puede ubicar el histograma de  $X_0^{(m)}$  en la parte posterior del gráfico y el histograma de  $X_1^{(m)}$  a la derecha, respecto del gráfico de dispersión.

c) Obtenga la matriz de covarianza  $C_X$  del vector aleatorio definido previamente y luego las matrices de autovectores  $V$  y de autovalores asociados  $\Lambda$ .

d) Obtenga un conjunto de vectores  $Y_m$  cuyas componentes están descorrelacionadas. Luego haga un gráfico de dispersión de  $Y$  y los histogramas asociados a cada elemento de los vectores resultantes de forma análoga al punto b). Observe los resultados y saque conclusiones.

Note que solo se pide proyectar en el espacio de autovectores, pero no aplicamos la reducción de componentes aún.

### Ejercicio 2: Compresión

En este caso vamos a aplicar la compresión mediante PCA con un vector aleatorio de mayor dimensión. Vamos a fijar un tamaño de ventana  $L = 1323$  muestras correspondiente a 30 ms de señal a 44100 Hz).

a) Aplicar el método PCA: Asumiendo  $CR = 70\%$ , obtener la colección de vectores  $\hat{Y}_m$  de componentes principales. Utilizar como señal el archivo *audio\_02\_2024a.wav*.

b) Volver a transformar los vector  $\hat{Y}$  al espacio original y reconstruir la señal de audio. Reproducir las señales original y reconstruida para evaluar subjetivamente las diferencias.

c) Repita los puntos anteriores pero para tasas de compresión  $CR = 90\%$ ,  $CR = 95\%$ . Analice nuevamente los resultados al reproducir las señales reconstruidas luego de la compresión.

### Ejercicio 3: MSE vs CR

a) Aplicar la compresión seguida de la descompresión para los archivos *audio\_01\_2024a.wav*, *audio\_02\_2024a.wav* y *audio\_03\_2024a.wav*, utilizando diferentes tasas (recuerde normalizar cada señal). Suponga los siguientes casos:  $CR = \{10\%, 20\%, \dots, 90\%\}$ . Haga un gráfico MSE vs CR para cada audio y compárelos en simultáneo.

b) Qué puede concluir acerca de la calidad de la reconstrucción observando el MSE entre los distintos audios?

## 2. Ejercicio 1

En este ejercicio se buscó un primer análisis de los audios mediante la correlación. Para esto se tomó el audio indicado en el enunciado y se lo normalizó según:

$$x(n) = \frac{\text{audio}(n)}{\text{RMS}(\text{audio}(n))} \quad (1)$$

Una vez normalizado el audio se segmentó la señal utilizando un ancho de ventana  $L = 2$  para obtener un vector aleatorio.

La interpretación de esto es la siguiente: el proceso de segmentación es un experimento aleatorio que consta en tomar  $L$  muestras continuas. Por ende, se obtiene un vector aleatorio de  $L$  variables aleatorias. La cantidad de experimentos realizados (el largo de cada variable aleatoria) dependerá de la cantidad de muestras de cada audio.

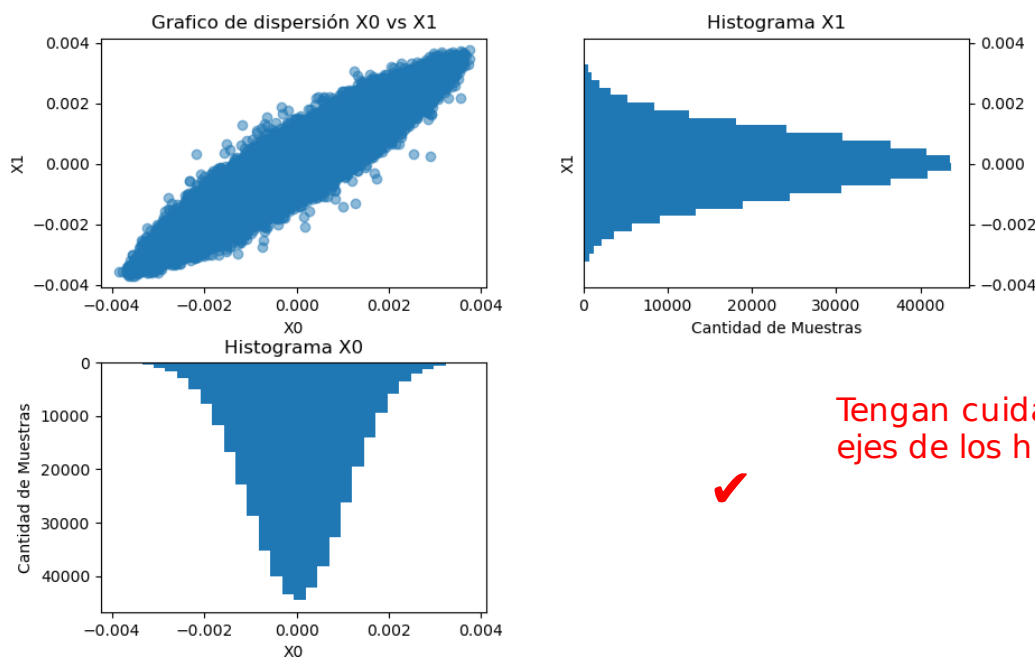
Dado el  $L$  mencionado anteriormente, el vector aleatorio será de dos componentes (dos variables aleatorias),

$$X = [X_0^{(m)} X_1^{(m)}]^T \quad (2)$$

Con  $m$  la cantidad de realizaciones.

Obtenido el vector, dado a que contiene solo dos variables aleatorias, es posible realizar un gráfico de dispersión.

**Cuidado. La norma de la señal no es el RMS!**



**Tengan cuidado con los ejes de los histogramas.**

Figura 1: Grafico de dispersión de  $X_0$  y  $X_1$  (centro). Histograma de  $X_0$  (abajo). Histograma de  $X_1$  (derecha)

Como se puede ver en la figura 2, el diagrama de dispersión sugiere una correlación positiva entre ambas variables dado a la tendencia lineal que muestran los puntos. Luego, por los histogramas, se puede inferir que ambas variables siguen una distribución normal estándar.

Luego se estima la matriz de covarianza según,

$$\hat{C}_X = \frac{1}{M} \sum_{m=0}^{M-1} (X_m - \hat{\mu}_X)(X_m - \hat{\mu}_X)^T \quad (3)$$

Donde  $\hat{\mu}_X$  es la esperanza estimada obtenida como,

$$\hat{\mu}_X = \frac{1}{M} \sum_{m=0}^{M-1} X_m \quad (4)$$

**Esto no es correcto. Por qué infieren esto?**

En este caso, la matriz tiene dimensión de  $\mathbb{R}^{2 \times 2}$ . Utilizando este estimador se obtuvo la siguiente matriz de correlación,

$$C_X = \begin{pmatrix} 1,0917 * 10^{-6} & 1,0439 * 10^{-6} \\ 1,0439 * 10^{-6} & 1,0915 * 10^{-6} \end{pmatrix} \quad (5)$$

Los elementos de la diagonal corresponden a la varianza de cada variable aleatoria los cuales tienen un valor similar pero no igual. Los otros términos de la matriz corresponden a la covarianza entre ambas variables aleatorias, estos dos términos son idénticos dado a que  $Cov(X_0, X_1) = Cov(X_1, X_0)$ , si bien este no es el coeficiente de correlación, se puede afirmar que hay correlación positiva tanto por el gráfico como por el signo de la covarianza. El coeficiente de correlación obtenido fué,

$$\rho = 0,956 \quad \checkmark \quad (6)$$

Ahora se busca un nuevo vector  $Y$  cuyas variables estén descorrelacionadas. Para esto, se realiza una descomposición espectral de la matriz  $C_X$ .

$$C_X = V \Lambda V^T \quad (7)$$

Con  $V$  matriz de autovectores de  $C_X$  y  $\Lambda$  matriz de autovalores donde cada autovalor corresponde a la varianza de cada variable aleatoria. Luego, definiendo la transformación,

$$Y = V^T X \quad (8)$$

Como  $C_Y = V^T C_X V$  entonces reemplazando por 2,

$$C_Y = \Lambda \quad (9)$$

Entonces  $Y$  es un vector cuyas componentes están descorrelacionadas. Probando esta transformación en el programa implementado para este trabajo se obtiene el siguiente gráfico de dispersión.

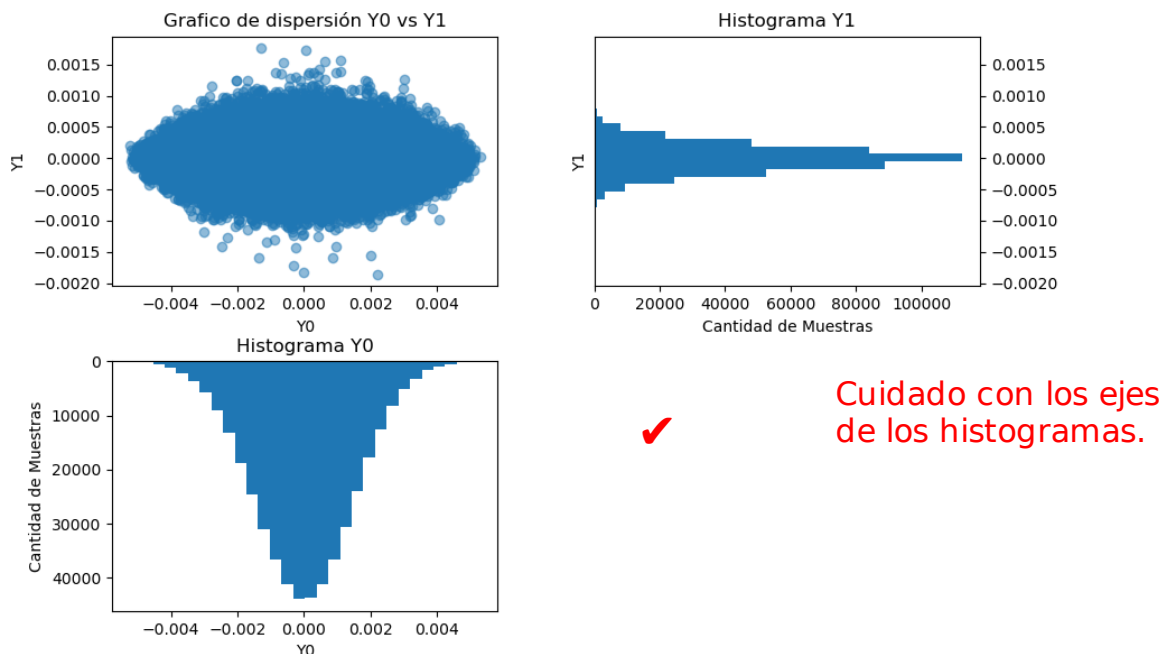


Figura 2: Grafico de dispersión de  $Y_0$  y  $Y_1$ (centro). Histograma de  $Y_0$ (abajo).Histograma de  $Y_1$ (derecha)

Aquí se puede ver que claramente que ambas variables aleatorias están descorrelacionadas. Por lo que coincide con el resultado teórico.

### 3. Ejercicio 2

En el presente ejercicio se realizó una compresión mediante PCA utilizando un vector aleatorio de mayor dimensión. Para ello se definió una matriz  $U$  que contenga los autovectores principales de  $V$  asociados a los autovalores de mayor peso, es decir, se conservan las componentes que presentan mayor variabilidad. Luego de aplicar PCA, se proyecta  $X_m$  en el espacio de  $U$  obteniendo:

$$\hat{Y}_m = U^T \cdot X_m \quad (10)$$

y por último se recuperó el vector original mediante la transformación inversa:

$$\hat{X}_m = U \cdot \hat{Y}_m \quad (11)$$

Para este caso se tomó un tamaño de ventana de  $L = 1323$  muestras y se realizó el mismo procedimiento considerando diferentes tasas de compresión  $CR = 70\%$ ,  $90\%$  y  $95\%$ , esto es, quitando un  $70\%$ ,  $90\%$ ,  $95\%$  de autovalores a la matriz diagonal según la siguiente expresión,

$$CR = \left( 1 - \frac{\text{cantidad de componetnes principales}(K)}{\text{cantidad de componentes totales}(L)} \right) \times 100\% \quad (12)$$

Se reprodujeron las señales originales y las reconstruidas para evaluar las diferencias y se pudo concluir que a medida que aumenta la tasa de compresión existe una perdida en la calidad del audio. Cuando la tasa de compresión es baja la señal no llega a distorsionarse demasiado ya que se elimina información redundante, pero cuando dicha tasa es alta como los valores evaluados es posible notar la perdida de información útil.

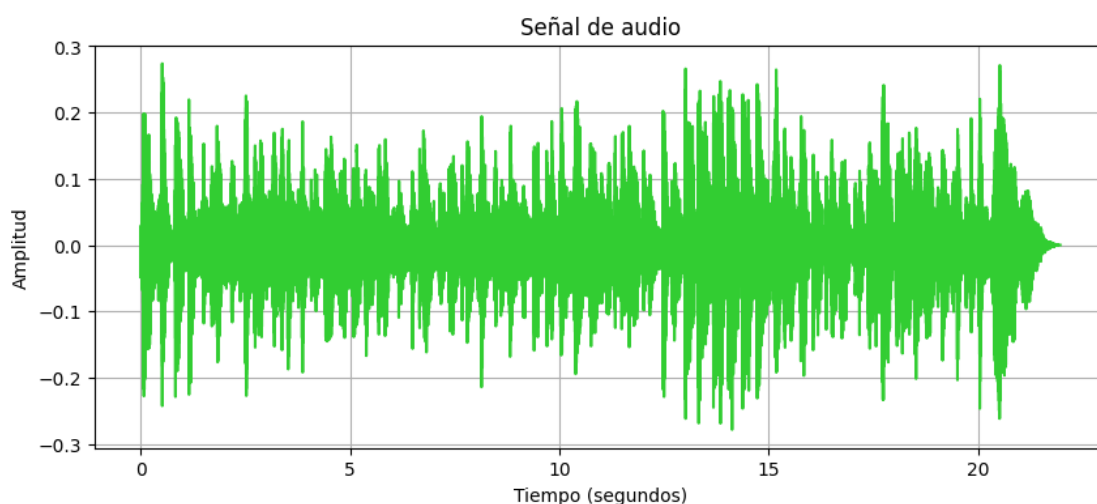


Figura 3: Gráfico de la señal original

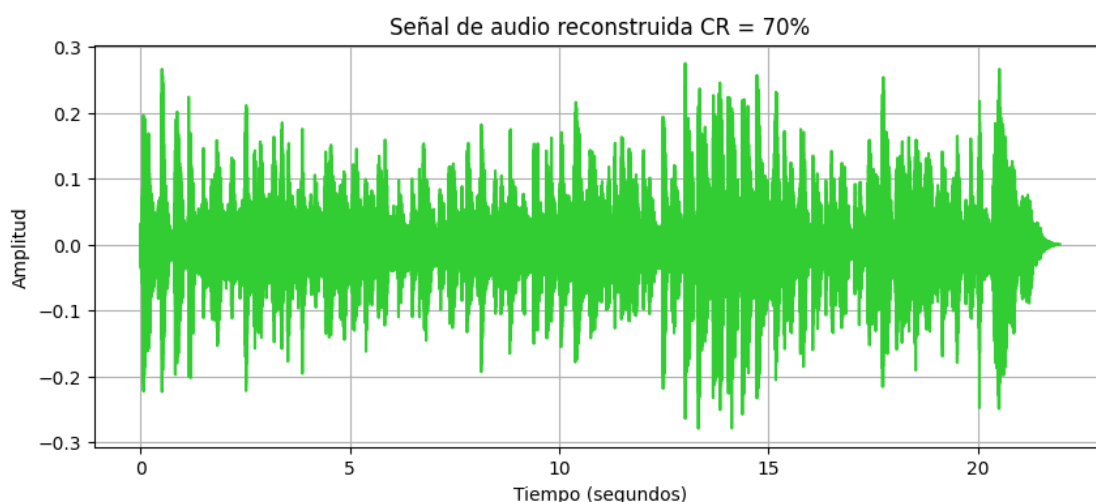


Figura 4: Gráfico de la señal reconstruida con  $CR = 70\%$

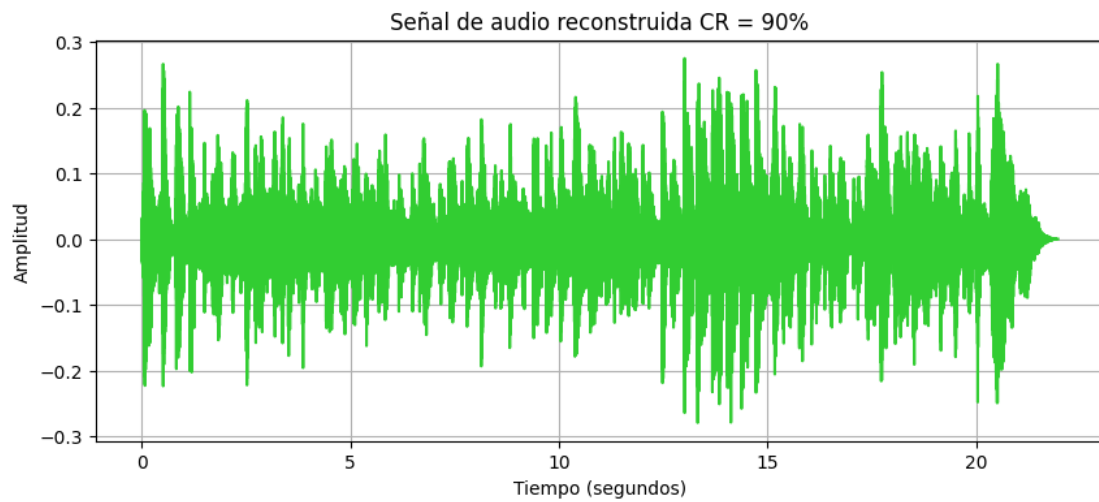


Figura 5: Gráfico de la señal reconstruida con CR = 90 %

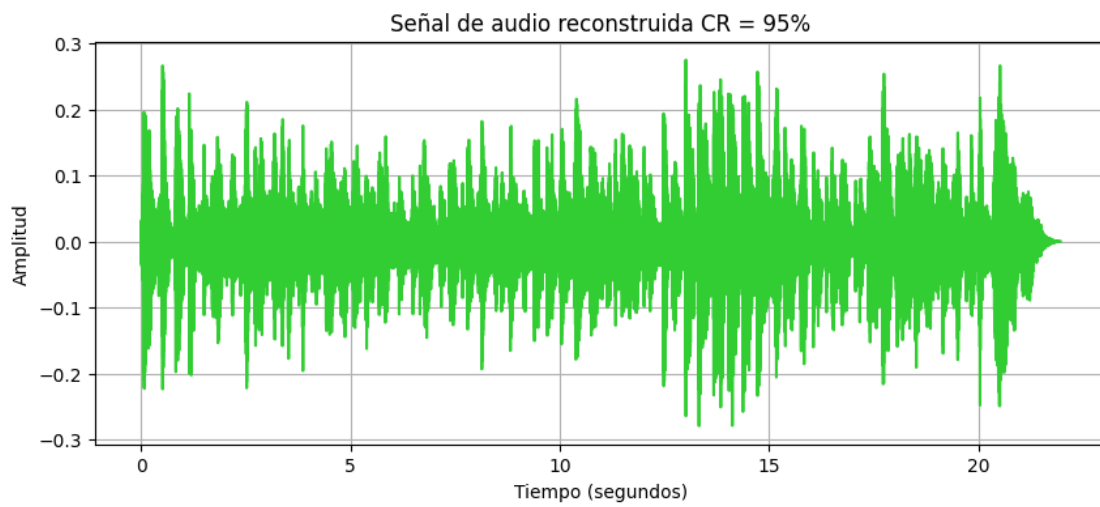


Figura 6: Gráfico de la señal original y la reconstruida con CR = 95 %

## 4. Ejercicio 3

En este ejercicio se repitieron los pasos del ejercicio 2 pero para tres audios, con tasas de compresión (*compression rates*) del 10 % al 90 %, con aumentos incrementales del 10 % (y una ultima con 99 %). A partir de los audios reconstruidos y los originales se efectuó el calculo del MSE (*mean square error*) con la siguiente formula:

$$MSE = \frac{1}{N} \sum_{n=0}^{N-1} (x(n) - \hat{x}(n))^2 \quad (13)$$

Los resultados fueron graficados contra las tasas de compresion y en simultaneo para todos los audios para facilitar comparacion:

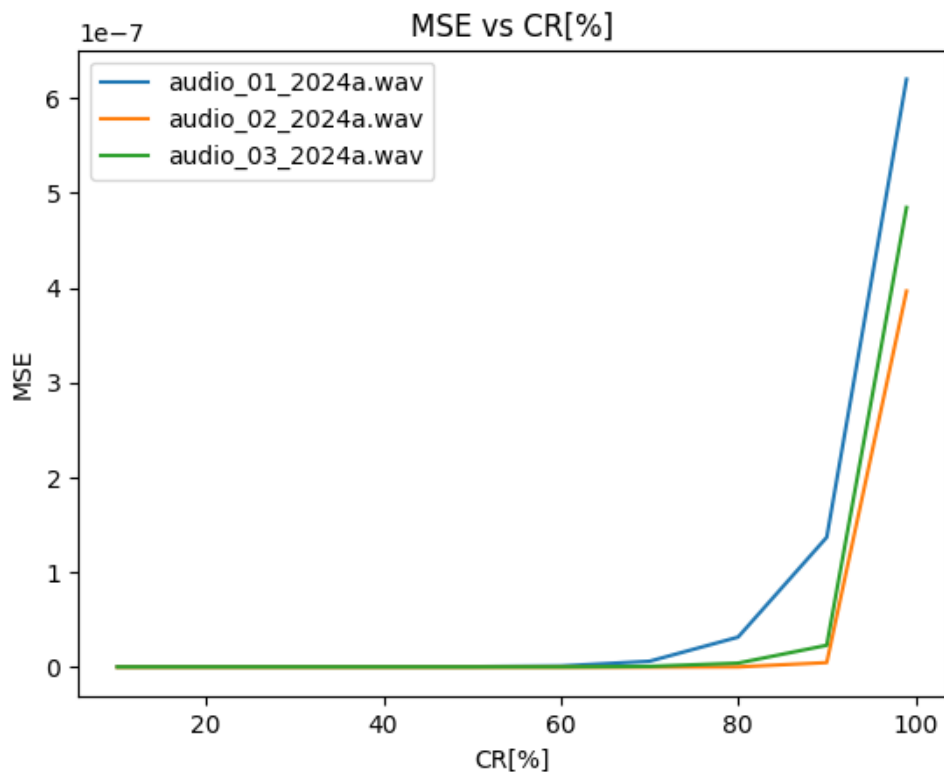


Figura 7: Comparacion de MSE y CR para tres audios

Se puede observar que el MSE mantiene niveles bajos hasta una cierta tasa de compresión, luego de la cual crece rapidamente. En todos los casos este valor se encuentra por encima del 60 %, lo cual verifica que todos estos audios son compresibles sin gran perdida de información. El valor exacto de la tasa varia según el audio

- Poco mas del 60 % para el audio 1
- Alrededor del 70 % para el audio 2
- Poco mas del 80 % para el audio 3

Claramente conviene realizar un analisis similar para decidir la compresión mas conveniente para un audio en particular. Por otro lado se desprende que el aumento de error se debe a la perdida de las componentes principales más importantes. ✓



## 5. Conclusiones

No se entiende. Uno utiliza PCA para describir un conjunto de datos cualesquiera en terminos de nuevas variables no correlacionadas en dimensiones más chicas.

Se puede concluir que el método PCA permite trabajar con componentes ortogonales descorrelacionadas que presentan la mayor variabilidad de los datos, lo que facilita eliminar componentes redundantes o ruido. Por lo tanto, una mayor correlación lineal entre los componentes implica una mayor reducción posible del tamaño total del audio mediante compresión.

El analisis del MSE vs. CR permite determinar a partir de que punto no conviene continuar la compresion. Para el caso especifico de los audios analizados, es posible reducir todos al 40 % de su tamaño original sin incurrir en un error grande. En general, a partir del 70 % de tasa de compresion se comienza a apreciar la degradacion en calidad, y luego del 90 % es notoria.

Las dificultades encontradas fueron las siguientes:

- La longitud del *array* resultante de la lectura del audio no era un multiplo del factor L elegido. Esto se soluciono removiendo elementos del *array* hasta alcanzar una longitud que fuese multiplo.
- Falta de familiaridad con las librerias utilizadas para procesar los datos (*numpy* y *matplotlib*), lo cual compenso con lectura de la documentacion oficial disponible *online*.

## Índice de figuras

1.	Grafico de dispersión de $X_0$ y $X_1$ (centro). Histograma de $X_0$ (abajo).Histograma de $X_1$ (derecha)	3
2.	Grafico de dispersión de $Y_0$ y $Y_1$ (centro). Histograma de $Y_0$ (abajo).Histograma de $Y_1$ (derecha) .	4
3.	Gráfico de la señal original . . . . .	5
4.	Gráfico de la señal reconstruida con CR = 70 % . . . . .	5
5.	Gráfico de la señal reconstruida con CR = 90 % . . . . .	6
6.	Gráfico de la señal original y la reconstruida con CR = 95 % . . . . .	6
7.	Comparacion de MSE y CR para tres audios . . . . .	7

## 6. Bibliografia

- [1] Steven M. Kay. Intuitive probability random processes using MATLAB. Springer 1951.