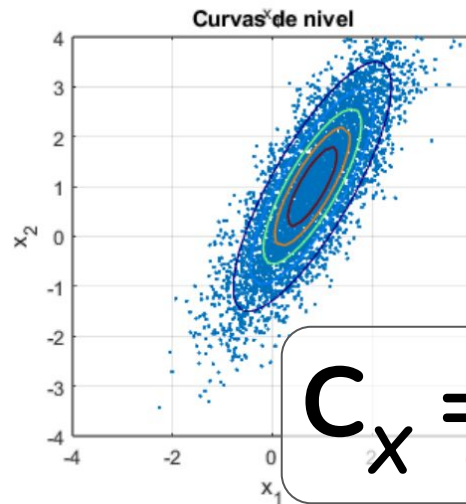
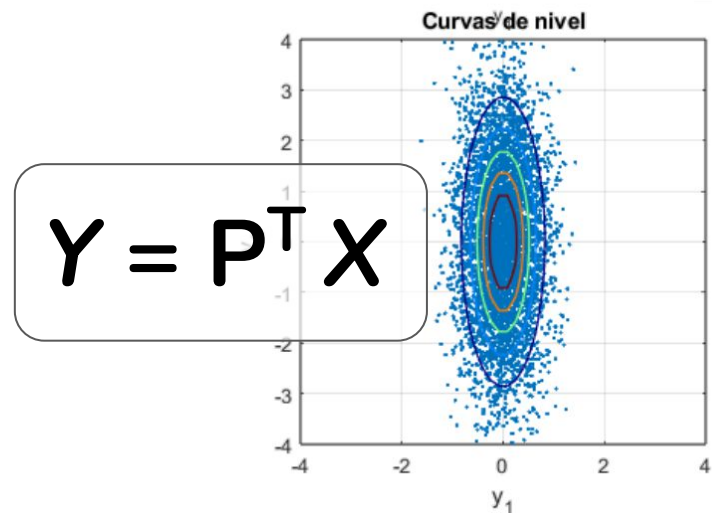


# Procesos estocásticos (86.09)

Análisis de componentes principales (PCA)



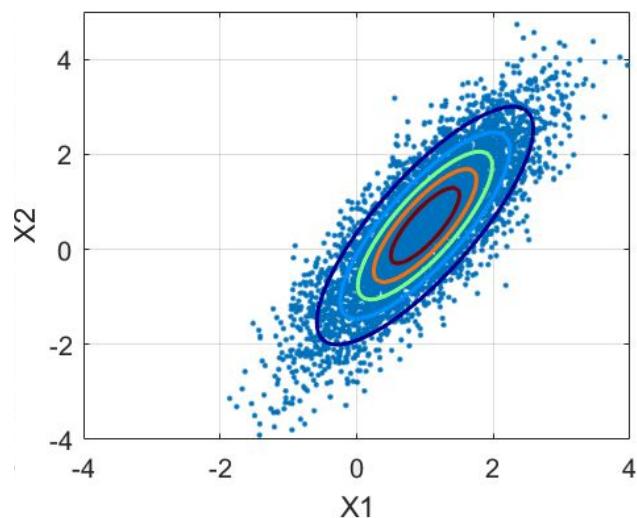
$$C_X = P \Lambda P^T$$



$$Y = P^T X$$

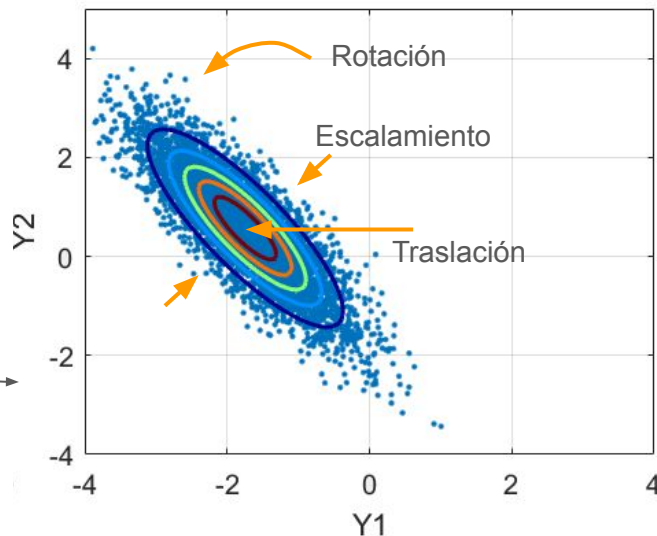
# Transformación Afín

Si  $\mathbf{X}$  es un VeA de media  $\mu_{\mathbf{x}}$  y covarianza  $C_{\mathbf{x}}$ . Podemos aplicar una transformación Afín para obtener otro VeA  $\mathbf{Y}$  de media  $\mu_{\mathbf{y}}$  y covarianza  $C_{\mathbf{y}}$ . Esto implica una traslación, escalamiento y rotación. Por ejemplo ( $\mathbf{X} \in \mathbb{R}^2$ ):



$$\mathbf{Y} = \mathbf{A} \mathbf{X} + \mathbf{b}$$

En particular, si  $\mathbf{X}$  es Normal,  $\mathbf{Y}$  es Normal



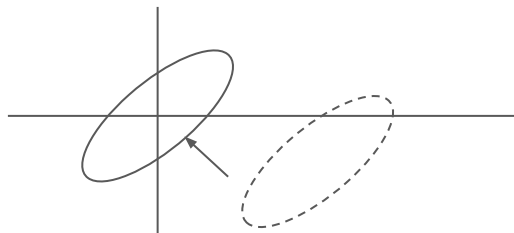
# Transformación Afín - Descorrelación y centrado

Buscamos generar un VeA  $\mathbf{Y}$ , **centrado** (media nula) y **descorrelacionado** (covarianza diagonal) a partir de otro VeA  $\mathbf{X}$  **arbitrario** mediante una **transformación afín**.

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \quad \mathbf{A} \in \mathbb{R}^{m \times n} \text{ y } \mathbf{b} \in \mathbb{R}^m \quad \text{tal que } C_{\mathbf{X}} = P_{\mathbf{X}} \Lambda_{\mathbf{X}} P_{\mathbf{X}}^T$$

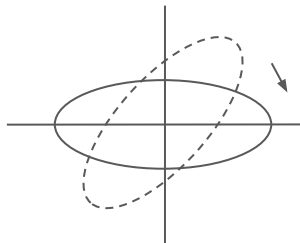
## Centrado

Si se define:  $\mathbf{b} = -\mathbf{A}\mu_{\mathbf{X}}$



## Descorrelación

Si se define:  $\mathbf{A} = P_{\mathbf{X}}^T$



# Análisis de componentes principales (PCA)

# PCA

Dado un VeA  $\mathbf{X} = [X_1, X_2, X_3, \dots, X_N]^T \sim N(\boldsymbol{\mu}_X, C_X)$  tal que  $C_X = P\Lambda P^T$ , con autovectores  $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N\}$  y autovalores  $\{\lambda_1, \lambda_2, \dots, \lambda_N\}$  (asumiendo  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$ ).

Si suponemos que hay un conjunto de autovalores  $\{\lambda_{r+1}, \dots, \lambda_N\}$  mucho más pequeños que el resto, tal que  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \gg \lambda_{r+1} \geq \dots \geq \lambda_N$ .

Podemos aproximar cualquier realización de  $\mathbf{X}$  descartando las direcciones de principales de menor peso  $\{\mathbf{p}_{r+1}, \dots, \mathbf{p}_N\}$  y conservando las direcciones principales  $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_r\}$  asociadas a los  $r$  autovalores más grandes (comprimir información)

# PCA

Dado  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \gg \lambda_{r+1} \geq \dots \geq \lambda_N$ .

$$Y_1 = \mathbf{p}_1^T \mathbf{X}$$

$$Y_2 = \mathbf{p}_2^T \mathbf{X}$$

:

$$Y_r = \mathbf{p}_r^T \mathbf{X}$$

Componentes con  
la mayor parte de  
la información

$$Y_{r+1} = \mathbf{p}_{r+1}^T \mathbf{X}$$

:

$$Y_N = \mathbf{p}_N^T \mathbf{X}$$

Componentes con menor  
variación (proyecciones  
sobre direcciones  
principales de menor peso)

$$Y_i = \mathbf{p}_i^T \mathbf{X} \quad / \quad Y_i \perp Y_j$$

Reconstrucción aproximada

$$\mathbf{X}_R = \sum_{i=1}^r \mathbf{Y}_i \mathbf{p}_i$$

Reconstrucción exacta

$$\mathbf{X} = \sum_{i=1}^r \mathbf{Y}_i \mathbf{p}_i + \sum_{i=r+1}^N \mathbf{Y}_i \mathbf{p}_i$$

# PCA

Dado  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r \gg \lambda_{r+1} \geq \dots \geq \lambda_N$ .

$$\mathbf{X} = [X_1, \dots, X_N]^T \quad \mathbf{Y} = [Y_1, \dots, Y_N]^T \text{ Variables}$$
$$\mathbf{Y}_R = [Y_1, \dots, Y_r]^T \text{ desacopladas}$$

Reconstrucción exacta

$$\mathbf{X} = \sum_{i=1}^N \mathbf{Y}_i \mathbf{p}_i \quad \left\{ \begin{array}{l} \mathbf{Y} = \mathbf{P}^T \mathbf{X} \\ \mathbf{X} = \mathbf{P} \mathbf{Y} \end{array} \right.$$

Reconstrucción aproximada

$$\mathbf{X}_R = \sum_{i=1}^r \mathbf{Y}_i \mathbf{p}_i \quad \left\{ \begin{array}{l} \mathbf{Y}_R = \mathbf{V}^T \mathbf{X} \\ \mathbf{X}_R = \mathbf{V} \mathbf{Y}_R \end{array} \right.$$

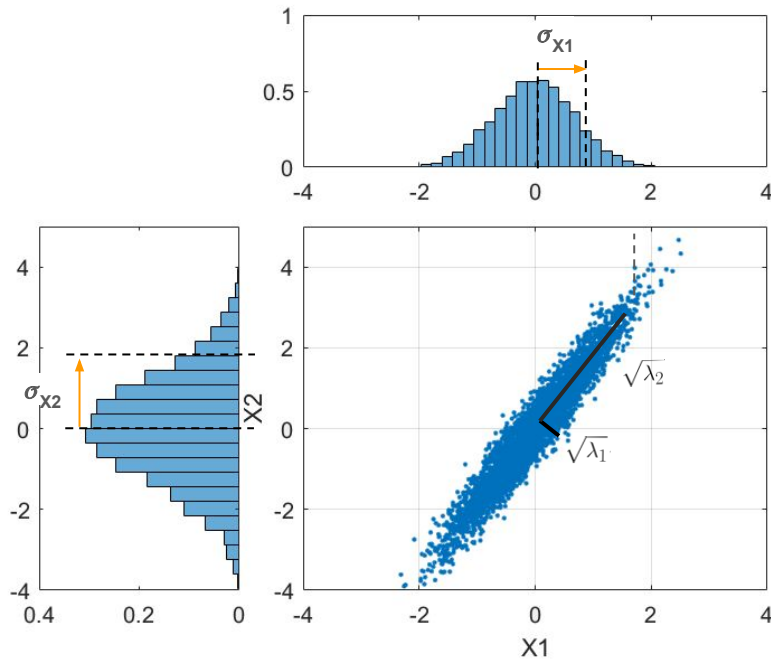
$$\mathbf{P} = \begin{pmatrix} | & | & \dots & | & | & \dots & | \\ \mathbf{p}_1 & \mathbf{p}_2 & \dots & \mathbf{p}_r & \mathbf{p}_{r+1} & \dots & \mathbf{p}_N \\ \downarrow & \downarrow & & \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda_N \end{pmatrix}$$

$$\mathbf{V} = \begin{pmatrix} | & | & \dots & | \\ \mathbf{p}_1 & \mathbf{p}_2 & \dots & \mathbf{p}_r \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

# PCA

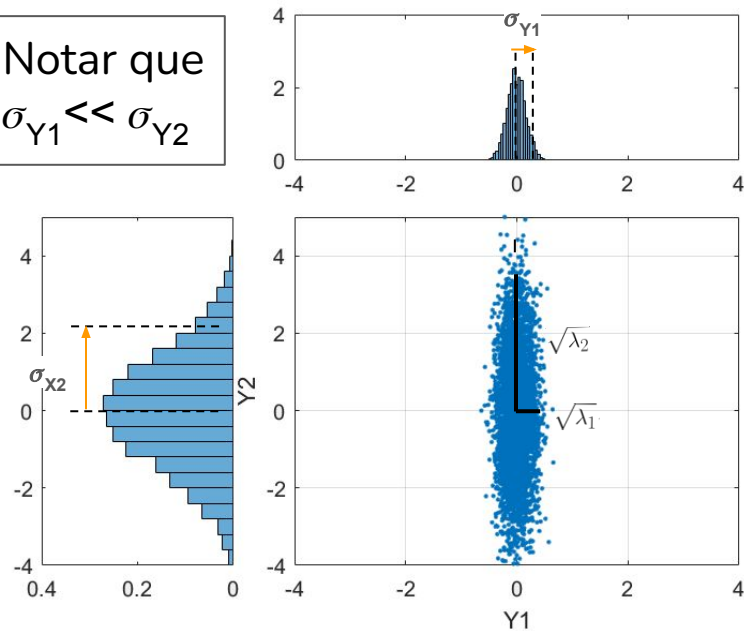
$$\mathbf{X} = [X_1, X_2]^T \quad / \quad \mathbf{C}_X = \mathbf{P} \mathbf{\Lambda} \mathbf{P}^T$$



$$\mathbf{Y} = [Y_1, Y_2]^T \quad / \quad \mathbf{Y} = \mathbf{P}^T \mathbf{X}$$

Notar que  
 $\sigma_{Y_1} \ll \sigma_{Y_2}$

$$\mathbf{Y} = \mathbf{P}^T \mathbf{X}$$





**Suponiendo el caso más general, donde  $\mathbf{X}$  posee una media  $\boldsymbol{\mu}_x$**

Proyección en las direcciones principales con reconstrucción exacta

$$\mathbf{Y} = \mathbf{P}^T (\mathbf{X} - \boldsymbol{\mu}_x)$$

$$\mathbf{X} = \mathbf{P} \mathbf{Y} + \boldsymbol{\mu}_x$$

Proyección en las direcciones principales con reconstrucción aproximada

$$\mathbf{Y}_R = \mathbf{V}^T (\mathbf{X} - \boldsymbol{\mu}_x)$$

$$\mathbf{X}_R = \mathbf{V} \mathbf{Y}_R + \boldsymbol{\mu}_x$$

# PCA

## Resumen del método PCA:

1. Obtenemos realizaciones de  $\mathbf{X} = [X_1, X_2, \dots, X_N]^T$ ,
2. Estimamos media  $\boldsymbol{\mu}_x$  y matriz de covarianza  $C_x$
3. Obtener matrices de autovectores  $P$  y autovalores  $\Lambda$ .
4. Ordenar autovectores en orden decreciente de autovalores.
5. Definir cantidad  $r$  de componentes a conservar y definir matriz truncada  $V$ .
6. Calcular las componentes proyectadas y reducidas  $\mathbf{Y}_R = V^T(\mathbf{X} - \boldsymbol{\mu}_x)$ .
7. Guardar las componentes  $\mathbf{Y}_R$  para más tarde.
8. Reconstruir  $X$  a partir de las componentes desacopladas  $\mathbf{X}_R = \mathbf{V} \mathbf{Y}_R + \boldsymbol{\mu}_x$

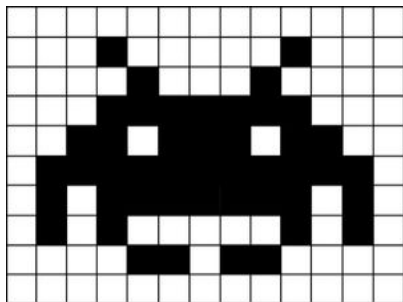
# Actividad 1:

## PCA aplicado a imágenes

# Actividad 1: PCA aplicado a imágenes

## Representación de una imagen

Blanco &  
Negro



0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	1	0	0	0	0	0
0	0	0	0	1	0	0	0	1	0	0	0	0	0	0
0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	1	1	0	1	1	1	0	1	1	0	0	0	0
0	1	1	1	1	1	1	1	1	1	1	1	0	0	0
0	1	0	1	1	1	1	1	1	1	0	1	0	0	0
0	1	0	1	0	0	0	0	0	1	0	1	0	0	0
0	0	0	0	1	1	0	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Escala  
de grises



128	128	160	128	128	96	128	128	128	128	128
128	160	128	64	32	0	0	192	160	128	128
160	160	32	32	0	0	32	64	128	160	160
192	64	128	255	255	255	192	96	64	192	192
160	96	192	255	192	192	255	96	64	192	192
160	64	160	160	192	192	128	64	0	160	160
160	64	160	0	160	128	0	32	0	128	128
192	160	128	255	192	128	160	64	64	255	255
160	192	96	255	160	64	128	32	192	255	255
160	192	128	192	192	96	96	64	255	192	192
160	192	160	64	128	64	32	96	255	192	192
160	192	255	0	0	0	96	128	255	255	255

# Actividad 1: PCA aplicado a imágenes

## Ayudas

Abrir una imagen:

```
img = imread('img_01.jpg');
```

Convertir a escala de grises

```
img_gris = rgb2gray(img);
```

Convertir a tipo de dato double

```
data = double(img_gris);
```

Aplanar una matriz

```
tira = data(:);
```

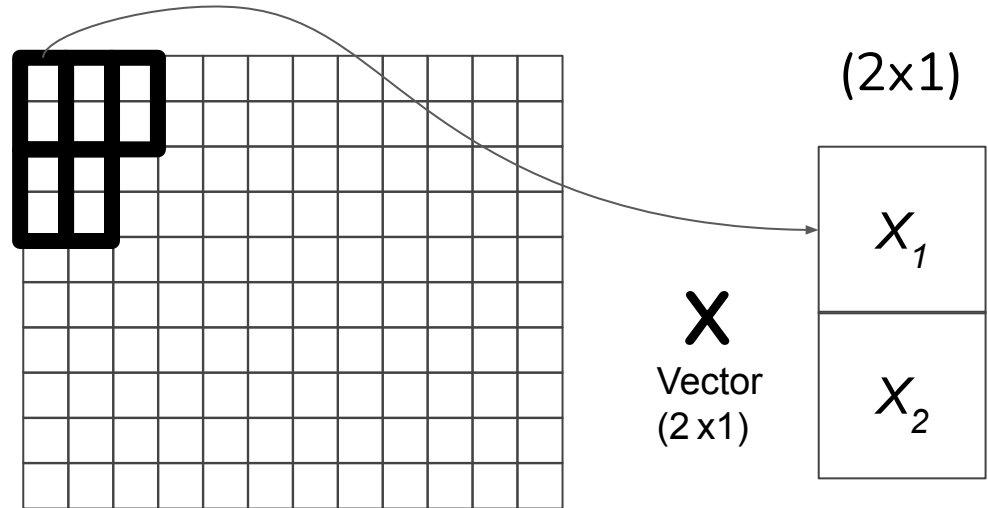
Rearmar imagen

```
img_out = reshape(XR, [fil, col]);
```

mostrar imagen

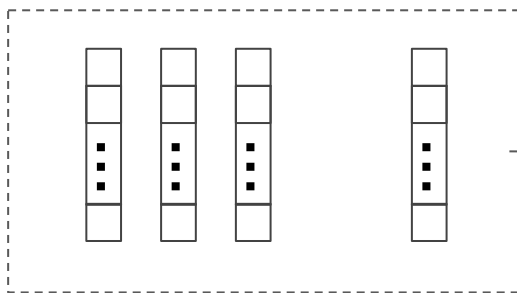
```
imshow(uint8(img_out))
```

Generar observaciones de vectores (2x1) a partir de cada par de píxeles contiguos de una imagen



# Actividad 1: PCA aplicado a imágenes

Observaciones de  $\mathbf{X} = [X_1 \ X_2 \ \dots \ X_N]^T$



$$Cx = P\Lambda P^T$$

$$P =$$

$$\begin{pmatrix} \downarrow & \downarrow & \downarrow & & \downarrow \\ \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 & \dots & \mathbf{p}_{64} \\ \downarrow & \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

$$\Lambda =$$

$$\begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \lambda_{64} \end{pmatrix}$$

Conservamos los  $r$  autovalores asociados a los mayores  $r$  autovalores

$$V = [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3 \dots \mathbf{p}_r] \quad (\text{suponer: } \lambda_1 > \lambda_2 > \dots > \lambda_r > \lambda_{r+1} > \dots > \lambda_N)$$

# Actividad 1: PCA aplicado a imágenes

Para la imagen **img\_01.png** provista en el campus, defina realizaciones de un vector  $\mathbf{X} = [X_1, X_2]^T$  seleccionando cada par de píxeles contiguos en la imagen (ver esquema en siguiente filmina)

1. Haga un gráfico de dispersión del vector  $\mathbf{X}$ .
2. Estime la matriz de covarianza de  $\mathbf{X}$  y compute las proyecciones  $\mathbf{Y}$  en las direcciones principales. Haga un gráfico de dispersión de  $\mathbf{Y}$ . Determine cuál de las componentes puede ser descartada.
3. Defina la matriz de proyección  $V$  que descarta el autovector asociado al menor autovalor. Obtenga la proyección  $\mathbf{Y}_R$  de  $\mathbf{X}$  en ese espacio reducido.
4. Reconstruya el conjunto de vectores  $\mathbf{X}_R$  con la transformación inversa.
5. Rearme la imagen a partir de los vectores reconstruidos y gráfíquela

# Trabajo práctico 1: Ej 4

**Original**



**Comprimida (CR = 5 %)**



**Comprimida (CR = 10 %)**



**Comprimida (CR = 15 %)**



**Comprimida (CR = 20 %)**



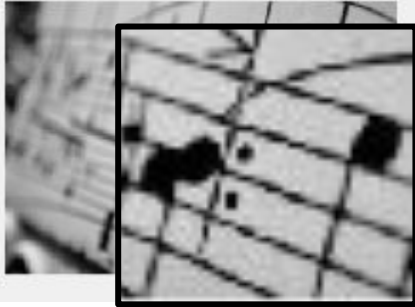
**Comprimida (CR = 25 %)**



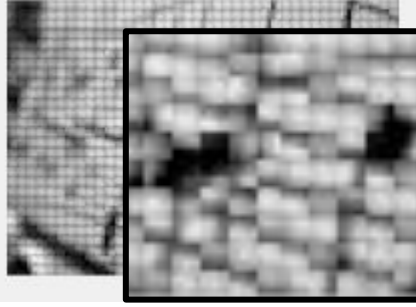


# Trabajo práctico 1: Ej 4

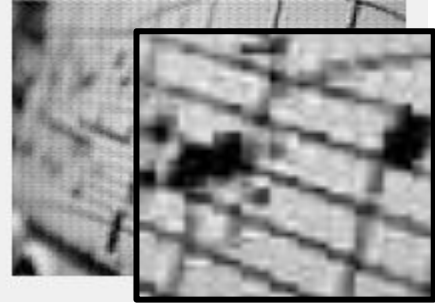
Original



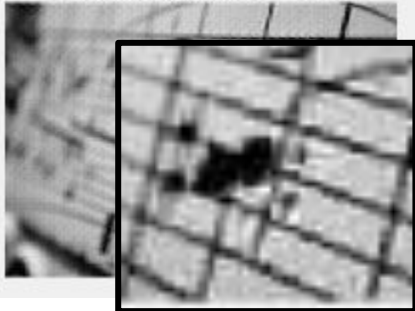
Comprimida (CR = 5 %)



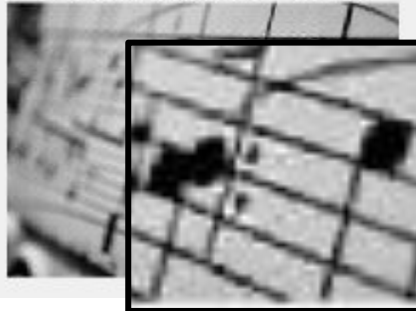
Comprimida (CR = 10 %)



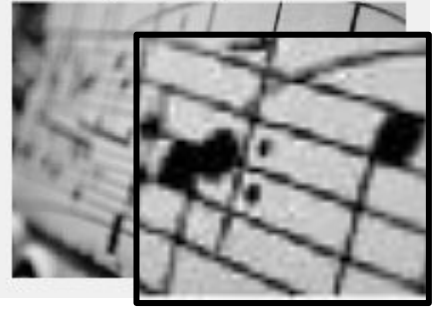
Comprimida (CR = 15 %)



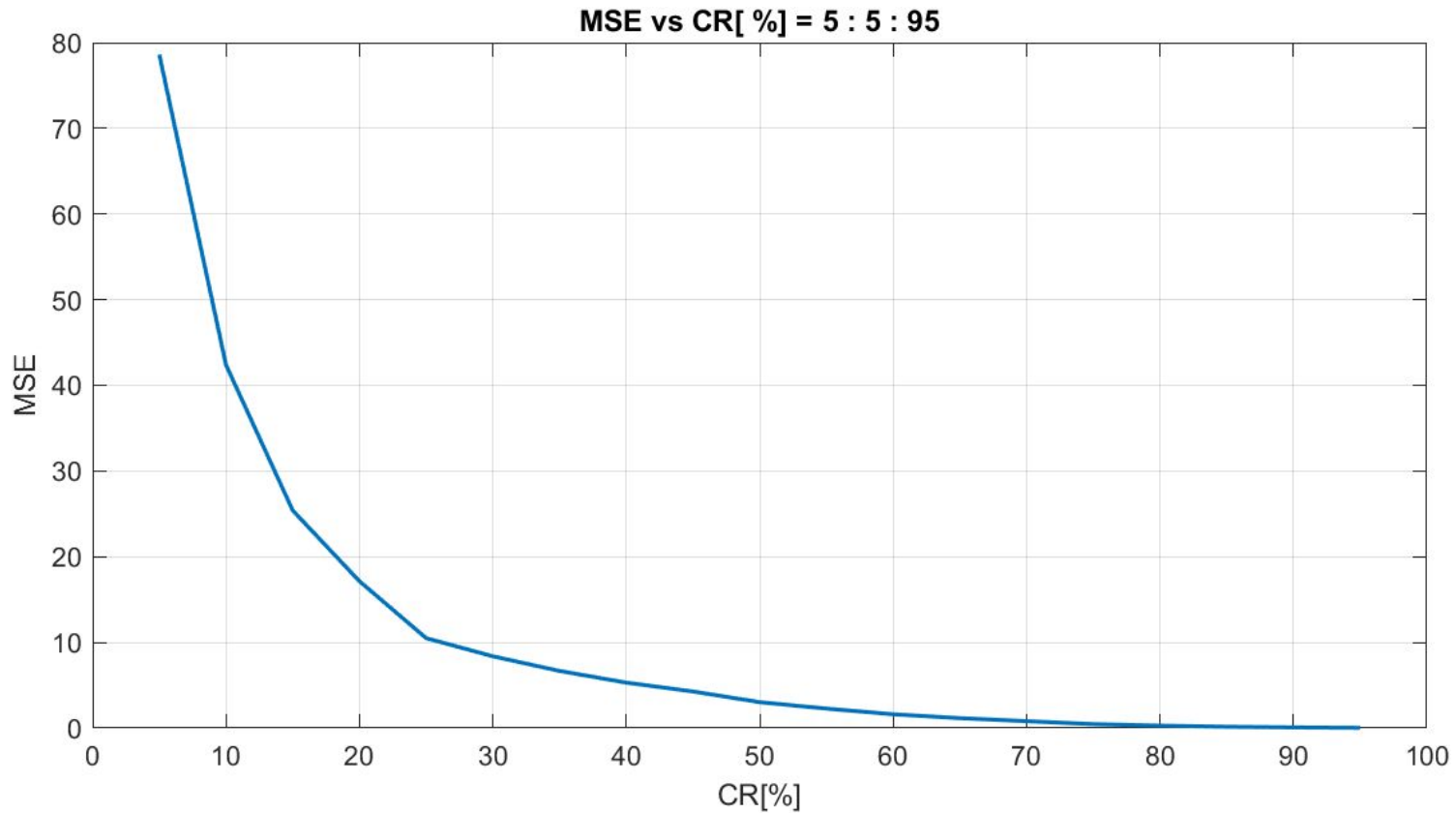
Comprimida (CR = 20 %)



Comprimida (CR = 25 %)



# Trabajo práctico 1: Ej 4



# Actividad 2

# Actividad 2

Sea un vector aleatorio  $\mathbf{X}$ , con  $C_{\mathbf{X}} = \begin{bmatrix} 1 & -0.25 & 0.1 \\ -0.25 & 2 & -0.25 \\ 0.1 & -0.25 & 3 \end{bmatrix}$  con media  $\boldsymbol{\mu}_{\mathbf{X}} = \mathbf{0}$ .

1. Genere  $M=200$  realizaciones de  $\mathbf{X}$  y estime las varianzas de  $X_1$ ,  $X_2$  y  $X_3$ .
2. Aplique la proyección en las componentes principales, definiendo la matriz de proyección  $V$  para retener sólo las componentes asociadas a los  $r$  autovalores de mayor peso. Considere los casos  $r=1$  y  $r=2$ . Compare el MSE (valor promedio de  $\|\mathbf{X}_R - \mathbf{X}\|^2$  para todas las realizaciones) para ambos valores de  $r$ . Nota:  $\mathbf{X}_R = V \mathbf{Y}_R$ .