



## Cuestionario 1

Santiago Neiszer, Marco Osornio

27 de mayo de 2024

### 1. Ejercicio 1

- (1) Verdadero, porque el PCA halla la máxima variabilidad y con eso forma la cámara de Weyl
- (2) Verdadero, porque son ortogonales
- (3) Verdadero, por Hotelling-Fisher, se trabaja con los datos reduciendo lo más posible la dimensión  $p \ll n$

### 2. Ejercicio 2

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

**Calcula la primer componente principal**

**Primer paso: Calcular los eigenvalores**

$$\det(\mathbf{XX}^T - \lambda \mathbf{I}) = 0$$

$$\det \left( \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) = 0$$

$$\det \begin{pmatrix} 1-\lambda & 0 & 0 \\ 0 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{pmatrix} = 0$$

$\lambda_1 = 2$  con multiplicidad 2 y  $\lambda_2 = 1$

**Segundo paso: Meterlos a la cámara de Weyl**

**Tercer paso: Calcular los eigenvectores**

$$\det(XX^T - 2I) \cdot \mathbf{v}_1 = 0$$

$$\det \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0$$

Así:  $\mathbf{v} = (1, 0, 0)$

### 3. Ejercicio 3

Suponga que se tiene la siguiente tabla:

Observación	$X_1$	$X_2$
1	-2	2
2	2	-2

- La primer componente (loading) del eigenvector que resuelve el problema de optimización del PCA para  $X_1$  es 0.7071.
- La primer componente (loading) del eigenvector que resuelve el problema de optimización del PCA para  $X_2$  es negativa.

**Calcula  $Xb_1$**

$$b_1 = \begin{pmatrix} 0.7071 \\ Xb_1 \end{pmatrix}$$

Así el eigenvector puede ser cualquier vector en el espacio generado por el primer eje, también sabemos que:

$$b_{1_1}^2 + b_{1_2}^2 = 1$$

$$1 - b_{1_1}^2 = b_{1_2}^2$$

$$b_{1_2}^2 = 1 - 0.7071^2$$

$$b_{1_2} = \sqrt{1 - 0.7071^2}$$

$$b_{1_2} = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 0.7071 \\ -0.7071 \end{pmatrix}$$

$$b_{1_2} = \begin{pmatrix} 1 - 2(0.7071) + 2(-0.7071) \\ 2(0.7071) - 2(0.7071) \end{pmatrix}$$

$$b_{1_2} = \begin{pmatrix} -1.8284 \\ 2.8284 \end{pmatrix}$$

### 4. Ejercicio 4

Es fácil ver que en la figura izquierda que los ejes coordenados son los que corresponden a los de los componentes principales, al buscar la ortogonalidad mientras estos se rotan para que no exista una correlación.

## 5. Ejercicio 5

Como  $X$ ,  $X_1$ ,  $X_2$  son variables aleatorias gaussianas con media cero y matriz de covarianza identidad  $W(0, 1)$  y  $Y_1 = 0.5X + 2X_1$ , y  $Y_2 = -0.5X + 2X_2$ , así  $Y$ :

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 0.5 & 2 \\ -0.5 & 2 \end{pmatrix}$$

Dado que  $Y$  comp  $Y$  es una com. lineal de variables gaussianas,  $Y$  es gaussiano.

$$\text{Cov}(Y) = \begin{pmatrix} 0.5 & 2 \\ -0.5 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0.5 & -0.5 \\ 2 & 2 \end{pmatrix}$$

Para los valores propios :

— $\text{Cov}(Y) - \lambda I = 0$ ,  $Y$  dan como resultado:  $\lambda_1 = 9.25$ ,  $\lambda_2 = 0.75$ . Los eigenvectores que tenemos son:

$$\lambda_1 = 9.25 :$$

$$v_1 = \begin{pmatrix} 0.803 \\ -0.596 \end{pmatrix}$$

$$\lambda_2 = .75 :$$

$$v_2 = \begin{pmatrix} -.596 \\ .803 \end{pmatrix}$$

$$\text{Cov}(\epsilon_1, \epsilon_2) = 0$$

## 6. Ejercicio 6

Supongamos que el eigenvector de la primera componente principal del conjunto de datos es  $(0.707, -0.5, -0.5)$ . Calcula la proporción de la varianza explicada por el primer componente.

$$v_1 = (0.707)^2 + (-0.5)^2 + (-0.5)^2$$

$$v_1 = 0.5 + 0.25 + 0.25$$

$$\lambda_1 = 1 :$$

La proporción de la varianza es:

$$1/1 = 1 :$$

## 7. Ejercicio 7

- (a) Falso, La proposición de varianza explicada por un componente principal adicional sí puede disminuir a medida que se agregan más componentes principales.
- (b) Verdadero, La proposición acumulativa de la varianza explicada nunca decrece a medida que se agregan más componentes principales.
- (c) Falso, Usar todas las posibles componentes principales no necesariamente provee de un mejor entendimiento de los datos.
- (d) Verdadero, La gráfica scree sí provee un método para determinar el número de componentes principales a usar.

## 8. Ejercicio 8

Podemos ver que todas las aseveraciones son falsas, ya que en el PCA se hace una transformación ortogonal y en ninguna de las gráficas observamos dicha transformación ortogonal.

## 9. Ejercicio 9

- (1) Falso.
- (2) Verdadero.
- (3) Verdadero.

## 10. Ejercicio 10

Para poder saber cual de los vectores representa mejor la primera componente principal, debemos calcular los vectores y valores propios de la matriz de covarianza de los datos.

Supongamos que ya normalizamos los datos sacando su desviacion estandar y su media cero.

Tenemos la matriz covarianza  $M$ :

$$M = (1/n)X^T X$$

Siendo  $n$  el numero de observaciones.

Entonces calculamos los valores y vectores propios de  $M$ .

Despues tenemos que el vector propio asociado con el valor mas grande es la direccion de la primera componente principal.

Como no tenemos ni los datos ni la matriz de covarianza podemos analizar las características de los vectores pedidos.

- (A) (1,1,1,1) Nos dice que las características tienen el mismo peso, por lo que no hay una característica sobresaliente.
- (B) (0.5,0.5,-0.5,0.5) Este vector parece tener un equilibrio teniendo características positivas y negativas.
- (C) (1,-1,1,-1) Este es similar al anterior pero con la diferencia de que sus características son cotrapuestas con un peso de 1.
- (D) (0.7071,0,-0.7071,0) En este vector podemos ver que está ponderado más las características diagonales de la matriz de covarianza.
- (E) (0.5,0.5,0.5,0.5) Similar al primero, todas las características tienen el mismo peso.

Si la matriz de covarianza refleja una relación lineal más fuerte entre las características, esperaríamos que el vector propio esté más alineado con esas características.

Por lo que, el vector propio que representa la mejor primer componente principal probablemente es D o B, ya que parecen sugerir una dirección donde las características están relacionadas de manera más significativa. Pero la verdadera elección final dependerá de cómo están distribuidas todas las características y de los datos.