

# Taller 2 - Matrices y Sistemas de Ecuaciones

Álgebra en ciencia de datos

Universidad EAFIT

Fecha de Entrega: Octubre 8 de 2023

## Problema 1

1. Un estudiante del curso de álgebra para ciencia de los datos debe hacer varios cálculos durante tres días en tres procesadores A B C. El primer día gasta 2, 4 y 5 horas en los procesadores A B C, respectivamente y por ello paga 220 dólares. El segundo día, las horas contratadas en cada procesador son 6, 9 y 8, respectivamente, y paga 490 dólares. El tercer día contrata 4.1, 5 y 3 horas, respectivamente en A B C y paga 274 dólares. ¿Cuál es el precio por hora en cada procesador?
2. El problema es que el estudiante comete un error y en lugar de 4.1 horas en el primer procesador del tercer día, pone 4.2 horas. El error es mínimo se podría pensar. Resuelva el sistema para ver como cambia el valor de la hora en cada procesador, antes y después del error. Explique si el cambio es grande o no y por qué se da este cambio. Consulte una cota superior para el error relativo en la solución de un sistema cuando hay un cambio en los coeficientes de la matriz del sistema. Luego calcule una cota superior para el error relativo en el valor de la hora en los procesadores.

## Problema 2

1. Descargue un conjunto de datos y elija 4 variables numéricas según su interés. Calcule el vector de medias, la matriz de covarianzas y su inversa.
2. Llame  $Z$  a la matriz de datos obtenida al restar la media de cada observación. ¿Cuál es la relación entre  $Z^T Z$  y la matriz de covarianza?

3. Calcule la distancia de Mahalanobis de cada observación respecto al vector de medias. Realice una gráfica de dispersión de variables dos a dos coloreando cada observación de acuerdo con la distancia calculada.
4. Identifique la bola de radio 1 con centro en el vector de medias respecto a la distancia de Mahalanobis en la gráfica del punto anterior.

### Problema 3

Implemente el algoritmo de substitución hacia adelante para encontrar la solución al sistema lineal  $Lx = b$ , donde  $L$  es una matriz triangular inferior. En caso de infinitas soluciones, el algoritmo debe retornar una solución particular y una base para el espacio nulo de  $L$ .

### Problema 4

1. Sea  $H_n(i, j) = \frac{1}{i+j-1}$ , llamada la matriz de Hilbert. Simule 1000 datos normales con matriz de covarianza  $H_n$ . Estime la matriz de covarianzas desde los datos simulados. Realice una gráfica de  $n$  en el eje  $x$  con el número condición de la matriz de covarianza estimada en el eje  $y$ . Recuerde que el número de condición es  $\|A\| \|A^{-1}\|$ . ¿Que tipo de comportamiento observa?
2. Considere  $x = [1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ 15]$ . Defina  $b = H_{15}x$  y resuelva numéricamente el sistema con la forma  $x = H^{-1}b$ . ¿Qué conclusión obtiene? Busque alternativas para resolver el problema observado.