

Notas importantes

- Todos los ejercicios resueltos serán subidos a los repositorios de github. La forma en la que el estudiante debe presentar este laboratorio es:
 - Debes crear una carpeta local llamada Practica1.
 - Dentro de la carpeta Practica1, se deben crear las carpetas Ejercicio1, Ejercicio2, Ejercicio3, Ejercicio4, ... donde se deben alojar las soluciones con la extensión Rmd para los comentarios que son parte de las preguntas y .R del lenguaje R para los otros. Cualquier otra extensión, incluyendo letras minúsculas, será motivo de una rebaja de puntos en el ejercicio.
- Los archivos de respuesta deben llevar un comentario inicial con tu nombre y código. Por ejemplo.

```
> # Nombre : Cesar Lara Avila 20122345J
> # Respuesta1: El codigo muestra el uso del SVD en el calculo
> #de la inversa de matrices
>
>
> library(MASS)
>
> a <- matrix(c(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0,
+ 0, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1), 9, 4)
>
> a.svd <- svd(a)
> a.svd$d
```

- El laboratorio calificado es individual, cualquier copia de código puede implicar la descalificación de este laboratorio.

Preguntas

1. (a) Crea y almacena una secuencia de valores del 5 al -11 que progresa en pasos de tamaño 0.3.
(b) Sobrescribe el objeto en (a) usando la misma secuencia con el orden invertido.
(c) Repite el vector

```
> c(-1,3, -5,7, -9)
```

dos veces, con cada elemento repetido 10 veces y almacena el resultado. Visualiza el resultado ordenado de mayor a menor.

- (d) Crea y almacena un vector que contenga, en cualquier configuración, lo siguiente:
 - Una secuencia de enteros del 6 al 12 (inclusive)
 - Una repetición triple del valor 5.3
 - El número -3
 - Una secuencia de nueve valores que comienza en 102 y termina en el número que es la longitud total del vector creado en (c).
 - Confirma que la longitud del vector creado en (d) es 20.
2. (a) Construye y almacena una matriz 4×2 con sus valores ordenados por filas con los valores 4.3, 3.1, 8.2, 8.2, 3.2, 0.9, 1.6 y 6.5, en ese orden.
(b) Confirma que las dimensiones de la matriz en (a) son 3×2 si se elimina una fila.

- (c) Sobrescribe la segunda columna de la matriz en (a) con la misma columna ordenada de menor a mayor.
- (d) ¿Qué devuelve R si se elimina la cuarta fila y la primera columna de (c)? Usa la función `matrix` para asegurarte que el resultado sea una matriz de una sola columna, en lugar de un vector.
- (e) Almacena los cuatro elementos inferiores de (c) como una nueva matriz 2×2 .
- (f) Sobrescribe, en este orden, los elementos de (c) en las posiciones (4,2), (1,2), (4,1) y (1,1) con $-\frac{1}{2}$ de los dos valores en la diagonal de (e).
- (g) Para

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Verifica que $A^{-1}A - I_4$ es una matriz 4×4 de ceros.

3. El gobierno de un país se compone de los partidos políticos: Nacional, Laborista, Verdes, Maoríes y con varios partidos más pequeños etiquetados como Otros. Supongamos que se pregunta a 20 habitantes de ese país con cuál de esos partidos se identificaron más y se obtuvo la siguiente información:

- Hubo 12 hombres y 8 mujeres; los individuos numerados 1, 5 – 7, 12 y 14 – 16 eran mujeres.
- Las personas numeradas 1, 4, 12, 15, 16 y 19 estaban identificadas con los Laboristas; nadie estaba identificado con los maoríes; los individuos numerados 6, 9 y 11 estaban identificados con los Verdes; 10 y 20 estaban identificados con los Otros y el resto identificado con los Nacionales.

- (a) Utiliza tu conocimiento de vectores para crear dos vectores de caracteres: sexo con entradas M (masculino) y F (femenino) y partido con las entradas: Nacional, Laboral, Verdes, Maori y Otro. Asegúrate de que las entradas se coloquen en las posiciones correctas como se describió anteriormente.
- (b) Crea dos factores diferentes basados en el vector sexo y partido. ¿Tiene sentido usar `ordered = TRUE` en cualquier caso? ¿Cómo R organiza los niveles?.
- (c) Use el subconjunto de factores para hacer lo siguiente:
 - Devuelve el factor de los partidos elegidos solo por participantes hombres.
 - Retorna el factor de géneros para quienes eligen al partido Nacional.
- (d) Otras seis personas se unieron a la encuesta, con los resultados

```
> c("Nacional", "Maori", "Maori", "Laborista", "Verdes", "Laboristas")
```

para un partido preferido y

```
> c("M", "M", "F", "F", "F", "M")
```

por su género.

Combina estos resultados con los factores originales en (b).

Supongamos que también se pidió a todas las personas que expresen su confianza en que los Laboristas **ganarán más escaños** en el parlamento que los Nacionales en las próximas elecciones y que **otorgen un porcentaje subjetivo a esa confianza**.

Para esas preguntas se obtuvieron los siguientes 26 resultados:

93, 55, 29, 100, 52, 84, 56, 0, 33, 52, 35, 53, 55, 46, 40, 56, 45, 64, 31, 10, 29, 40, 95, 18, 61.

- Crea un factor con niveles de confianza de la siguiente manera: Bajo para porcentajes [0,30]; Moderado para porcentajes (30,70] y Alto para porcentajes (70,100].

- Del ítem anterior, extrae los niveles correspondientes a aquellos individuos que originalmente dijeron que se identificaban con los Laboralistas. Haz esto también para los del partido Nacional. ¿Qué notaste?
4. (a) Identifica la clase de los siguientes objetos. Para cada objeto, también indica si la clase está explícita o implícitamente definida.

```
> f1 <- array(data=1:36,dim=c(3,3,4))
> bar <- as.vector(f1)
> baz <- as.character(bar)
> qux <- as.factor(baz)
> quux <- bar+c(-0.1,0.1)
```

- (b) Para cada objeto definido en (a), encuentre la suma del resultado de la llamada de `is.numeric` y `is.integer` por separado. Por ejemplo, `is.numeric(f1) + is.integer(f1)` calcularía la suma

```
> f1 <- array(data=1:36,dim=c(3,3,4))
```

Convierte la colección de los cinco resultados en un factor con niveles 0, 1 y 2, identificados por los mismos resultados. Compara este factor con el resultado de coercion en un vector numérico.

- (c) Almacena la matriz

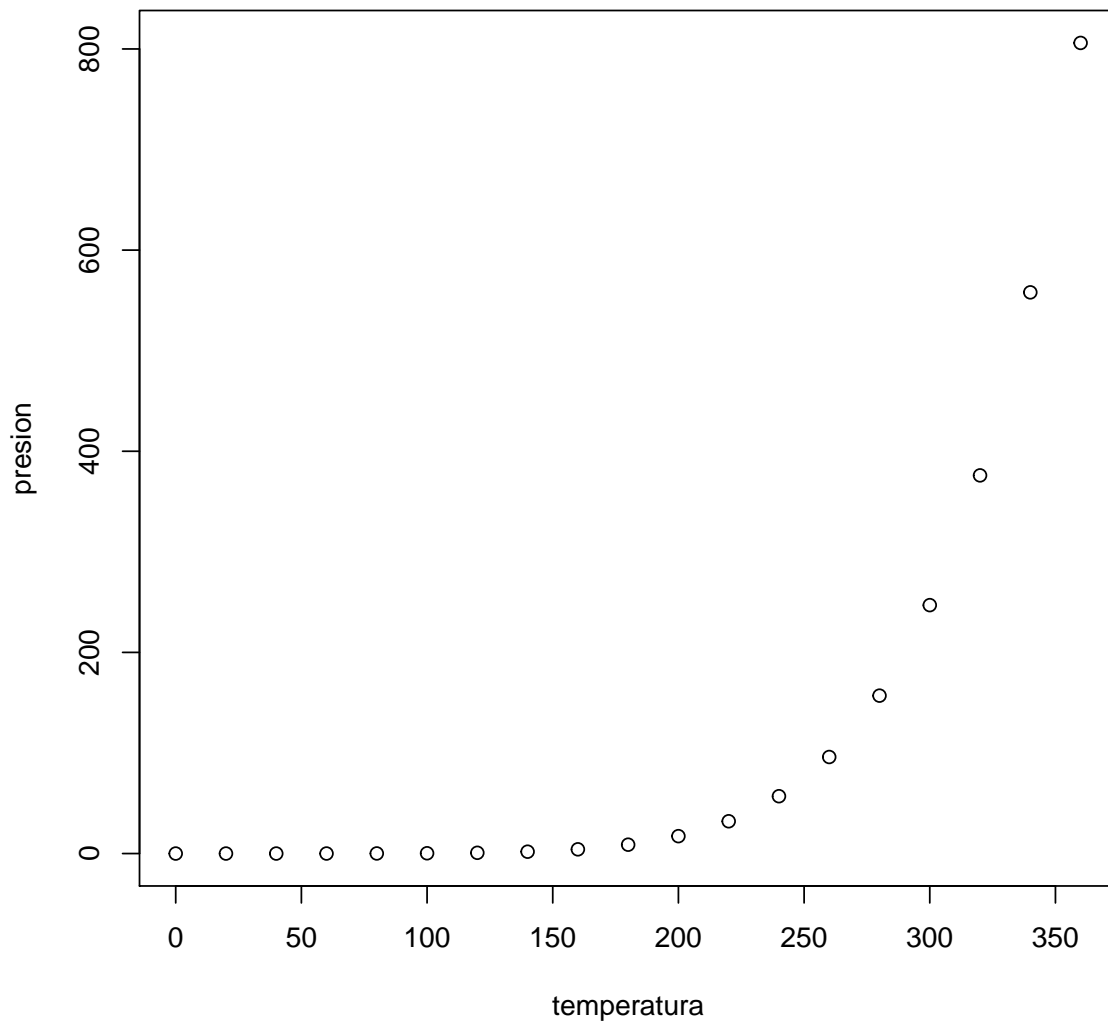
$$A = \begin{bmatrix} 34 & 0 & 1 \\ 23 & 1 & 2 \\ 33 & 1 & \\ 42 & 0 & 1 \\ 41 & 0 & 2 \end{bmatrix}$$

Luego, realiza lo siguiente:

- Realiza una coerción de la matriz a un data frame.
 - Como un data frame, coerce la segunda columna para que tenga un valor lógico.
 - Como un data frame, realiza una coerción a la tercera columna para tener un valor de factor.
5. El siguiente código R produce un diagrama de dispersión simple

```
> plot.new()
> plot.window(range(pressure$temperature),
+             range(pressure$pressure))
> box()
> axis(1)
> axis(2)
> points(pressure$temperature, pressure$pressure)
> mtext("temperatura", side=1, line=3)
> mtext("presion", side=2, line=3)
> mtext("Presion de vapor de Mercurio\ncomo una funcion de la Temperatura",
+       side=3, line=1, font=2)
```

**Presion de vapor de Mercurio
como una funcion de la Temperatura**



- Describe lo que está haciendo cada llamada de función en el código R anterior, incluye una explicación del significado de cada argumento en las llamadas a función. Tu respuesta también debe incluir una explicación de las diferentes regiones y sistemas de coordenadas que se crean con este código.
 - Describe cómo se podría producir la misma gráfica usando ventanas gráficas, diseños, unidades y gráficas primitivas en el sistema de gráficos grid. Tu descripción debe incluir además, una mención de las funciones grid que se requieren y lo que hace cada una de estas funciones.
6. (a) En la biblioteca de conjuntos de datos integrados de R `datasets`, se encuentra el data frame `quakes`. Asegurate de que puedes acceder a este objeto y ver el archivo de ayuda correspondiente para tener una idea de lo que representa esta información. Luego, realiza lo siguiente:
- Selecciona solo los registros que correspondan a una magnitud (`mag`) mayor o igual que 5 y escribe esa información en un archivo de formato de tabla llamado `q5.txt` en una carpeta existente. Usa un caracter delimitador de `!` y no incluir cualquier nombre de fila.
 - Lee el archivo nuevamente en tu espacio de trabajo R, nombrando el objeto como `q5.dframe`.
- (b) En el paquete `car`, hay un data frame llamado `Duncan`, que proporciona datos históricos sobre el prestigio y otras características de 45 ocupaciones de los Estados Unidos en 1950. Instala el

paquete car y accede al conjunto de datos de Duncan y su archivo de ayuda. Luego, realiza lo siguiente:

- Escribe el código R que dibujará education en el eje x e income en el eje y , con los ejes x e y fijados en $[0, 100]$. Escribe etiquetas de ejes apropiados. Para trabajos con un valor de prestige menor o igual a 80, usa \circ como el carácter de punto. Para trabajos con un prestige superior a 80, usa \bullet .
 - Agrega una leyenda que explique la diferencia entre los dos tipos de puntos y luego guarde esa información en un archivo de 500×500 píxeles con formato .png de la imagen.
- (c) Cree una lista llamada exer que contenga los tres conjuntos de datos: quakes, q5.dframe y Duncan. Luego, realiza lo siguiente:
- Escribe el objeto de la lista directamente en el disco, llamándolo Ejercicio_df.txt. Inspecciona brevemente el contenido del archivo en un editor de texto.
 - Lee el archivo Ejercicio_df.txt en tu espacio de trabajo; llama al objeto resultante lista.de.dataframes. Verifica que lista.de.dataframes contenga efectivamente los tres objetos data frames.

7. (a) Almacena el siguiente vector:

```
> f1 <- c(13563, -14156, -14319, 16981, 12921, 11979, 9568, 8833, -12968, 8133)
```

Entonces realiza lo siguiente:

- Muestra la salida todos los elementos de f1 que, cuando se eleva a una potencia de 75, no son infinitos.
 - Devuelve los elementos de f1, excluyendo aquellos que resultan en infinito negativo cuando se eleva a una potencia de 75.
- (b) Almacena la siguiente matriz 3×4 , como el objeto varMatriz

$$\begin{pmatrix} 77875.40 & 27551.45 & 23764.30 & -36478.88 \\ -35466.25 & -73333.85 & 36599.69 & -70585.69 \\ -39803.81 & 55976.34 & 76694.82 & 47032.00 \end{pmatrix}$$

Entonces realiza lo siguiente:

- Identifica los índices específicos de las entradas de varMatriz que son NaN cuando se eleva varMatriz a una potencia de 65 y se divide por infinito.
 - Devuelve los valores en varMatriz que NO son NaN cuando se eleva varMatriz a una potencia de 67 y se añade infinito al resultado. Confirma que esto es idéntico a identificar aquellos valores en varMatriz que, cuando aumentan a una potencia de 67, no son iguales al infinito negativo.
 - Identifique los valores en varMatriz que sean infinito negativo o finito cuando eleva varMatriz a una potencia de 67.
- (c) Considere la siguiente línea de código:

```
> f2 <- c(4.3, 2.2, NULL, 2.4, NaN, 3.3, 3.1, NULL, 3.4, NA)
```

Decide cuál de las siguientes afirmaciones son verdaderas y cuáles son falsas y luego usa R para confirmar:

- La longitud de f2 es 8.
 - Llamando a which(x=is.na(x=f2)), no resultará en 4 y 8.
 - Verificando is.null(x=f2), proporciona la localización de dos valores NULL, presentes.
8. (a) Crea y almacena este data frame como dframe en tu espacio de trabajo:
Las variables Persona, Sexo y Puntuacion deberían ser un vector de caracteres, Sexo debe ser un factor con niveles F y M y Puntuacion debe ser un factor con niveles Alto, Medio y Bajo.
- (b) Stan y Francine tienen 41 años, Steve tiene 15, Hayley tiene 21 y Klaus tiene 60. Roger es extremadamente viejo, 1.600 años. Agrega estos datos como una nueva variable columna numérica en dframe llamada edad.

Persona	Sexo	Puntuacion
Stan	M	Alto
Francine	F	Medio
Steve	M	Bajo
Roger	M	Alto
Hayley	F	Medio
Klaus	M	Medio

- (c) Utiliza tu conocimiento de reordenar las variables de columna en función de las posiciones de índice de columna para sobrescribir dframe, ajustándolo al data frame misdatos.

```
> misdatos <- data.frame(person=c("Peter","Lois","Meg","Chris","Stewie"),
+ edad=c(42,40,17,14,1),
+ sexo= factor(c("M","F","F","M","M")))
```

Es decir, la primera columna debe ser Persona, la segunda columna edad, la tercera columna Sexo y la cuarta columna son Puntuacion.

- (d) Dirige tu atención a misdatos tal como quedó después de que se haya incluido la variable edad.mon.

```
> misdatos$edad.mon <- misdatos$edad*12
```

Crea una nueva versión de misdatos llamada misdatos eliminando la columna edad.mon.

- (e) Ahora combina misdatos2 con dframe, llamándolo misdatosframe.