

Ejercicios 1: probando lo visto en la intro

Leonardo Collado Torres

October 27, 2008

1. Empiezen por hacer el ejercicio simple de tarea que les puse en la clase.
2. Abran R y utilicen los datos 2 5 4 10 8 para:
 - (a) Almacenarlos en un vector de datos x.
 - (b) Encuentren el cuadrado de cada número.
 - (c) Substraigan 3 de cada número.
 - (d) Substraigan 5 de cada número y luego encuentren su raíz.

Hagan esto usando funciones (aprovéchense del reciclaje).

3. Siendo x & y:

```
> x <- c(1, 3, 5, 7, 9)
> y <- c(2, 3, 5, 7, 11, 13)
```

Sin correr los comandos, cual es el resultado de:

- (a) x+1
 - (b) y*2
 - (c) length(x) y length(y)
 - (d) x + y (hay reciclaje)
 - (e) sum(x>5) y sum(x[x>5])
 - (f) sum(x>5 | x<3)
 - (g) y[3]
 - (h) y[-3]
 - (i) y[x] (Que es NA?)
 - (j) y[y>=7]
4. Usen el conjunto de datos `islands` que viene en R. Este contiene el tamaño de los bloques de tierra cuya área excede las 10 mil millas cuadradas. Encuentren las 7 más grandes y las 2 más chicas. Les recomiendo que usen lo visto en índices de vectores y la función `sort`.

5. Qué imprime el siguiente código?

```
> funej <- function(x) {
+   x <- x^2 + x
+   return(x)
+ }
> x <- 5
> funej(x)
> x
```

6. Con los números del 65 al 89, construye una matrix 5x5 llamada M.

- Encuentra la matrix M5 al multiplicar M por el vector 5.
- Multiplica M5 por M para encontrar N. Encuentra "manualmente" el valor de la casilla (3,4) de N y luego verifica tu resultado con R.
- Encuentra el resultado de la ecuación (1):

$$N = M(M^T M)^{-1} M^T N \quad (1)$$

7. Ahora si viene el ejercicio bueno :D.

- Almacenen en un `data frame` el archivo de datos `fagos_grandes_codon.txt`. Usen cualquiera la forma para leer archivos que mas se acomode a los datos o que les guste usar.
- Escriban una función para encontrar la suma de cada columna: tienen que usar el ciclo `for`. Estos resultados los pueden almacenar en un vector; relativícenlos con la suma total de las todas columnas para encontrar la frecuencia relativa de cada columna. NOTA: Si quieren hacer un ejercicio más avanzado, relativicen las columnas por amino ácido codificado; por ejemplo la columna D_GAT tendría el valor de $(D_GAT) / (D_GAT + D_GAC)$.
- Encuentren los nombres de las columnas donde la frecuencia relativa¹ es mayor a 3%². Para los que hagan la versión avanzada, encuentren el codón más frecuente para todos los amino ácidos.

¹La que encontraron en el anterior punto

²Son 6 columnas :)