

# Taller 2 Regresión lineal Multiple

Andrés Felipe Palomino - David Stiven Rojas

2023-04-21

## 1 Introducción

La base de datos "yarn" obtenida de la librería (PLS) contiene información sobre espectros NIR y mediciones de densidad de hilos de PET, consta de 28 individuos (hilos de PET), 268 variables predictoras (NIRS) y una variable de respuesta (densidad). Se ajustará un modelo lineal múltiple para estimar la densidad del hilo PET, mediante mediciones NIR

```
#Importación de librerías necesarias
library(car)
library(MASS)
library(xtable)
library(lmtest)
library(readxl)
library(lmridge)
library(pls)
```

### 1.1 Base de datos

En la siguiente tabla se encuentra un encabezado de la base de datos que se trabajara, esta consta de 30 covariables predictoras, las cuales estarán desde NIR1 hasta NIR30. De primera mano se observa que los valores de los NIR disminuyen a medida que la covariable aumenta

```
library(readxl)
data <- read_excel("C:/Users/david/OneDrive/Escritorio/Octavo semestre github/OctavoSemestre/Estadística")
X <- data
X<- cbind(X[,1:30],X[,colnames(X)=='density'])
xtable(head(X[,1:11]))
```

% latex table generated in R 4.3.0 by xtable 1.8-4 package % Wed Apr 26 23:52:21 2023

	NIR1	NIR2	NIR3	NIR4	NIR5	NIR6	NIR7	NIR8	NIR9	NIR10	NIR11
1	3.07	3.09	3.11	3.10	3.00	2.83	2.62	2.40	2.19	2.01	1.84
2	3.07	3.09	3.10	3.07	2.98	2.84	2.68	2.51	2.35	2.22	2.12
3	3.08	3.10	3.09	3.03	2.88	2.69	2.48	2.27	2.08	1.92	1.77
4	3.08	3.10	3.10	3.07	2.99	2.87	2.74	2.61	2.50	2.42	2.38
5	3.10	3.10	3.08	3.02	2.89	2.72	2.54	2.38	2.24	2.13	2.05
6	3.08	3.08	3.05	2.93	2.73	2.51	2.29	2.10	1.93	1.79	1.67

```
xtable(head(X[,12:21]))
```

% latex table generated in R 4.3.0 by xtable 1.8-4 package % Wed Apr 26 23:52:21 2023

	NIR12	NIR13	NIR14	NIR15	NIR16	NIR17	NIR18	NIR19	NIR20	NIR21
1	1.69	1.58	1.50	1.44	1.34	1.22	1.14	1.12	1.13	1.16
2	2.04	1.98	1.96	1.94	1.89	1.82	1.75	1.71	1.68	1.65
3	1.65	1.55	1.49	1.44	1.35	1.26	1.20	1.18	1.19	1.21
4	2.35	2.35	2.37	2.40	2.40	2.38	2.33	2.28	2.21	2.11
5	1.99	1.95	1.94	1.93	1.90	1.85	1.80	1.76	1.73	1.68
6	1.56	1.48	1.43	1.39	1.32	1.25	1.20	1.19	1.19	1.19

```
xtable(head(X[,22:31]))
```

% latex table generated in R 4.3.0 by xtable 1.8-4 package % Wed Apr 26 23:52:21 2023

	NIR22	NIR23	NIR24	NIR25	NIR26	NIR27	NIR28	NIR29	NIR30	density
1	1.16	1.15	1.15	1.13	1.07	1.02	1.01	1.03	1.08	100.00
2	1.58	1.51	1.45	1.38	1.29	1.20	1.15	1.13	1.14	80.22
3	1.20	1.18	1.17	1.15	1.10	1.07	1.06	1.08	1.12	79.49
4	1.98	1.85	1.75	1.63	1.51	1.40	1.30	1.23	1.20	60.80
5	1.60	1.52	1.46	1.39	1.31	1.24	1.19	1.16	1.17	59.97
6	1.18	1.15	1.14	1.12	1.09	1.06	1.06	1.07	1.11	60.48

$2^{30}$

[1] 1073741824

## 1.2 Selección de variables

En el proceso de selección de variables se procede a ajustar todos los posibles modelos  $2^{30}$ , del cual se observa el  $R_{adj}^2$ , el AIC y el BIC. Además se realizan los algoritmos de selección (forward selection, backward selection, stepwise selection) para identificar si hay diferencias en modelos.