Tarea3-Arturo González Bencomo

Arturo González Bencomo

18 de julio de 2017

Esta parte del proyecto será sobre regresión lineal. Supongamos que quieren explicar una variable estadistica Y (por ejemplo altura) utilizando la información de p variables X1, …, Xp (peso, ancho de huesos, etc). Si se toma una muestra de N individuos, cada variable está representada por un vector de tamaño N. La información de las variables explicativas se pueden juntar en una matriz

$$X = [X^1 \mid \ldots \mid X^p]$$

de tamaño n x p donde cada columna es una variable y cada fila uno de los individuos de la muestra. Tienen que contestar lo siguiente:

- Plantear el problema de regresión lineal como un problema de mínimos cuadrados, encontrar el vector beta que resuelva

$$\hat{\beta} = argmin_{\beta \in R^p} \mid\mid Y - X\beta \mid\mid^2$$

y encontrar la solución teórica. ¿Por qué este planteamiento nos da un ajuste lineal a nuestros datos?

Nos da un ajuste lineal debido a que se puede describir una variable ej 'Y' en términos de combinaciones lineales de los vectores columnas multiplicados con sus respectivos coeficientes numéricos (escalares) en las coordenadas del vector de coeficientes $\beta = [\beta_0, \ldots \beta_p]$, en el caso de una regresión lineal univariable se describe una variable dependiente en términos de una independiente de la siguiente manera:

$$Y = \beta_0 + \beta 1X$$

En regresión lineal multivariada se puede describir una variable mediante una combinación lineal de varias variables de la siguiente manera:

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p + \epsilon$$

¿Podríamos usarlo para ajustar polinomios (ej $y = x^2$)?

Del desarrollo de la ecuación

$$\mid\mid Y - X\beta \mid\mid^2$$

, se llega a la forma matricial para obtener el vector β de mínimos cuadrados lo cual es válido.

$$\bigtriangledown \mid\mid Y - X\beta \mid\mid^2 = \bigtriangledown(Y^T Y + \beta^T X^T X\beta - 2\beta^T X^T Y)$$

Derivamos con respecto a Beta para obtener el gradiente e igualamos a cero.

$$0 = 2X^T X\beta - 2X^T Y$$

Despejamos β de la ecuación

$$\beta_{OLS} = (X^T X)^{-1} X^T Y$$

Sin embargo para poder realizar una aproximación a un polinomio lo único distinto es que los vectores columna deben ser elevadas a la potencia a la cual se quiera aproximar, una vez formada esta matriz X, se puede proceder de la misma forma para encontrar el vector de β.

$$Y_{m,n} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}, X = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \dots & & \\ 1 & x_n & x_n^2 \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \dots \\ \epsilon_n \end{pmatrix}$$

De manera que se puede formar un polinomio de grado n con sus respectivos coeficientes

$$Y = \beta_0 + \beta_1 X_i^1 + \ldots + \beta_p X_i^p + \epsilon_i$$

- Argumentar la relación entre la solución encontrada y un problema de proyección en subespacios vectoriales de álgebra lineal ¿Cuál es la relación particular con el teorema de Pitágoras?

En la relación lineal entre variables, existen varianzas en los valores predecidos por la regresión lineal y la varianza del error, de manera que la varianza de la variable dependiente original es igual a la suma de la varianza de la variable predicida + la varianza del error, y la varianza al ser una medida cuadrática se asemeja al teorema de Pitágoras

$$var(y) = \sigma^2 = E[(Y - \mu)^2]$$

$$\sigma_y^2 = \sigma_{ypred}^2 + \sigma_\epsilon^2$$

$$Var(y) = var(ypredecida) + var(\epsilon)$$

Ejemplo: Sea x = {1,2,3,4,5,6,7,8,9,10}

$$Var(x) = 8.25$$

Sea y = {1.5, 2.3, 3,4, 5.6, 7.1, 7.8, 8.3,9,10}

$$Var(y) = 8.387$$

y predecida = [1,2,3,4,5,6,7,8,9,10]

$$Var(Ypredecida) = 8.25$$

Y recordando que el Error es la diferencia entre el valor estimado y el valor real, se puede obtener un vector de errores y calculares la varianza.

$$\epsilon^2 = [0.25, 0.09, 0, 0, 0.36, 1.21, 0.64, 0.09, 0, 0]$$

$$Var(\epsilon) = 0.1385$$

$$Var(Ypredecida) + Var(\epsilon) = Var(Yoriginal)$$

$$8.25 + 0.1385 = 8.3885 \sim 8.387$$

- ¿Que logramos al agregar una columna de unos en la matriz X? es decir, definir mejor

$$X = [1_n \mid X^1 \mid \ldots \mid X^p]$$

con

$$1_n = [1, 1, \ldots, 1]^T$$

Mediante la incorporación del vector columna 1_n a la matrix X, se obtiene un parámetro β_0 al realizar el ajuste a mínimos cuadrados. Este parámetro incorpora un 'offset' o desplazamiento constante en el valor numérico en la variable a predecir ya que no necesariamente la variable a predecir tiene un valor 0 cuando las demás variables estén en cero.

- Plantear el problema de regresión ahora como un problema de estadística

$$Y = \beta_0 + \beta 1 X_i^1 + \ldots + \beta_p X_i^p + \epsilon_i$$

donde los errores son no correlacionados con distribución

$$\epsilon \sim N(0, \sigma^2)$$

- ¿Cual es la función de verosimilitud del problema anteriror? Hint: empiecen por escribir el problema como

Sea

$$Y = X\beta + \epsilon$$

con

$$\epsilon \sim N(0, \sigma^2 I_n)$$

con I_n la matriz identidad. Y concluyan entonces que

$$Y \sim N(X\beta, \sigma^2 I_n)$$

Escriban entonces la verosimilitud como

$$L(\beta, \sigma^2) = f(Y|\beta, \sigma^2, X)$$

$$L(\beta, \sigma^2; Y, X) = \prod_{i=1}^{p} f_y(y_i|X; \beta, \sigma^2)$$

$$= \prod_{i=1}^{p} \frac{1}{\sqrt{2\pi\sigma^2}} e^{(-\frac{1}{2}\frac{(y_i - x_i\beta)^2}{\sigma^2})}$$

$$= \frac{1}{(2\pi\sigma^2)^{N/2}} e^{(-\frac{1}{2\sigma^2}\sum_{i=1}^{p}(y_i - x_i\beta)^2)}$$

En representación matricial:

$$L(\beta, \sigma^2; Y, X) = (2\pi\sigma^2)^{-N/2} e^{(-\frac{1}{2\sigma^2}(Y - X\beta)^2)}$$

$X\beta$ corresponde a la media de la distribución normal de Y, mientras que en término σ^2 en la formula se refiere a la varianza de la distribución.

- Mostrar que la solución de máxima verosimilitud es la misma que la del problema de mínimos cuadrados.

Esto se puede mediante plantear la función log de verosimilitud, posteriormente derivar respecto a una beta, igualar a cero, despejar β y comprobar que esta ecuación tenga la misma forma de $(X^TX)^{-1}X^Ty$

$$l(\beta, \sigma^2; Y, X) = -\frac{N}{2}ln(2\pi) - \frac{N}{2}ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - x_i\beta)^2$$

$$\bigtriangledown_\beta l(\beta, \sigma^2; Y, X)$$

$$\bigtriangledown_\beta(-\frac{N}{2}ln(2\pi) - \frac{N}{2}ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - x_i\beta)^2)$$

Al obtener el gradiente de β, se deriva cada uno de los términos de la suma (derivada de constantes = 0 y queda solamente el término de la suma, y hubo que hacer una derivada de una sumatoria donde se aplicó regla de la cadena) y se obtiene la siguiente expresión:

$$= \frac{1}{\sigma^2}\sum_{i=1}^{N} x_i^T(y_i - x_i\beta)$$

$$= \frac{1}{\sigma^2}(\sum_{i=1}^{N} x_i^T y_i - \sum_{i=1}^{N} x_i^T x_i\beta)$$

Que es igual a cero solo si

$$\sum_{i=1}^{N} x_i^T y_i - \sum_{i=1}^{N} x_i^T x_i\beta = 0$$

Despejamos β de la ecuación y podemos ver que corresponde con

$$\beta = \frac{\sum_{i=1}^{N} x_i^T y_i}{(\sum_{i=1}^{N} x_i^T x_i)} = (X^TX)^{-1}X^Ty$$

- Investiga el contenido del Teorema de Gauss-Markov sobre mínimos cuadrados.

Este teorema afirma que de todos los estimadores lineales insesgados de β, $ _{OLS} (Ordinary least squares) $ es el que proporciona varianza mínima comparado con otros estimadores lineales (BLUE). Esto implica reducir al mínimo los errores RSS (residual sum of squares) (real - predicción). En términos de literatura, BLUE es un acrónimo para Best Linear Unbiased Estimator (Mejor estimador lineal insesgado) Este teorema está sujeto a algunas condiciones.

Las condiciones son:

- Correcta especificación: La variable a estimar ha de ser una combinación lineal de los parámetrosβ multiplicado por la matriz de proyección `H\hat`
- Valor estimado de los errores igual a cero: $E(u_i|X_i') = 0$
- Los distintos términos de error son no correlacionados $Cov(\epsilon_i, \epsilon_j) = 0$
- La varianza de los errores son finitas (homocedasticidad) $Var(\epsilon_i) = \sigma^2 < \infty$

β se obtiene de la siguiente manera:

$\beta = (X^T X)^{-1} X^T Y$

Parte aplicada

Para esta parte pueden usar la base de datos diamonds que sugirieron, aunque hay puntos adicionales si usan alguna base original interesante.

Cargar la base que se encuentra en el paquete ggplot2. Los comandos que pueden usar para cargar la base diamonds a su ambiente de trabajo en R son:

```
#install.packages("ggplot2")
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.3.3
```

```
data(diamonds)
head(diamonds)
```

```
## Source: local data frame [6 x 10]
##
##   carat       cut  color clarity depth table price     x     y     z
##   <dbl>    <fctr> <fctr>  <fctr> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.23     Ideal      E     SI2  61.5    55   326  3.95  3.98  2.43
## 2  0.21   Premium      E     SI1  59.8    61   326  3.89  3.84  2.31
## 3  0.23      Good      E     VS1  56.9    65   327  4.05  4.07  2.31
## 4  0.29   Premium      I     VS2  62.4    58   334  4.20  4.23  2.63
## 5  0.31      Good      J     SI2  63.3    58   335  4.34  4.35  2.75
## 6  0.24 Very Good      J    VVS2  62.8    57   336  3.94  3.96  2.48
```

Posteriormente deben hacer una regresión lineal. Su objetivo es explicar la variable price usando las demás variables. Noten que algunas variables no son numéricas, por lo que no pueden incluirse en un análisis crudo de regresión lineal. Para este proyecto NO es necesario saber transformar las variables no numéricas para poder usarlas en la regresión; hacerlo es

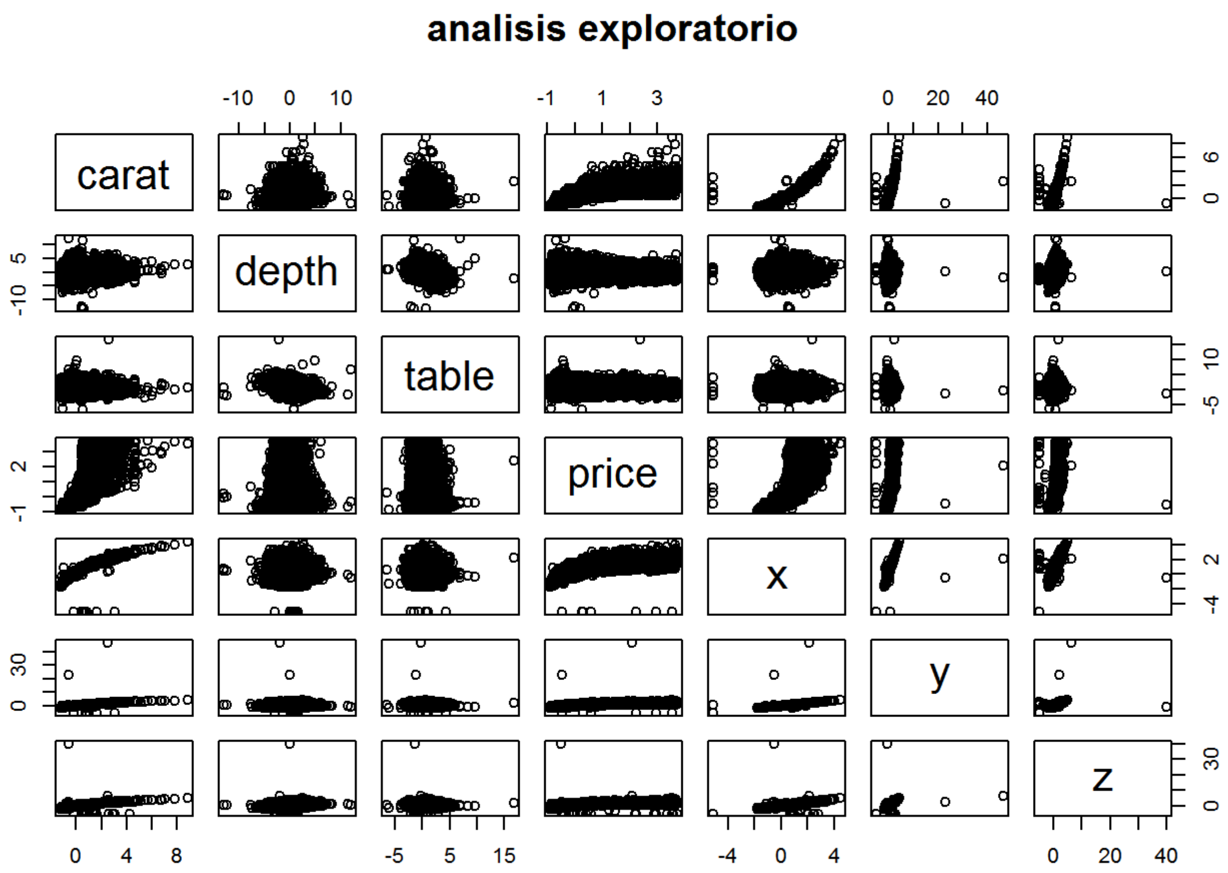
optativo, de hecho, las paqueterías lo hacen por ustedes pero deben ser cuidadosos. Pueden usar la función lm de R para su análisis de regresión.

Creamos un nuevo data_frame que solo contiene los datos que vamos a utilizar, es decir, las variables numéricas centradas mediante la funcion scale

```
diamantes_num = diamonds[ ,c(1,5,6,7,8,9,10)]
diamantes_num_esc = scale(diamantes_num)
diamantes_num_esc <- as.data.frame(diamantes_num_esc)
```

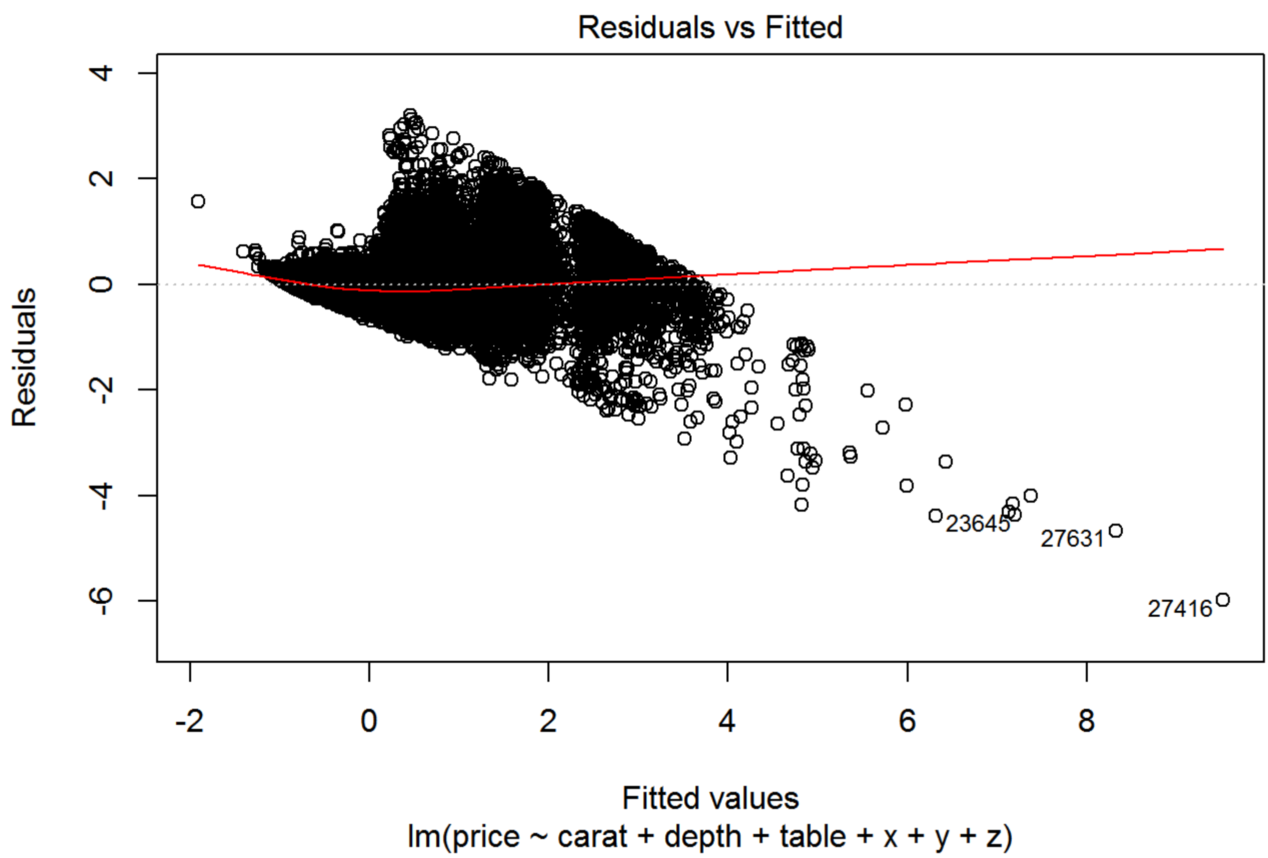
Utilizamos función pairs para tener una apreciación visual de las variables del dataset o análisis exploratorio.

```
pairs(diamantes_num_esc, main="analisis exploratorio")
```



Aplicamos la regresión lineal con la función 'lm' para pronosticar el precio del diamante en base a sus demás dimensiones numéricas.

```
predicted_price = lm(price ~ carat + depth + table + x + y + z, data = diamantes_num_esc)
plot(predicted_price, which=1)
```




```
summary(predicted_price)
```

```
##
## Call:
## lm(formula = price ~ carat + depth + table + x + y + z, data = diamantes_num_esc)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.9853 -0.1542 -0.0127  0.0872  3.1982
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.359e-14  1.616e-03   0.000  1.00000
## carat        1.270e+00  7.509e-03 169.085  < 2e-16 ***
## depth       -7.295e-02  1.976e-03 -36.910  < 2e-16 ***
## table       -5.738e-02  1.727e-03 -33.216  < 2e-16 ***
## x           -3.699e-01  1.211e-02 -30.547  < 2e-16 ***
## y            1.899e-02  7.307e-03   2.599  0.00937 **
## z            7.364e-03  7.837e-03   0.940  0.34744
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3752 on 53933 degrees of freedom
## Multiple R-squared:  0.8592, Adjusted R-squared:  0.8592
## F-statistic: 5.486e+04 on 6 and 53933 DF,  p-value: < 2.2e-16
```

- ¿Qué tan bueno fue el ajuste?

Del resumen y la gráfica del ajuste a la regresión lineal podemos observar el error estándar residual que es un valor relativamente cercano al cero (0.3752) por lo que se puede decir que es un buen modelo, en un modelo perfecto, el error debería ser cero. Adicionalmente en el resumen aparecen los coeficientes para cada una de las variables así como errores, valores t, y valores de probabilidad > t. Este último parámetro indica la relevancia de la variable para el modelo lineal. De aquí podemos observar que carat, depth, table y x son las más relevantes (tip: también se puede observar con '*' los grados de relevancia de las variables). Así mismo se puede observar el valor de $R^2 = 0.8592$ que equivale al coseno del ángulo entre el valor proyectado mediante mínimos cuadrados y el valor real del precio. A medida que R^2 se acerca a uno, el ajuste a mínimos cuadrados se acerca poder encontrar la solución exacta del sistema.

- ¿Qué medida puede ayudarnos a saber la calidad del ajuste? ¿Cúal fue el valor de $ ^{2} $ que ajustó su modelo y que relación tienen con la calidad del ajuste?

Dentro del resumen de la regresión se puede observar el valor de $R^2 = 0.8592$ que equivale al coseno del ángulo entre valor proyectado mediante mínimos cuadrados y el valor real del precio. A medida que R^2 se acerca a uno es decir $\angle 0$, el ajuste a mínimos cuadrados se acerca a poder encontrar la solución exacta del sistema. Este 0.8592 se podría interpretar como una medida de porcentaje es decir que el modelo explica en un 85.9% el precio en función de las demás variables numéricas.

- ¿Cual es el angulo entre Y y $\hat{Y}$ estimada? Hint: usen la R^2 cuadrada y el arcocoseno

```
angulo_radianes <- acos(sqrt(0.8592))
angulo_grados = angulo_radianes * 180/pi
angulo_grados
```

```
## [1] 22.03873
```

- Definan una función que calcule la logverosimilitud de unos parámetros β y σ^2.

```
head(diamonds)
```

```
## Source: local data frame [6 x 10]
##
##   carat       cut  color clarity depth table price     x     y     z
##   <dbl>    <fctr> <fctr>  <fctr> <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.23     Ideal      E     SI2  61.5    55   326  3.95  3.98  2.43
## 2  0.21   Premium      E     SI1  59.8    61   326  3.89  3.84  2.31
## 3  0.23      Good      E     VS1  56.9    65   327  4.05  4.07  2.31
## 4  0.29   Premium      I     VS2  62.4    58   334  4.20  4.23  2.63
## 5  0.31      Good      J     SI2  63.3    58   335  4.34  4.35  2.75
## 6  0.24 Very Good      J    VVS2  62.8    57   336  3.94  3.96  2.48
```

```r
diamonds_m <- data.matrix(diamonds[ ,c(5,6,7,8,9,10)])

n <- length(diamonds_m)
sigma_sq <- 0.8563
mod <- lm(formula = price ~ carat + x + y + z + depth, data=diamonds)
summary(mod)$coefficients[,1]
```

```
## (Intercept)       carat           x           y           z       depth
## 12196.68697 10615.49551 -1369.67016    97.59636    64.19545  -156.62430
```

```r
Beta_1 <- c(12196.7, 10615.5, -1369.7, 97.6, 64.2, -156.6)

options(max.print=1000)

funcionVerosimilitud <- function(bet, sig){
  -(n/2)*(log(2*pi))-((n/2)*log(sig))-((1/(2*sig))*((diamonds$price-(diamonds_m*bet))*(diamon
ds$price-(diamonds_m*bet))))
}

funcionVerosimilitud(Beta_1,sigma_sq)
```

```
##              depth         table         price             x
##     [1,] -3.282477e+11 -2.625018e+11 -9.229810e+12 -1337253996.7
##     [2,] -2.350616e+11 -2.445951e+11 -6.991629e+12  -980303362.5
##     [3,] -3.576756e+09 -4.662635e+09 -1.173074e+11   -20421328.8
##     [4,] -1.961967e+07 -1.684056e+07 -6.081151e+08     -275667.0
##     [5,] -8.391183e+06 -6.977084e+06 -2.620109e+08     -274156.9
##     [6,] -6.067091e+07 -5.036476e+07 -1.637598e+09     -802615.9
##     [7,] -3.368380e+11 -2.819415e+11 -9.804740e+12 -1336695244.4
##     [8,] -2.518611e+11 -1.988151e+11 -7.471417e+12 -1073303196.9
##     [9,] -4.677985e+09 -4.109413e+09 -1.245919e+11   -18831276.7
##    [10,] -1.767595e+07 -1.868580e+07 -6.227613e+08     -273904.7
##    [11,] -8.570440e+06 -6.221656e+06 -2.682987e+08     -274856.5
##    [12,] -6.071843e+07 -4.872776e+07 -1.676814e+09     -805328.2
##    [13,] -3.165918e+11 -3.229159e+11 -1.015804e+13 -1289092094.5
##    [14,] -2.543042e+11 -1.916424e+11 -7.785026e+12 -1226888468.2
##    [15,] -4.003540e+09 -4.245489e+09 -1.305774e+11   -18168541.5
##    [16,] -1.857606e+07 -1.677220e+07 -6.488119e+08     -276274.5
##    [17,] -7.976570e+06 -5.951919e+06 -2.827191e+08     -275269.6
##    [18,] -6.197199e+07 -4.556618e+07 -1.787051e+09     -871983.8
##    [19,] -3.532471e+11 -2.721189e+11 -1.069970e+13 -1533405714.2
##    [20,] -2.584058e+11 -2.287929e+11 -8.105082e+12 -1148268455.5
##    [21,] -4.425256e+09 -3.467135e+09 -1.351587e+11   -22615890.7
##    [22,] -2.042541e+07 -1.496356e+07 -6.753964e+08     -272631.0
##    [23,] -7.685821e+06 -6.655742e+06 -2.908937e+08     -278146.6
##    [24,] -5.470404e+07 -5.939169e+07 -1.807471e+09     -904431.6
##    [25,] -2.929197e+11 -3.335853e+11 -1.082199e+13 -1690379416.3
##    [26,] -2.397837e+11 -2.210965e+11 -8.244222e+12 -1019988076.0
##    [27,] -4.314961e+09 -3.591850e+09 -1.382568e+11   -19865596.7
##    [28,] -1.933489e+07 -1.609883e+07 -6.947123e+08     -274454.8
##    [29,] -7.536364e+06 -7.669186e+06 -2.975174e+08     -278468.2
##    [30,] -5.743104e+07 -5.059217e+07 -1.848659e+09     -829813.2
##    [31,] -3.123592e+11 -2.818879e+11 -1.403491e+13 -1367251672.6
##    [32,] -2.350053e+11 -2.135000e+11 -1.063152e+13 -1054192035.5
##    [33,] -4.075595e+09 -3.851588e+09 -1.772889e+11   -20184832.8
##    [34,] -1.733185e+07 -1.642025e+07 -8.808150e+08     -272367.3
##    [35,] -7.722401e+06 -6.714568e+06 -3.771759e+08     -285498.2
##    [36,] -5.314895e+07 -5.455035e+07 -2.344008e+09     -901181.8
##    [37,] -3.565319e+11 -3.020284e+11 -1.403491e+13 -1252603352.2
##    [38,] -2.691976e+11 -1.916036e+11 -1.063152e+13 -1189974351.4
##    [39,] -4.089067e+09 -3.851682e+09 -1.781721e+11   -21714503.5
##    [40,] -1.877168e+07 -1.466634e+07 -8.852013e+08     -273025.9
##    [41,] -7.531999e+06 -6.222401e+06 -3.790534e+08     -279989.0
##    [42,] -5.832790e+07 -4.940029e+07 -2.355683e+09     -986731.2
##    [43,] -3.688790e+11 -2.720775e+11 -1.410482e+13 -1360312313.4
##    [44,] -2.241230e+11 -2.608453e+11 -1.068448e+13 -1134622954.0
##    [45,] -4.402722e+09 -3.471816e+09 -1.781721e+11   -23798361.7
##    [46,] -1.915861e+07 -1.641411e+07 -8.852013e+08     -272369.8
##    [47,] -7.691482e+06 -5.985477e+06 -3.790534e+08     -281240.5
##    [48,] -6.335586e+07 -4.940029e+07 -2.355683e+09     -960675.5
##    [49,] -3.476817e+11 -3.123575e+11 -1.417491e+13 -1367138665.9
##    [50,] -2.421353e+11 -2.365794e+11 -1.073757e+13 -1212356384.1
##    [51,] -4.102561e+09 -4.115979e+09 -1.790574e+11   -20668420.2
##    [52,] -1.882931e+07 -1.410031e+07 -8.895984e+08     -272545.2
##    [53,] -7.448813e+06 -5.981824e+06 -3.809355e+08     -280474.5
##    [54,] -5.925499e+07 -5.282361e+07 -2.367387e+09     -879129.5
##    [55,] -3.051077e+11 -3.335403e+11 -1.417491e+13 -1305819655.3
##    [56,] -2.445354e+11 -2.287535e+11 -1.079079e+13 -1195418216.8
```

```
##    [57,] -3.890944e+09 -4.116077e+09 -1.799449e+11   -24736136.8
##    [58,] -1.927515e+07 -1.580834e+07 -8.940065e+08     -272357.5
##    [59,] -8.007142e+06 -6.456536e+06 -3.828224e+08     -282203.5
##    [60,] -6.224455e+07 -4.775818e+07 -2.379119e+09     -941504.5
##    [61,] -3.216755e+11 -2.817662e+11 -2.646279e+13 -1755115238.8
##    [62,] -2.574259e+11 -2.286460e+11 -2.004570e+13 -1148854175.7
##    [63,] -4.334760e+09 -3.609915e+09 -3.342777e+11   -24398980.8
##    [64,] -1.799310e+07 -1.436985e+07 -1.660534e+09     -282524.8
##    [65,] -7.007589e+06 -6.386048e+06 -7.109233e+08     -307385.9
##    [66,] -6.064431e+07 -5.102021e+07 -4.435406e+09    -1126005.4
##    [67,] -3.334088e+11 -2.651960e+11 -2.655876e+13 -1639881356.6
##    [68,] -2.632193e+11 -2.133934e+11 -2.011840e+13 -1204441511.2
##    [69,] -4.238926e+09 -3.736869e+09 -3.354899e+11   -25027445.8
##    [70,] -1.711866e+07 -1.550623e+07 -1.666555e+09     -287554.4
##    [71,] -7.003623e+06 -6.627201e+06 -7.135005e+08     -324198.6
##    [72,] -6.385914e+07 -5.103110e+07 -4.451461e+09    -1149692.5
##    [73,] -3.431649e+11 -3.019008e+11 -2.665490e+13 -1557847389.2
##    [74,] -2.566019e+11 -2.133927e+11 -2.019123e+13 -1181975403.3
##    [75,] -4.501921e+09 -3.610099e+09 -3.367044e+11   -24311254.5
##    [76,] -1.620675e+07 -1.668664e+07 -1.672586e+09     -284587.5
##    [77,] -6.598647e+06 -6.138449e+06 -7.160823e+08     -320588.5
##    [78,] -6.253026e+07 -4.767818e+07 -4.451461e+09    -1103246.9
##    [79,] -3.420739e+11 -3.122294e+11 -2.665490e+13 -1365555199.7
##    [80,] -2.574244e+11 -2.286446e+11 -2.019123e+13 -1057186390.3
##    [81,] -4.459885e+09 -3.866017e+09 -3.367044e+11   -21523425.9
##    [82,] -1.798024e+07 -1.668664e+07 -1.672586e+09     -287375.3
##    [83,] -7.360653e+06 -6.138449e+06 -7.160823e+08     -323432.0
##    [84,] -6.102871e+07 -5.103110e+07 -4.451461e+09    -1232251.9
##    [85,] -2.907399e+11 -3.122294e+11 -2.665490e+13 -1514021333.0
##    [86,] -2.549609e+11 -2.286446e+11 -2.019123e+13  -999965603.2
##    [87,] -4.157649e+09 -3.736963e+09 -3.367044e+11   -21620033.5
##    [88,] -1.717366e+07 -1.608789e+07 -1.672586e+09     -287559.0
##    [89,] -6.797613e+06 -6.138449e+06 -7.160823e+08     -322328.7
##    [90,] -6.347803e+07 -5.449857e+07 -4.451461e+09    -1163194.6
##    [91,] -3.368542e+11 -2.799803e+11 -6.601324e+14 -2603014953.7
##    [92,] -1.978905e+11 -3.109195e+11 -5.000537e+14 -2521698836.1
##    [93,] -4.433139e+09 -3.687020e+09 -8.338778e+12   -65438845.4
##    [94,] -6.752016e+06 -4.863880e+06 -4.147674e+10    -3110326.8
##    [95,] -1.316014e+06 -6.805873e+05 -1.775368e+10    -3596478.9
##    [96,] -8.107338e+07 -8.215262e+07 -1.103979e+11    -8158809.4
##    [97,] -3.041501e+11 -3.314650e+11 -6.610905e+14 -2612376971.3
##    [98,] -2.869728e+11 -2.508188e+11 -5.007795e+14 -2377046566.3
##    [99,] -4.433444e+09 -4.078475e+09 -8.356936e+12   -66664438.0
##   [100,] -6.411013e+06 -5.185673e+06 -4.150681e+10    -3040465.1
##   [101,] -1.255222e+06 -6.796114e+05 -1.776655e+10    -3601055.8
##   [102,] -8.633318e+07 -7.373232e+07 -1.104780e+11    -8264266.9
##   [103,] -3.271853e+11 -2.605995e+11 -6.615698e+14 -2698740902.6
##   [104,] -2.483864e+11 -2.193709e+11 -5.011426e+14 -2056394985.9
##   [105,] -4.616483e+09 -3.687298e+09 -8.356936e+12   -69595674.2
##   [106,] -6.672434e+06 -4.243844e+06 -4.150681e+10    -3065337.7
##   [107,] -1.235864e+06 -8.891254e+05 -1.776655e+10    -3583178.8
##   [108,] -9.175881e+07 -7.581118e+07 -1.105580e+11    -8113689.5
##   [109,] -2.980117e+11 -2.799771e+11 -6.620493e+14 -3037914678.2
##   [110,] -2.508173e+11 -1.971668e+11 -5.015058e+14 -1725114584.0
##   [111,] -4.405715e+09 -3.815594e+09 -8.362993e+12   -69957445.4
##   [112,] -6.668568e+06 -4.543504e+06 -4.153689e+10    -3085377.3
##   [113,] -1.234364e+06 -8.133494e+05 -1.777942e+10    -3539685.9
##   [114,] -9.154464e+07 -8.437460e+07 -1.106381e+11    -8091178.1
```

```
## [115,] -3.379325e+11 -2.702002e+11 -6.625290e+14 -2669612617.1
## [116,] -2.565362e+11 -1.830218e+11 -5.018691e+14 -2019639989.4
## [117,] -4.447597e+09 -3.946084e+09 -8.369052e+12   -70320158.0
## [118,] -6.740351e+06 -3.946289e+06 -4.156698e+10    -3107999.8
## [119,] -1.024426e+06 -8.122258e+05 -1.779230e+10    -3619187.8
## [120,] -8.636153e+07 -8.437460e+07 -1.106381e+11    -8279681.1
## [121,] -3.346896e+11 -2.702002e+11 -6.625290e+14 -2631229817.3
## [122,] -2.532599e+11 -1.900281e+11 -5.018691e+14 -2019639989.4
## [123,] -4.447597e+09 -4.630924e+09 -8.369052e+12   -64490118.9
## [124,] -7.561512e+06 -4.854061e+06 -4.156698e+10    -3145801.7
## [125,] -1.466768e+06 -6.176015e+05 -1.779230e+10    -3637161.6
## [126,] -9.017792e+07 -8.658374e+07 -1.106381e+11    -8037732.5
## [127,] -3.250538e+11 -2.701994e+11 -6.630088e+14 -2986597976.5
## [128,] -2.316949e+11 -2.118361e+11 -5.022326e+14 -1848395233.9
## [129,] -4.832601e+09 -3.815782e+09 -8.375113e+12   -72629222.4
## [130,] -8.048511e+06 -5.858340e+06 -4.159709e+10    -3033207.9
## [131,] -1.137689e+06 -6.766905e+05 -1.780519e+10    -3593314.3
## [132,] -9.480346e+07 -8.222178e+07 -1.107984e+11    -8039580.9
## [133,] -3.303883e+11 -3.000478e+11 -6.634888e+14 -2592969337.6
## [134,] -2.557128e+11 -2.348276e+11 -5.029600e+14 -1890345729.3
## [135,] -4.392237e+09 -4.213745e+09 -8.387242e+12   -68617047.6
## [136,] -6.172701e+06 -5.851118e+06 -4.165733e+10    -3166252.2
## [137,] -9.704021e+05 -1.134848e+06 -1.783098e+10    -3625783.8
## [138,] -9.067593e+07 -8.441665e+07 -1.108786e+11    -8077235.9
## [139,] -3.229301e+11 -2.799738e+11 -6.639690e+14 -2776356962.7
## [140,] -2.500016e+11 -2.044329e+11 -5.033238e+14 -1827335222.9
## [141,] -4.337022e+09 -3.816066e+09 -8.393310e+12   -66909943.6
## [142,] -5.777292e+06 -4.527726e+06 -4.168747e+10    -3075686.3
## [143,] -1.113913e+06 -6.131259e+05 -1.785678e+10    -3642181.5
## [144,] -8.665474e+07 -8.011052e+07 -1.110391e+11    -8159539.0
## [145,] -3.271792e+11 -2.605940e+11 -6.649299e+14 -2640406812.7
## [146,] -2.516258e+11 -2.673256e+11 -5.036878e+14 -1890212848.1
## [147,] -4.674383e+09 -4.079255e+09 -8.405452e+12   -62565278.5
## [148,] -5.910910e+06 -4.521423e+06 -4.174777e+10    -3095824.6
## [149,] -1.050736e+06 -8.055090e+05 -1.786969e+10    -3648592.3
## [150,] -8.868264e+07 -8.012417e+07 -1.111194e+11    -8116839.6
## [151,] -3.282454e+11 -3.000445e+11 -6.654106e+14 -2611674043.7
## [152,] -2.451560e+11 -2.044308e+11 -5.044162e+14 -2033764625.6
## [153,] -4.560801e+09 -4.214241e+09 -8.417603e+12   -64756117.0
## [154,] -6.336419e+06 -5.486908e+06 -4.180812e+10    -3100963.0
## [155,] -1.211388e+06 -7.341911e+05 -1.789552e+10    -3630795.2
## [156,] -9.211964e+07 -8.015149e+07 -1.112800e+11    -8158972.4
## [157,] -3.176512e+11 -2.701939e+11 -6.663725e+14 -2668980979.1
## [158,] -2.403593e+11 -2.193638e+11 -5.047806e+14 -2070438751.7
## [159,] -4.420534e+09 -3.946852e+09 -8.417603e+12   -67130907.3
## [160,] -5.797966e+06 -4.827928e+06 -4.180812e+10    -3078430.0
## [161,] -8.925389e+05 -8.771810e+05 -1.789552e+10    -3607460.9
## [162,] -8.738039e+07 -7.804028e+07 -1.113604e+11    -8183446.6
## [163,] -3.282428e+11 -3.103391e+11 -6.668537e+14 -2688195099.8
## [164,] -2.499978e+11 -2.118304e+11 -5.051452e+14 -1968319513.4
## [165,] -4.476480e+09 -3.562291e+09 -8.423682e+12   -70260092.3
## [166,] -5.967373e+06 -5.479930e+06 -4.186852e+10    -3106106.0
##                    y            z
##    [1,] -1.357779e+09  -501961096.3
##    [2,] -9.550738e+08  -342113349.9
##    [3,] -2.060969e+07    -7388458.0
##    [4,] -2.759316e+05     -275791.5
##    [5,] -2.741149e+05     -286961.2
```

```
##      [6,] -8.061073e+05     -578682.9
##      [7,] -1.357216e+09  -518452322.1
##      [8,] -1.094666e+09  -410947183.8
##      [9,] -1.802854e+07    -8472788.1
##     [10,] -2.742172e+05     -278706.7
##     [11,] -2.747099e+05     -287955.3
##     [12,] -8.000992e+05     -579417.6
##     [13,] -1.262464e+09  -460554468.5
##     [14,] -1.238279e+09  -472025226.9
##     [15,] -1.781608e+07    -7239269.1
##     [16,] -2.766594e+05     -276365.9
##     [17,] -2.751114e+05     -290377.0
##     [18,] -8.831554e+05     -621943.7
##     [19,] -1.555379e+09  -624717316.0
##     [20,] -1.181486e+09  -454343201.6
##     [21,] -2.301343e+07    -9910892.3
##     [22,] -2.728479e+05     -279360.5
##     [23,] -2.779975e+05     -295257.2
##     [24,] -9.120658e+05     -612494.5
##     [25,] -1.713447e+09  -570000198.2
##     [26,] -1.040816e+09  -371941315.6
##     [27,] -1.958861e+07    -8431727.3
##     [28,] -2.745954e+05     -277728.5
##     [29,] -2.783914e+05     -296344.7
##     [30,] -8.351871e+05     -588577.7
##     [31,] -1.388005e+09  -491069212.1
##     [32,] -1.064752e+09  -373656723.9
##     [33,] -2.056022e+07    -8338204.5
##     [34,] -2.723207e+05     -288734.6
##     [35,] -2.850512e+05     -307451.2
##     [36,] -9.049833e+05     -621331.3
##     [37,] -1.265831e+09  -511933762.3
##     [38,] -1.201192e+09  -484273287.0
##     [39,] -2.200628e+07    -8948162.9
##     [40,] -2.731084e+05     -282424.9
##     [41,] -2.799032e+05     -302243.9
##     [42,] -9.989201e+05     -679627.0
##     [43,] -1.381013e+09  -577096561.0
##     [44,] -1.162108e+09  -386271487.4
##     [45,] -2.410394e+07   -10424228.1
##     [46,] -2.723967e+05     -284469.9
##     [47,] -2.806029e+05     -302926.9
##     [48,] -9.646526e+05     -684222.2
##     [49,] -1.387891e+09  -546148989.5
##     [50,] -1.235054e+09  -445744228.1
##     [51,] -2.076307e+07    -8526320.4
##     [52,] -2.725008e+05     -288363.7
##     [53,] -2.802106e+05     -303022.7
##     [54,] -8.716950e+05     -628832.6
##     [55,] -1.285693e+09  -450578174.8
##     [56,] -1.212303e+09  -442294527.5
##     [57,] -2.422120e+07    -9521244.1
##     [58,] -2.723960e+05     -284478.0
##     [59,] -2.818168e+05     -303464.1
##     [60,] -9.493587e+05     -672445.0
##     [61,] -1.794375e+09  -649895340.9
##     [62,] -1.170952e+09  -447821034.4
##     [63,] -2.470843e+07   -10752199.2
```

```
##    [64,] -2.820774e+05     -321554.4
##    [65,] -3.062919e+05     -348805.5
##    [66,] -1.132652e+06     -807103.8
##    [67,] -1.662603e+09  -626319800.3
##    [68,] -1.187612e+09  -472135105.4
##    [69,] -2.471598e+07   -10689475.7
##    [70,] -2.871882e+05     -330186.7
##    [71,] -3.235302e+05     -363733.4
##    [72,] -1.145215e+06     -829452.7
##    [73,] -1.528561e+09  -603179994.1
##    [74,] -1.159773e+09  -451194827.8
##    [75,] -2.410643e+07   -10966473.5
##    [76,] -2.840965e+05     -328768.8
##    [77,] -3.188791e+05     -362420.2
##    [78,] -1.114191e+06     -804723.7
##    [79,] -1.393244e+09  -536481144.1
##    [80,] -1.073068e+09  -410801944.1
##    [81,] -2.191117e+07    -9834408.2
##    [82,] -2.857721e+05     -327359.9
##    [83,] -3.225484e+05     -361538.7
##    [84,] -1.236947e+06     -860235.4
##    [85,] -1.535856e+09  -502506680.9
##    [86,] -9.846363e+08  -378588804.6
##    [87,] -2.113919e+07    -9135635.4
##    [88,] -2.879297e+05     -330554.9
##    [89,] -3.229893e+05     -363898.9
##    [90,] -1.158682e+06     -834824.8
##    [91,] -2.622069e+09  -971563956.7
##    [92,] -2.424890e+09  -699689353.5
##    [93,] -6.493289e+07   -33564865.9
##    [94,] -3.097777e+06    -3668041.5
##    [95,] -3.589327e+06    -3980865.6
##    [96,] -8.192450e+06    -6586075.5
##    [97,] -2.660227e+09  -875226796.1
##    [98,] -2.130688e+09  -955478299.2
##    [99,] -6.683511e+07   -34320247.2
##   [100,] -3.050401e+06    -3643416.1
##   [101,] -3.610012e+06    -3997052.6
##   [102,] -8.237227e+06    -6710798.6
##   [103,] -2.747372e+09  -983073440.0
##   [104,] -2.012493e+09  -729921607.5
##   [105,] -6.838103e+07   -36302598.4
##   [106,] -3.082815e+06    -3651637.0
##   [107,] -3.599266e+06    -3989482.1
##   [108,] -8.140520e+06    -6732913.7
##   [109,] -3.007173e+09  -988877629.1
##   [110,] -1.758968e+09  -628498787.9
##   [111,] -7.065817e+07   -35686182.8
##   [112,] -3.075379e+06    -3657190.8
##   [113,] -3.546784e+06    -3958443.8
##   [114,] -8.151523e+06    -6724634.6
##   [115,] -2.640799e+09  -988829577.5
##   [116,] -1.983353e+09  -743044345.4
##   [117,] -6.962113e+07   -35819973.8
##   [118,] -3.097962e+06    -3670993.5
##   [119,] -3.613805e+06    -4033337.0
##   [120,] -8.320367e+06    -6742885.3
##   [121,] -2.593125e+09  -965525983.3
```

```
##   [122,] -2.041570e+09  -743044345.4
##   [123,] -6.398786e+07   -33247371.9
##   [124,] -3.155924e+06    -3670993.5
##   [125,] -3.646167e+06    -4001059.1
##   [126,] -7.997769e+06    -6645845.9
##   [127,] -2.865614e+09 -1048255153.1
##   [128,] -1.862370e+09  -616283209.9
##   [129,] -7.227354e+07   -39018085.0
##   [130,] -3.043131e+06    -3572786.7
##   [131,] -3.598681e+06    -4007802.5
##   [132,] -8.066284e+06    -6738485.3
##   [133,] -2.621521e+09  -948136591.8
##   [134,] -1.911564e+09  -703662165.1
##   [135,] -6.896359e+07   -34978709.4
##   [136,] -3.158645e+06    -3737515.0
##   [137,] -3.620396e+06    -4047951.1
##   [138,] -8.104004e+06    -6699828.9
##   [139,] -2.805898e+09  -994554331.7
##   [140,] -1.848198e+09  -661253924.6
##   [141,] -6.639829e+07   -33765311.4
##   [142,] -3.095701e+06    -3704343.5
##   [143,] -3.634981e+06    -4042453.4
##   [144,] -8.206659e+06    -6689379.5
##   [145,] -2.659597e+09  -953742452.8
##   [146,] -1.911431e+09  -690732982.6
##   [147,] -6.406108e+07   -33904289.2
##   [148,] -3.085808e+06    -3704479.4
##   [149,] -3.635984e+06    -4053048.6
##   [150,] -8.136962e+06    -6693251.5
##   [151,] -2.573712e+09  -936506774.8
##   [152,] -2.063133e+09  -720814120.8
##   [153,] -6.408549e+07   -33922017.5
##   [154,] -3.118551e+06    -3699090.7
##   [155,] -3.627200e+06    -4026606.2
##   [156,] -8.192614e+06    -6774022.9
##   [157,] -2.697947e+09  -936413252.9
##   [158,] -2.033696e+09  -712089988.4
##   [159,] -6.644801e+07   -34287286.8
##   [160,] -3.085931e+06    -3704615.3
##   [161,] -3.616426e+06    -4057085.2
##   [162,] -8.196915e+06    -6710946.0
##   [163,] -2.717265e+09  -976713187.3
##   [164,] -1.997213e+09  -716384065.5
##   [165,] -6.938723e+07   -35902028.0
##   [166,] -3.091060e+06    -3710280.4
##  [ reached getOption("max.print") -- omitted 53774 rows ]
```

- Utilicen la función optim de R para numericamente el máximo de la función de verosimilitud. Si lo hacen correctamente, su solución debe coincidir con la del método lm.

```r
verosimilitud <- function(b0){
  #Aqui se tiene que ingresar la funcion de verosimilitud
  regresa = 1/(2*pi*sigma(diamantes_num_esc)^(-N/2))*exp^(-1/(2*sigma(diamantes_num_esc)))
}

# a optimizar b0, b1, b2, b3...
#result <- optim(c(1,1,1,1,1,1), verosimilitud, data=diamantes_num_esc)
```

Referencias

https://onlinecourses.science.psu.edu/stat501/node/324 (https://onlinecourses.science.psu.edu/stat501/node/324) \ https://lx.it.pt/~mtf/Figueiredo_Linear_Regression.pdf (https://lx.it.pt/~mtf/Figueiredo_Linear_Regression.pdf) \ https://stats.stackexchange.com/questions/123651/geometric-interpretation-of-multiple-correlation-coefficient-r-and-coefficient (https://stats.stackexchange.com/questions/123651/geometric-interpretation-of-multiple-correlation-coefficient-r-and-coefficient) \ http://www.magesblog.com/2013/03/how-to-use-optim-in-r.html (http://www.magesblog.com/2013/03/how-to-use-optim-in-r.html) \ Notas de clase de propedeutico