

Regresion no parametrica y metodos de regularizacion en modelo lineal

Comerso Salzer Delfina Belen, Ortiz Pedro y Sanchez Goicochea Jhosept Levyt

2023-12-11

En primer lugar importaremos las librerías usadas posteriormente para realizar los ajustes y modelos requeridos. Una vez ya obtenidas las librerías necesarias pasaremos a importar el dataset utilizado, "Body" con datos de morfología corporal humana, nuestro objetivo principal va a ser estudiar la relación entre el peso y las distintas variables morfológicas basadas en las observaciones del mismo.

A su vez, realizamos los cambios necesarios para trabajar con el dataset, agregamos los nombres a las columnas, y llevamos la categorías de Genero al tipo factor (Masculino o Femenino)

```
body <- read_xls("body.xls", col_names = F)

colnames(body) <- c("BIAC", "BIIL", "BITRO", "CHEST1", "CHEST2", "ELBOW", "WRIST",
"KNEE", "ANKLE", "SHOUL", "CHESTG", "WAISTG", "NAVEL", "HIP", "GLUTE", "BICEP",
"FLOREA", "KNEEG", "CALF", "ANKLEG", "WRISTG", "AGE", "WEIG", "HEIG", "GEN"))

body <- body %>% mutate(GEN = as.factor(if_else(GEN == 1, "Male", "Female")) , .keep = "unused") # Consultar
```

Luego chequeamos si existe algún valor faltante en nuestros datos, pues en caso afirmativo no podríamos realizar los modelos, y tendríamos que tomar alguna decisión sobre la o las observaciones que posean algún NA.

```
any(is.na(body))
```

```
## [1] FALSE
```

Por suerte vemos que no hay ningún valor faltante en el dataset.

Ahora sí, yendo al objetivo del estudio, en primer lugar surgió la hipótesis de que el peso (variable WEIG) depende del género de la persona observada, esto viene dado por la intuición de las experiencias personales, donde esto ocurre.

Para empezar a estudiar esto, observaremos las medianas del peso según género, de acuerdo a las observaciones del dataset, separando los datos y viendo los valores de las medianas (valor de la variable en la posición central, si ordenásemos los datos), medida robusta a outliers, a diferencia de lo que puede llegar a ser el promedio.

```
body_male = body %>% filter(GEN == "Male")
body_female = body %>% filter(GEN == "Female")

weig_male <- median(body_male$WEIG)
weig_female <- median(body_female$WEIG)
```

El valor de la mediana para el género masculino fue 77.3, mientras que para el género femenino fue 59. Se puede observar claramente que la mediana es mayor por más de 15 kg., pero estos valores, ¿Son fijos? ¿Varían con otros datos?

Estos valores no son fijos, dependen de los datos, por lo que dadas las observaciones, suena más sensato dar un intervalo en los que se muevan ambos coeficientes, puesto que nos dará mayor información, a la hora de observar nuevos datos.

Para esto utilizaremos la técnica de bootstrap no paramétrico, dado que no conocemos la distribución de la variable peso, utilizaremos dos métodos , en primer caso pensamos que el estimador de la mediana es asintoticamente normal, por lo que para una cantidad de muestras grande, se comportará similar a una normal, por lo que se podrá calcular con esta información un intervalo de confianza. Para el segundo método utilizaremos los cuantiles muestrales de los estadísticos bootstrap del estimador , para calcular otra vez un intervalo de confianza. En ambos casos utilizaremos un intervalo de confianza de nivel 95%

```
N <- 1000
```

```
est_mediana_male_weig <- vector("numeric",N)
est_mediana_female_weig <- vector("numeric",N)

for (i in 1:N){
  num_bootstrap_body <- sample(1:nrow(body),507,replace = T)
  bootstrap_body <- body[num_bootstrap_body,]
  bootstrap_body_male <- bootstrap_body %>% filter(GEN == "Male") %>% select(WEIG)
  bootstrap_body_female <- bootstrap_body %>% filter(GEN == "Female") %>% select(WEIG)
  est_mediana_male_weig[i] <- median(bootstrap_body_male$WEIG)
  est_mediana_female_weig[i] <- median(bootstrap_body_female$WEIG)
}

# Intervalos asumiendo normalidad (TCL)
alpha <- 0.05
z_alpha_sobre_2 <- qnorm(1-alpha/2)

centrado_male_weig <- mean(est_mediana_male_weig)
semi_length_male_weig <- z_alpha_sobre_2 * sd(est_mediana_male_weig)
intervalo_normal_male_weig <- c(centrado_male_weig-semi_length_male_weig, centrado_male_weig+semi_length_male_weig)

centrado_female_weig <- mean(est_mediana_female_weig)
semi_length_female_weig <- z_alpha_sobre_2 * sd(est_mediana_female_weig)
intervalo_normal_female_weig <- c(centrado_female_weig-semi_length_female_weig, centrado_female_weig+semi_length_female_weig)

# Intervalos por cuantiles (TCL)

intervalo_percentil_male_weig <- c(quantile(est_mediana_male_weig,alpha/2),quantile(est_mediana_male_weig,1-alpha/2))
intervalo_percentil_female_weig <- c(quantile(est_mediana_female_weig,alpha/2),quantile(est_mediana_female_weig,1-alpha/2))

#plot()
```

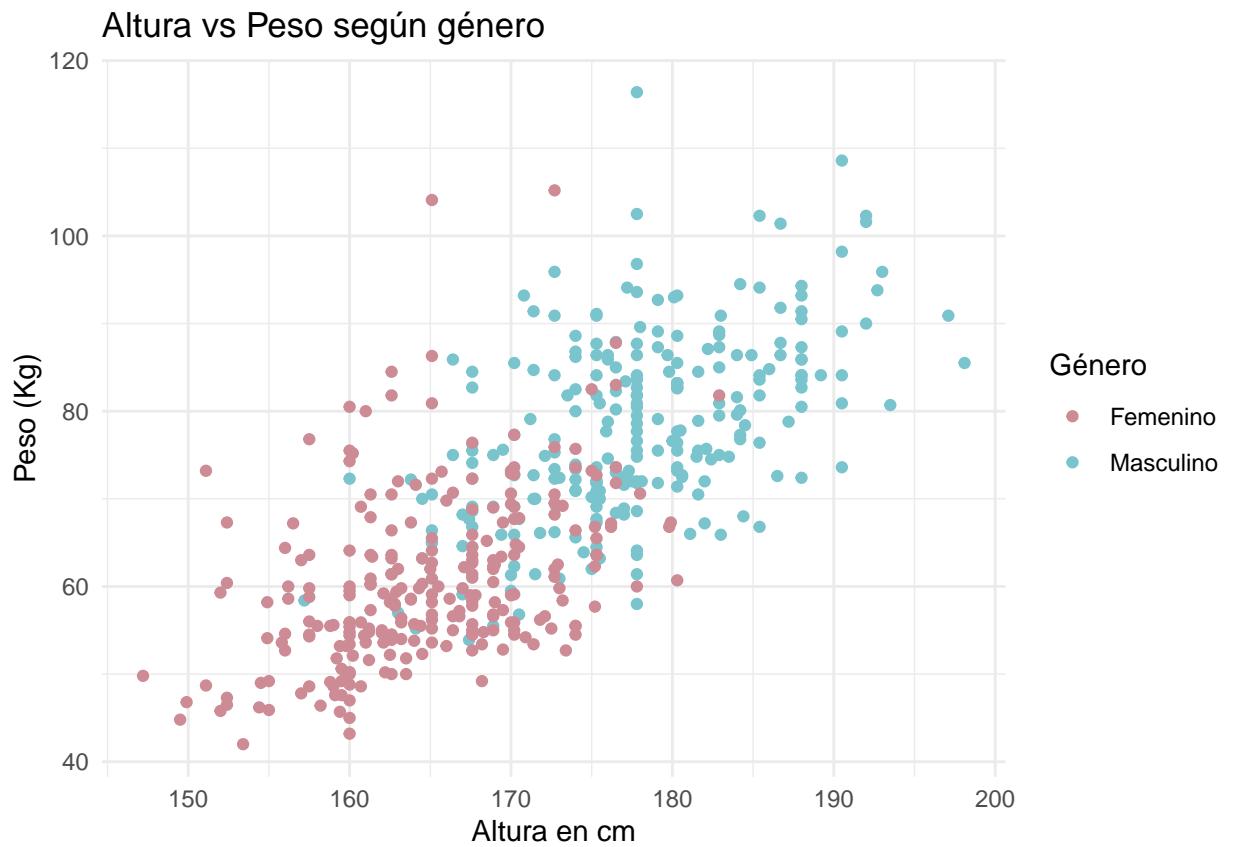
AGREGAR QNORM PARA JUSTIFICAR QUE EL ESTIMADOR DE LA MEADIANA TIENE DISTRIBUCION NORMAL

Notamos que una vez vistos los intervalos, en agregado a los conocimientos sobre regiones de confianza podemos garantizar, con un 95% de probabilidad que los intervalos mostrados cubren a la mediana real. En el caso del género Femenino, en valores alrededor de entre 55 y 60 kilos, mientras que en los hombres en valores alrededor de los 75 y 80 kilogramos. Por lo que nuestra primer hipótesis vemos que se cumple, en general el peso de los hombres es mayor que el peso de las mujeres

Prosiguiendo, nos interesa observar la relación existente entre el peso y la altura de las personas, pero teniendo en cuenta las observaciones hechas anteriormente sobre las diferencias vistas en los pesos según su género, esta relación la veremos desagregada por género.

```
color_f = "#CD8C95"
color_m = "#7AC5CD"

body %>% ggplot(aes(x = HEIG , y = WEIG , color = GEN)) +
  geom_point() +
  labs(title = "Altura vs Peso según género" , x = "Altura en cm" , y = "Peso (Kg)") +
  guides(color = guide_legend(title = "Género")) +
  scale_color_manual(values = c(color_f, color_m) , labels = c("Femenino", "Masculino"))
```



Estudiando el gráfico, en general se observa una correlación positiva entre las variables WEIG y HEIG, indistintamente del género al que se pertenezca. Ahora, discriminando según el género, claramente se nota una tendencia a poseer tanto pesos como alturas mayores del género masculino, por lo que a la hora de realizar modelos que ajusten o traten de encontrar parámetros que expliquen esta relación, el gráfico nos sugiere que podría ser de utilidad separar nuestras observaciones en género masculino y género femenino.

Siguiendo el estudio de la relación entre el peso y la altura, ya visto que dependerá del sexo del individuo, estudiaremos cada género por separado, tratando de ajustar un modelo mediante el estimador de Nadaraya-Watson, con una ventana fija de 10 cm. Este nos dará para cada altura una estimación del peso que tendrá cada individuo, mediante un promedio ponderado de las observaciones que se encuentran dentro de esta ventana

```
model_male_weig_heigh = ksmooth(x = body_male$HEIG, y = body_male$WEIG, kernel = "normal", bandwidth=10)
model_female_weig_heigh = ksmooth(x = body_female$HEIG, y = body_female$WEIG, kernel = "normal", bandwidth=10)
```

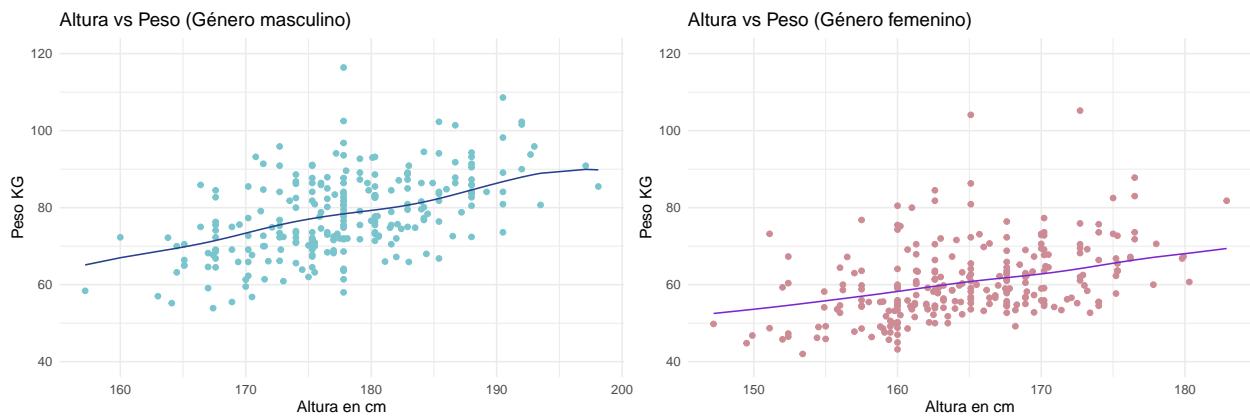
```

plot_HEIG_WEIG_MALE = body_male %>%
  ggplot(aes(x = HEIG , y = WEIG)) +
  geom_point(color = color_m) +
  labs(title = "Altura vs Peso (Género masculino)" , x = "Altura en cm" , y = "Peso KG") +
  ylim(40,120) +
  geom_line(aes(x = model_male_weig_heigh$x , y = model_male_weig_heigh$y), color = "#27408B")

plot_HEIG_WEIG_FEMALE = body_female %>%
  ggplot(aes(x = HEIG , y = WEIG)) +
  geom_point(color = color_f) +
  labs(title = "Altura vs Peso (Género femenino)" , x = "Altura en cm" , y = "Peso KG") +
  ylim(40,120) +
  geom_line(aes(x = model_female_weig_heigh$x , y = model_female_weig_heigh$y), color = "#7D26CD")

plot_grid(plot_HEIG_WEIG_MALE,plot_HEIG_WEIG_FEMALE)

```



Vemos ahora ambos gráficos con sus estimadores, se nota que, sobre todo, para el género femenino el estimador, que puede adoptar cualquier forma, resulta aproximadamente una recta, efecto que se observa en menor medida en la contraparte masculina. Por otro lado, comparando los estimadores en sí, y no las observaciones, se puede apreciar una mayor inclinación (en general), en el de los hombres que en el de las mujeres, por lo que podemos hipotetizar que el peso en los hombres es más sensible a la altura, pero esta apreciación no deja de ser mínima, y sería importante comprobarla realizando modelos lineales, relacionando HEIG y WEIG, para ambos géneros.

Por otro lado estos gráficos no dejan de reflejar las observaciones, y un estimador, realizado con una ventana arbitraria, en este caso 10 cm., por lo que podría no ser la óptima, es por esto, que mediante el proceso de convalidación cruzada elegiremos la óptima, en relación al mínimo error cuadrado. Es decir, la ventana que minimice la diferencia entre los predichos por el estimador, y los valores reales para las observaciones.

```

loocv <- function(X,Y,h){
  n.x <- length(X)
  errores <- vector("numeric",n.x)
  for (i in 1:n.x){
    estimacion_i <- ksmooth(X[-i],Y[-i],"normal",h,x.points = X[i])$y
    errores[i] <- (Y[i] - estimacion_i)**2
  }
  return(mean(errores))
}

h_opt_loocv <- function(X,Y,grilla = seq(5,20,0.5)){

```

```

n.h <- length(grilla)
error_cuadratico <- vector("numeric",n.h)

for (i in 1:n.h){
  error_cuadratico[i] <- loocv(X,Y,grilla[i])
}
datos_aplotear = data.frame("grilla" = grilla, "errores" = error_cuadratico)

return(list(datos_aplotear = datos_aplotear , h_optimo= grilla[which.min(error_cuadratico)]))
}

loocv_opt_male <- h_opt_loocv(body_male$HEIG,body_male$WEIG)
loocv_opt_female <- h_opt_loocv(body_female$HEIG ,body_female$WEIG)

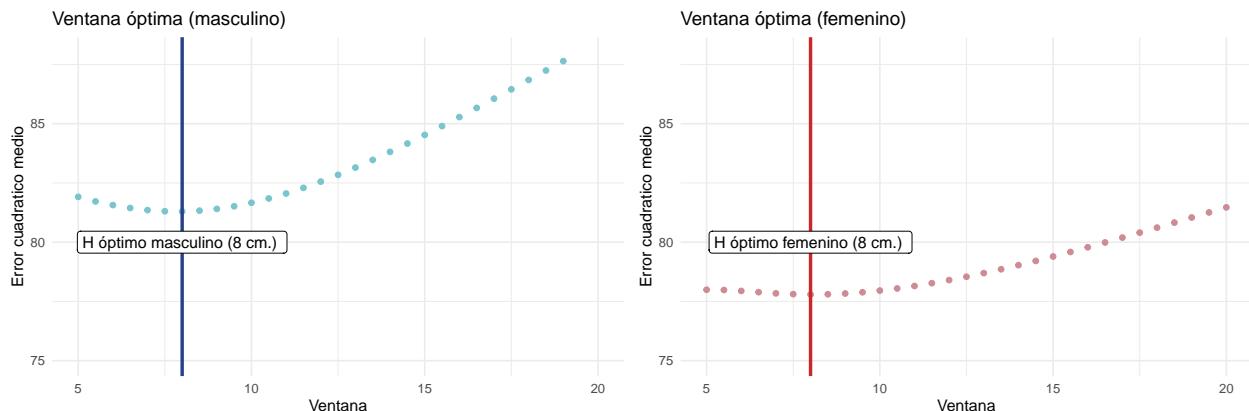
error_min_loocv_male <- min(loocv_opt_male$datos_aplotear[,2])
error_min_loocv_female <- min(loocv_opt_female$datos_aplotear[,2])

h_opt_male_Graph = loocv_opt_male$datos_aplotear %>% ggplot(aes(x = grilla, y = errores)) +
  geom_point( color = color_m) +
  labs(title = "Ventana óptima (masculino)",x = "Ventana", y = "Error cuadrático medio") +
  ylim(75,88) +
  geom_vline(xintercept = loocv_opt_male$h_optimo, color = "#27408B" , linewidth = 1.1) +
  geom_label(aes(x = 8, y = 80,label = "H óptimo masculino (8 cm.) "), stat = "unique")

h_opt_female_Graph = loocv_opt_female$datos_aplotear %>% ggplot(aes(x = grilla, y = errores)) +
  geom_point( color = color_f) +
  labs(title = "Ventana óptima (femenino)",x = "Ventana", y = "Error cuadrático medio") +
  ylim(75,88) +
  geom_vline(xintercept = loocv_opt_female$h_optimo, color = "#CD2626" , linewidth = 1.1) +
  geom_label(aes(x = 8, y = 80,label = "H óptimo femenino (8 cm.) "), stat = "unique")

plot_grid(h_opt_male_Graph,h_opt_female_Graph)

```



Vemos que aunque intuitivamente nos parezca incorrecto la ventana óptima para ambos géneros es de 8 cm., que nos brinda un error cuadrático medio de 81.3 para el género masculino, mientras que para el género femenino nos dio un error de 77.79.

##verificar respuesta la funcion train

```

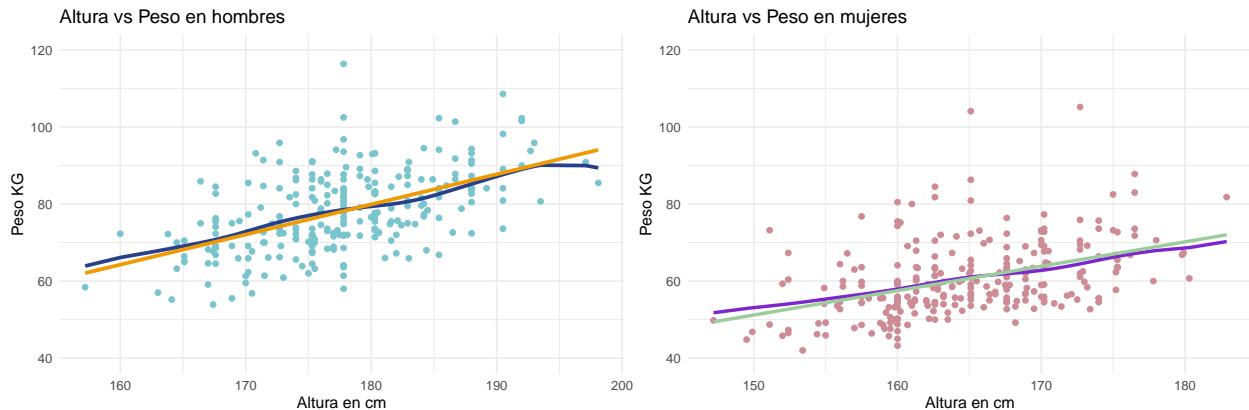
model_male_weig_heigh_opt = ksmooth(x = body_male$HEIG, y = body_male$WEIG, kernel = "normal", bandwidth= 1.5)
model_female_weig_heigh_opt = ksmooth(x = body_female$HEIG, y = body_female$WEIG, kernel = "normal", bandwidth= 1.5)

plot_superpuesto_male = body_male %>%
  ggplot(aes(x = HEIG , y = WEIG)) +
  geom_point(color = color_m) +
  ylim(40,120) +
  labs(title = "Altura vs Peso en hombres" , x = "Altura en cm" , y = "Peso KG") +
  geom_line(aes(x = model_male_weig_heigh_opt$x , y = model_male_weig_heigh_opt$y), color = "#27408B", size= 1.2) +
  geom_smooth(method = "lm", se = FALSE, color = "#EE9A00", linewidth= 1.2)

plot_superpuesto_female = body_female %>%
  ggplot(aes(x = HEIG , y = WEIG)) +
  geom_point(color = color_f) +
  ylim(40,120) +
  labs(title = "Altura vs Peso en mujeres" , x = "Altura en cm" , y = "Peso KG") +
  geom_line(aes(x = model_female_weig_heigh_opt$x , y = model_female_weig_heigh_opt$y), color = "#7D26C8", size= 1.2) +
  geom_smooth(method = "lm", se = FALSE, color = "#9BCD9B", linewidth= 1.1)

modelo_lm_HEiG_WEIG_female = lm(WEIG~HEIG, data =body_female )
modelo_lm_HEiG_WEIG_male = lm(WEIG~HEIG, data =body_male )
Coef_model_HEIH_WEIG_male = modelo_lm_HEiG_WEIG_male$coefficients[2]
Coef_model_HEIH_WEIG_female = modelo_lm_HEiG_WEIG_female$coefficients[2]
plot_grid(plot_superpuesto_male,plot_superpuesto_female)

```



En ambos casos el estimador de nadaraya Watson es muy parecida a la recta aunque tiene la libertad de tomar cualquier forma, esto nos quiere decir que el estimador ideal se asemeja mucho a una recta. entonces podemos utilizar el modelo de minimos cuadrados ya que es la mejor recta que aproxima los datos, sabiendo ademas toda la estructura que tenemos para esto, sabiendo ademas los coeficientes para las variables y sus interpretaciones, etc.

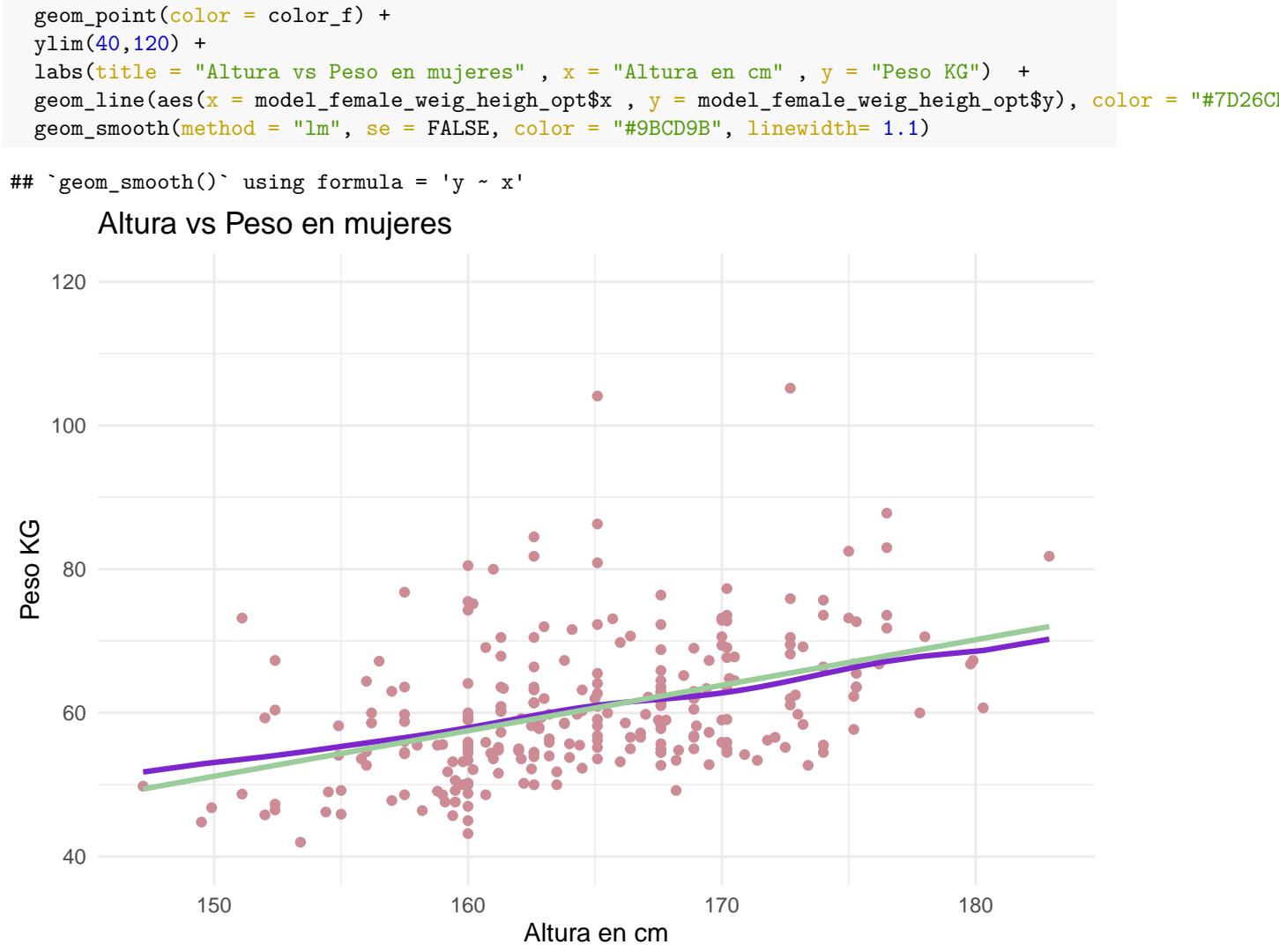
Tambien volviendo a la hipotesis que hicimos de nadaraya watson ahora si graficando y estudiando la recta de cuadrados minimos, podemos estudiar si en verdad el peso de un hombre es mas sensible a la altura con respecto a lo que pasa en el genero femenino.

observamos que el coeficiente de la altura para los hombres en la ecuacion de la recta de cuadrados minimos es 0.78 mientras que el de las mujeres es 0.63, como habiamos pensado anteriormente el peso masculino es mas sensible a la altura que el femenino segun los datos. Estos coeficientes representan la pendiente de la recta, por lo tanto ahí se ve la relacion entre lo predicho antes y la conclusion a partir de los coeficientes.

```

body_female %>%
  ggplot(aes(x = HEIG , y = WEIG)) +

```



```
decision_train_test <- pull(read.table("TrainTest.txt"), V1)
```

```
entrenamiento <- body %>% filter(decision_train_test)
testeo <- body %>% filter(!decision_train_test)
```

```
modelo_total <- lm(WEIG ~ ., data = entrenamiento)
summary(modelo_total)
```

```
##
## Call:
## lm(formula = WEIG ~ ., data = entrenamiento)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.3537 -1.2563 -0.0524  1.2437  8.6171
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.212e+02  3.024e+00 -40.061 < 2e-16 ***
```

```

## BIAC      8.889e-03 7.711e-02 0.115 0.908282
## BIIL      1.324e-01 7.268e-02 1.821 0.069304 .
## BITRO    -1.427e-01 1.024e-01 -1.394 0.164190
## CHEST1    3.387e-01 7.712e-02 4.392 1.45e-05 ***
## CHEST2    1.223e-01 9.272e-02 1.319 0.188070
## ELBOW     2.326e-02 2.165e-01 0.107 0.914479
## WRIST     2.597e-01 2.505e-01 1.037 0.300525
## KNEE      4.609e-01 1.555e-01 2.964 0.003227 **
## ANKLE     9.814e-02 1.812e-01 0.542 0.588279
## SHOUL     7.694e-02 3.537e-02 2.175 0.030231 *
## CHESTG    1.518e-01 4.440e-02 3.420 0.000694 ***
## WAISTG    3.636e-01 3.102e-02 11.721 < 2e-16 ***
## NAVEL     -2.973e-02 2.789e-02 -1.066 0.287001
## HIP       2.588e-01 5.285e-02 4.898 1.42e-06 ***
## GLUTE     2.142e-01 5.946e-02 3.602 0.000357 ***
## BICEP      4.899e-02 9.180e-02 0.534 0.593904
## FLOREA    6.033e-01 1.559e-01 3.869 0.000128 ***
## KNEEG     2.165e-01 8.882e-02 2.437 0.015252 *
## CALF      3.712e-01 7.649e-02 4.854 1.76e-06 ***
## ANKLEG    1.871e-02 1.121e-01 0.167 0.867519
## WRISTG   -3.176e-01 2.374e-01 -1.338 0.181815
## AGE       -5.570e-02 1.444e-02 -3.858 0.000134 ***
## HEIG      3.005e-01 2.060e-02 14.586 < 2e-16 ***
## GENMale   -1.536e+00 5.847e-01 -2.627 0.008948 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.124 on 388 degrees of freedom
## Multiple R-squared: 0.9764, Adjusted R-squared: 0.9749
## F-statistic: 669 on 24 and 388 DF, p-value: < 2.2e-16

```

el valor del estadistica del estadistico F me dice ???

El p-valor me dice que la regresion es significitva ya que es muy cercana a 0.

El modelo ajustado es este o como quieren que lo pongamos

Nos quedamos con todos los que tengan 3 estrellitas de significacion, buscamos una significacion mayor al 99%.

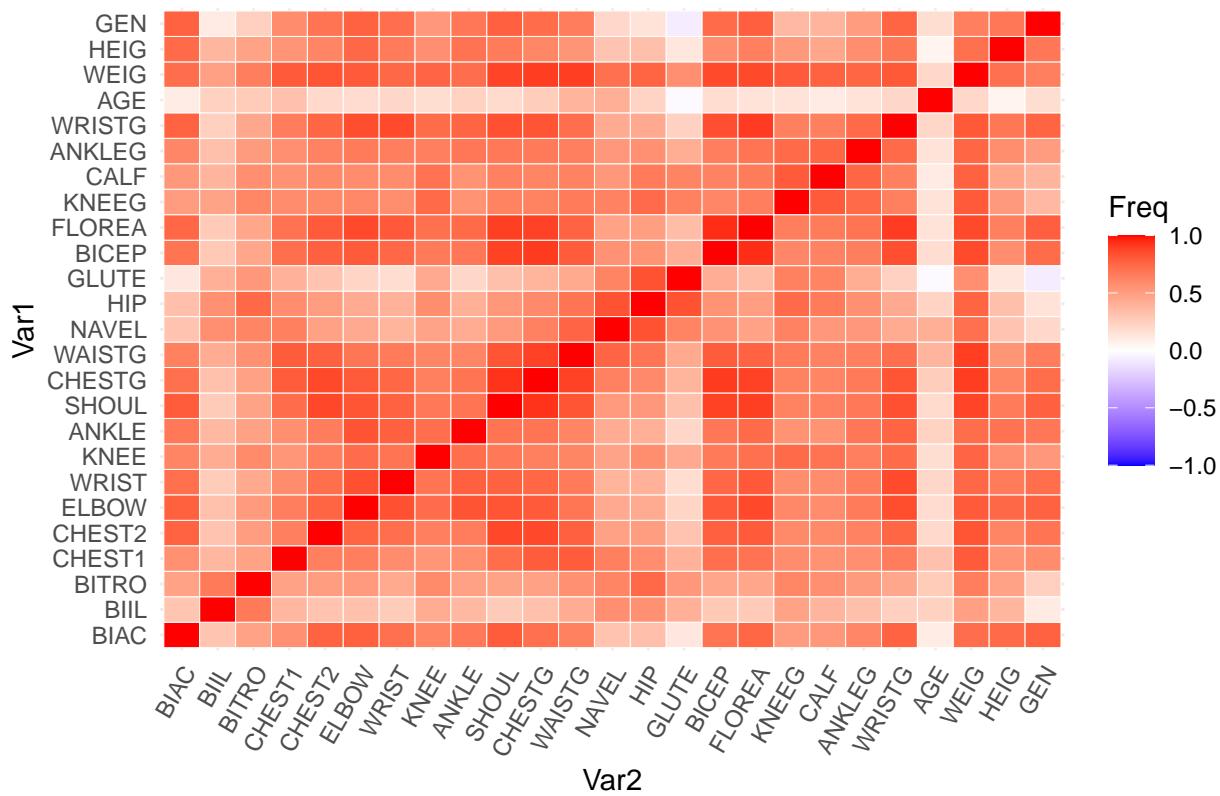
como lo relacionamos F con el item anterior.

Vimos antes en el ejercicio anterior que el modelo ksmooth se parece mucho a una recta pero es solo para la variable heigh y no para todas.

Ver como queda el estadistico f en el ejercicio anterior.

```
entrenamiento2 <- entrenamiento %>% mutate(GEN = if_else(GEN == "Male", 1, 0) , .keep = "unused")  
matriz_correlacion <- cor(entrenamiento2)  
  
ggplot(data = as.data.frame(as.table(matriz_correlacion)), aes(Var2, Var1, fill = Freq)) +  
  geom_tile(color = "white") +  
  scale_fill_gradient2(low = "blue", mid = "white", high = "red", midpoint = 0, limit = c(-1, 1)) +  
  theme_minimal() +  
  labs(title = "Heatmap de Correlación") +  
  theme(axis.text.x = element_text(angle = 60, hjust = 1))
```

Heatmap de Correlación



```
mean((testeo$WEIG - predict(modelo_total,newdata = testeo))^2)
```

```
## [1] 4.225356
```

Consultar que es el error empírico si el promedio esta bien

```
solve(matriz_correlacion)
```

```
##          BIAC        BIIL       BITRO      CHEST1      CHEST2
## BIAC  4.84889319 -0.380875707 -0.78306139  0.083065766 -1.159006632
## BIIL -0.38087571  2.360758749 -1.07683951 -0.001620463  0.065770983
## BITRO -0.78306139 -1.076839510  3.95564815 -0.043188374 -0.604653215
## CHEST1  0.08306577 -0.001620463 -0.04318837  3.613889580  0.742114880
## CHEST2 -1.15900663  0.065770983 -0.60465322  0.742114880  5.748241247
## ELBOW -0.44692900  0.319151972 -0.90407607 -0.304736709 -0.094474950
## WRIST  0.16758232  0.048158821  0.13709871  0.323001183 -0.279995799
## KNEE  -0.08243672 -0.255294879 -0.29905956  0.276248051 -0.281851821
## ANKLE  0.29480203 -0.365453011  0.20319093  0.183950073  0.090186096
## SHOUL -1.92199018  0.716873841  0.30013030  0.201503517 -1.545187000
## CHESTG 1.18291110  0.100820701  0.88906080 -1.533369041 -3.637581641
## WAISTG -0.19263289 -0.009358140  0.03620842 -0.041473198 -0.425001989
## NAVEL  0.02441457 -0.867801179  0.45060147 -0.488970681  0.620679705
## HIP   0.62805887  0.514002336 -3.20571594  0.658758491  0.602859298
## GLUTE 0.09376262  0.146422888  0.56541186 -0.397628525 -0.358454943
## BICEP 0.29106939 -0.001303919 -0.37608359  0.333210112  1.394452710
## FLOREA 0.43891749 -0.339951454  0.19706450  0.508654764 -0.280999016
## KNEEG 0.33402988 -0.231663124  0.33135766  0.143253214  0.006707594
```

```

## CALF   -0.13678900  0.042636373 -0.39331993 -0.013092918  0.053162178
## ANKLEG -0.38061851  0.333327186 -0.05292089  0.114058157  0.008292182
## WRISTG -1.01131800  0.603050719  0.01629392 -0.385800391  0.525151599
## AGE     0.37541626 -0.060948045 -0.39035981 -0.400788608 -0.142602235
## WEIG    -0.08389946 -0.921055506  0.91389361 -2.693588226 -1.042569240
## HEIG    -0.65482317 -0.369507515 -0.60361251  0.091898840  0.332698813
## GEN     -1.24896317  0.423544085  0.39028525 -0.704382528  0.078660511
##          ELBOW      WRIST      KNEE      ANKLE      SHOUL
## BIAC    -0.44692900  0.167582320 -0.08243672  0.29480203 -1.921990179
## BIIL    0.31915197  0.048158821 -0.25529488 -0.36545301  0.716873841
## BITRO   -0.90407607  0.137098709 -0.29905956  0.20319093  0.300130301
## CHEST1  -0.30473671  0.323001183  0.27624805  0.18395007  0.201503517
## CHEST2  -0.09447495 -0.279995799 -0.28185182  0.09018610 -1.545187000
## ELBOW   7.82403234 -1.414627914 -0.47494494 -1.74582390 -0.319229265
## WRIST   -1.41462791  5.076547148 -0.13367765 -0.81930694  0.072998041
## KNEE    -0.47494494 -0.133677654  4.09472955 -1.19153120 -0.259979102
## ANKLE   -1.74582390 -0.819306941 -1.19153120  4.61517431  0.744998263
## SHOUL   -0.31922926  0.072998041 -0.25997910  0.74499826 12.664487196
## CHESTG  -0.64504324 -0.062055967  1.24493702 -0.74548631 -5.501571133
## WAISTG  1.32900120 -0.179676543  0.77062762  0.71859941  0.771787457
## NAVEL   -0.73983108  0.468382579  0.37975130 -0.51493233  0.670759489
## HIP     0.29416872 -0.599336773 -0.18532744  0.26196973 -1.168609037
## GLUTE   0.52223747  0.849430063 -0.13554100 -0.35137811  1.182301900
## BICEP   0.24846432 -0.349163386 -0.26664187  0.48830008 -2.008890229
## FLOREA  -2.81021403  0.584876747  0.53516015  0.09043950 -0.515050844
## KNEEG   -0.39963469 -0.003048605 -1.22868437  0.89335117 -0.308134538
## CALF    0.19545569 -0.360659069 -0.31055586  0.12877261 -0.348572767
## ANKLEG  0.38737194  0.261065509  0.68574916 -1.21081098 -0.112277033
## WRISTG  0.31167710 -2.410867648 -1.12751079  0.05136895  0.540279109
## AGE     0.09400456 -0.080787727 -0.14689656 -0.32893311 -0.081164364
## WEIG   -0.09934449 -0.770935536 -1.96007962 -0.38453881 -2.542810523
## HEIG   -0.73647214  0.018782322  0.85329197 -0.56820830  0.001268627
## GEN     -0.60652796  0.229108699 -0.39318465 -1.00502359 -0.654053422
##          CHESTG      WAISTG      NAVEL      HIP      GLUTE
## BIAC    1.182911096 -0.19263289  0.02441457  0.62805887  0.09376262
## BIIL    0.100820701 -0.00935814 -0.86780118  0.51400234  0.14642289
## BITRO   0.889060798  0.03620842  0.45060147 -3.20571594  0.56541186
## CHEST1 -1.533369041 -0.04147320 -0.48897068  0.65875849 -0.39762853
## CHEST2 -3.637581641 -0.42500199  0.62067971  0.60285930 -0.35845494
## ELBOW   -0.645043237  1.32900120 -0.73983108  0.29416872  0.52223747
## WRIST   -0.062055967 -0.17967654  0.46838258 -0.59933677  0.84943006
## KNEE    1.244937020  0.77062762  0.37975130 -0.18532744 -0.13554100
## ANKLE   -0.745486306  0.71859941 -0.51493233  0.26196973 -0.35137811
## SHOUL   -5.501571133  0.77178746  0.67075949 -1.16860904  1.18230190
## CHESTG  18.743524765 -2.47291001 -1.47137519  0.66720270  0.39119032
## WAISTG -2.472910014  14.23947738 -2.96733342 -1.02704977  1.71681010
## NAVEL   -1.471375187 -2.96733342  6.49562751 -3.29556428  0.57130004
## HIP     0.667202699 -1.02704977 -3.29556428 12.50318448 -4.90324269
## GLUTE   0.391190318  1.71681010  0.57130004 -4.90324269  7.01780063
## BICEP   -3.915632621  0.43690064 -1.32308999  1.70479029 -2.91200351
## FLOREA -0.252353620  2.47605402  0.61400352  0.36049537  0.19504670
## KNEEG   0.650384836 -0.01036712 -0.49291343 -0.19831319 -0.72440366
## CALF    1.067084320  0.29938914  0.58870089  1.00576338 -1.22929453
## ANKLEG  0.069214185  0.16784188 -0.60311786 -0.13933387  0.25927458

```

```

## WRISTG -1.103897802 -0.25762332 0.36168951 -0.30949477 1.11563804
## AGE 0.007447309 -1.09291237 -0.66756287 -0.07431922 0.66683895
## WEIG -4.821081000 -12.56220966 0.89679907 -5.55499588 -3.10200563
## HEIG 0.977301304 3.28436860 -0.01437632 1.06605998 0.79199184
## GEN 0.827565340 -4.77742856 1.93059074 0.44484650 1.45469918
## BICEP FLOREA KNEEG CALF ANKLEG
## BIAC 0.291069390 0.4389175 0.334029881 -0.13678900 -0.380618510
## BIIL -0.001303919 -0.3399515 -0.231663124 0.04263637 0.333327186
## BITRO -0.376083590 0.1970645 0.331357663 -0.39331993 -0.052920889
## CHEST1 0.333210112 0.5086548 0.143253214 -0.01309292 0.114058157
## CHEST2 1.394452710 -0.2809990 0.006707594 0.05316218 0.008292182
## ELBOW 0.248464317 -2.8102140 -0.399634688 0.19545569 0.387371944
## WRIST -0.349163386 0.5848767 -0.003048605 -0.36065907 0.261065509
## KNEE -0.266641872 0.5351601 -1.228684371 -0.31055586 0.685749157
## ANKLE 0.488300081 0.0904395 0.893351168 0.12877261 -1.210810979
## SHOUL -2.008890229 -0.5150508 -0.308134538 -0.34857277 -0.112277033
## CHESTG -3.915632621 -0.2523536 0.650384836 1.06708432 0.069214185
## WAISTG 0.436900642 2.4760540 -0.010367122 0.29938914 0.167841876
## NAVEL -1.323089993 0.6140035 -0.492913435 0.58870089 -0.603117862
## HIP 1.704790293 0.3604954 -0.198313193 1.00576338 -0.139333865
## GLUTE -2.912003507 0.1950467 -0.724403665 -1.22929453 0.259274577
## BICEP 14.170955580 -8.5810931 1.421112741 0.57180294 0.149100136
## FLOREA -8.581093087 18.5551441 -0.760480372 -0.48688213 0.369413299
## KNEEG 1.421112741 -0.7604804 5.102584433 -0.98486126 -1.059166155
## CALF 0.571802944 -0.4868821 -0.984861259 4.59302909 -1.383843112
## ANKLEG 0.149100136 0.3694133 -1.059166155 -1.38384311 3.942745167
## WRISTG -0.772399288 -4.6733483 -0.497730168 -0.08396905 -1.539137978
## AGE -0.075567534 0.3797971 -0.076229587 -0.19925198 0.174766027
## WEIG -0.663688465 -5.4054966 -1.805778674 -3.33808951 -0.109541634
## HEIG 0.706117056 1.8640284 -0.355985264 1.07404993 -0.089780030
## GEN -1.412002655 -2.5453941 0.203651099 0.07243386 0.060786731
## WRISTG AGE WEIG HEIG GEN
## BIAC -1.01131800 0.375416263 -0.08389946 -0.654823173 -1.24896317
## BIIL 0.60305072 -0.060948045 -0.92105551 -0.369507515 0.42354409
## BITRO 0.01629392 -0.390359809 0.91389361 -0.603612506 0.39028525
## CHEST1 -0.38580039 -0.400788608 -2.69358823 0.091898840 -0.70438253
## CHEST2 0.52515160 -0.142602235 -1.04256924 0.332698813 0.07866051
## ELBOW 0.31167710 0.094004560 -0.09934449 -0.736472144 -0.60652796
## WRIST -2.41086765 -0.080787727 -0.77093554 0.018782322 0.22910870
## KNEE -1.12751079 -0.146896564 -1.96007962 0.853291966 -0.39318465
## ANKLE 0.05136895 -0.328933111 -0.38453881 -0.568208301 -1.00502359
## SHOUL 0.54027911 -0.081164364 -2.54281052 0.001268627 -0.65405342
## CHESTG -1.10389780 0.007447309 -4.82108100 0.977301304 0.82756534
## WAISTG -0.25762332 -1.092912373 -12.56220966 3.284368599 -4.77742856
## NAVEL 0.36168951 -0.667562869 0.89679907 -0.014376321 1.93059074
## HIP -0.30949477 -0.074319216 -5.55499588 1.066059980 0.44484650
## GLUTE 1.11563804 0.666838946 -3.10200563 0.791991842 1.45469918
## BICEP -0.77239929 -0.075567534 -0.66368846 0.706117056 -1.41200265
## FLOREA -4.67334830 0.379797103 -5.40549655 1.864028366 -2.54539406
## KNEEG -0.49773017 -0.076229587 -1.80577867 -0.355985264 0.20365110
## CALF -0.08396905 -0.199251979 -3.33808951 1.074049928 0.07243386
## ANKLEG -1.53913798 0.174766027 -0.10954163 -0.089780030 0.06078673
## WRISTG 9.71685551 -0.415427364 1.37489821 -0.708485979 0.72681484
## AGE -0.41542736 1.778418426 1.66871155 0.128540039 0.01928579

```

```

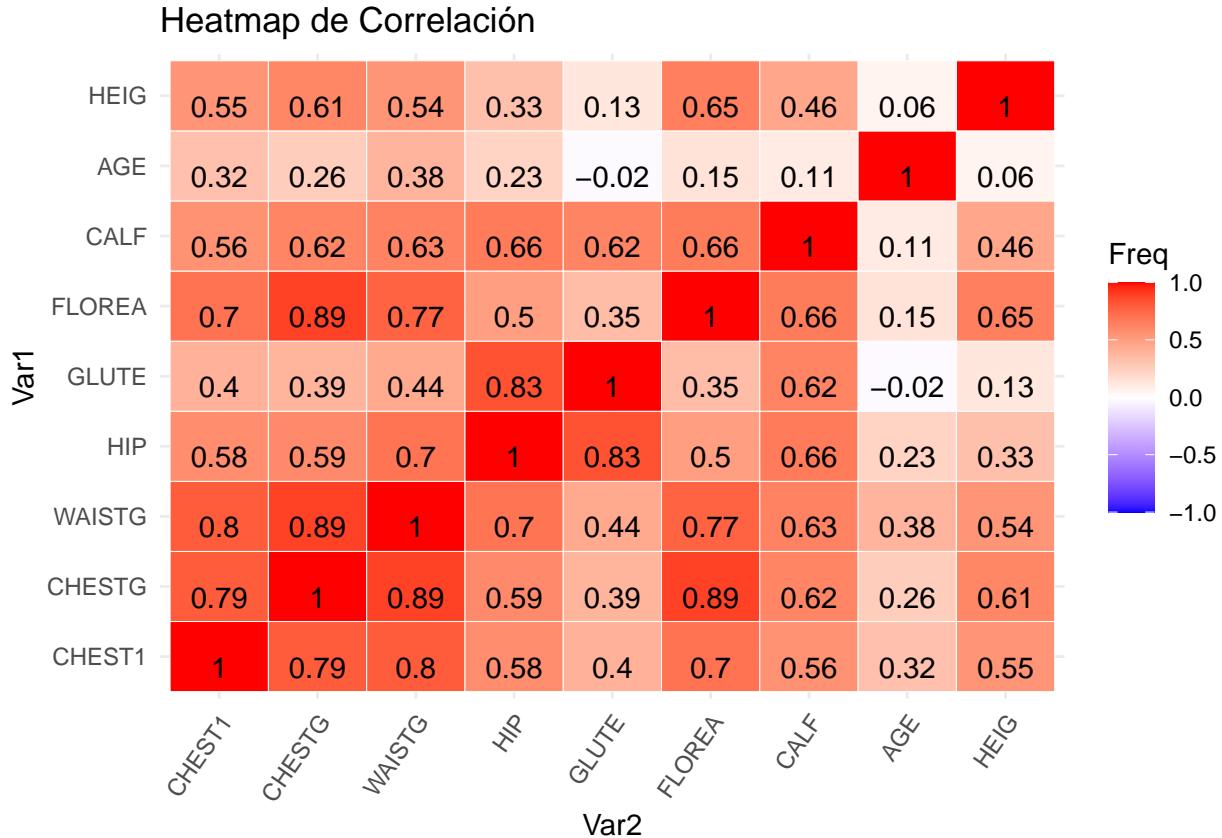
## WEIG      1.37489821  1.668711553  42.38399287 -9.017700311  2.42778380
## HEIG     -0.70848598  0.128540039  -9.01770031   5.417814603 -1.62382331
## GEN       0.72681484  0.019285794   2.42778380 -1.623823313  7.95588410
variables_tres_estrellas = names(summary(modelo_total)$coefficients[, "Pr(>|t|)"])[summary(modelo_total)$coefficients[, "Pr(>|t|)"] >= 0.05]

vector_tres_estrellas <- as.vector(variables_tres_estrellas[-1])

matriz_correlacion_2 <- cor(entrenamiento2[,vector_tres_estrellas])

# Crear el heatmap con ggplot2
ggplot(data = as.data.frame(as.table(matriz_correlacion_2)), aes(Var2, Var1, fill = Freq)) +
  geom_tile(color = "white") +
  geom_text(aes(label = round(Freq, 2)), vjust = 1) +
  scale_fill_gradient2(low = "blue", mid = "white", high = "red", midpoint = 0, limit = c(-1, 1)) +
  theme_minimal() +
  labs(title = "Heatmap de Correlación") +
  theme(axis.text.x = element_text(angle = 55, hjust = 1)) # Rotar etiquetas del eje x

```



Vemos que chest tiene mucha correlacion con otras variables por lo que decidimos sacarla y dejar las otras, veamos como queda la matriz de correlacion

```

vector_tres_estrellas_SINCG <- vector_tres_estrellas[-2]

matriz_correlacion_2 <- cor(entrenamiento2[,vector_tres_estrellas_SINCG])

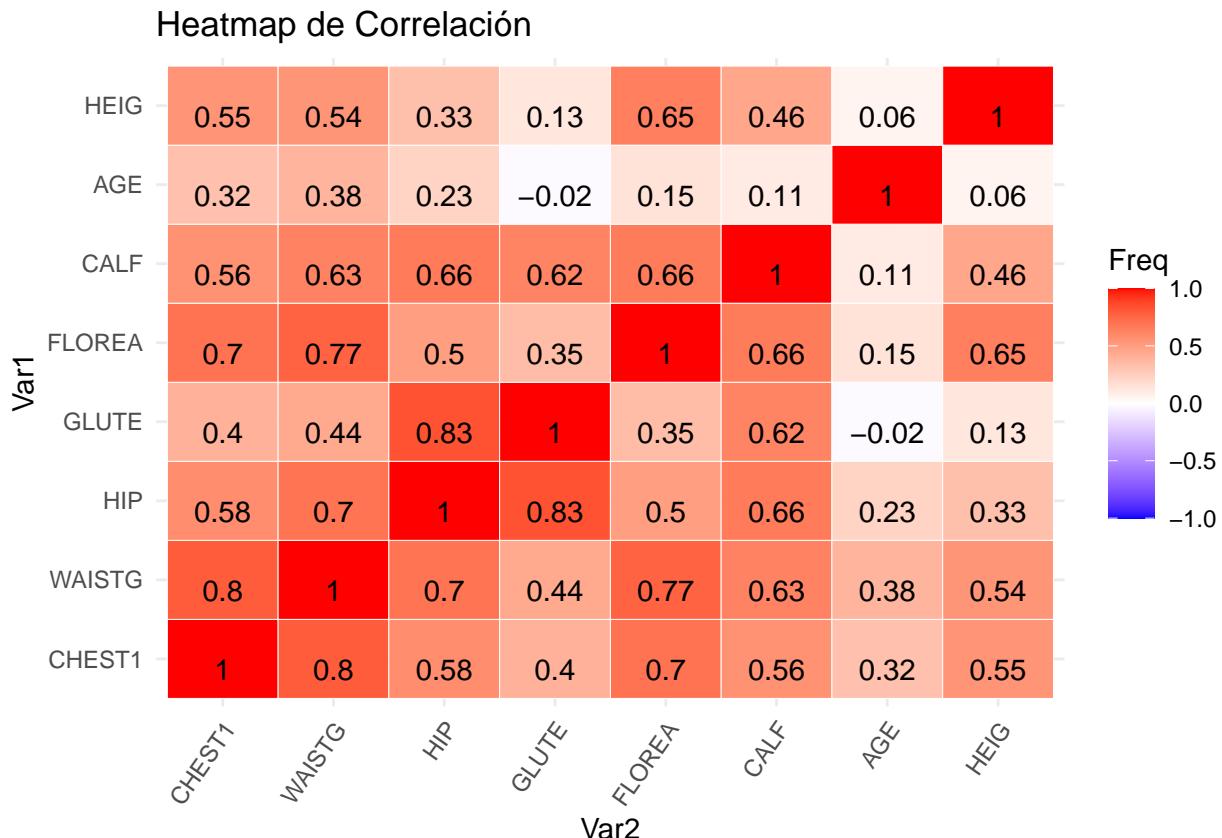
# Crear el heatmap con ggplot2

```

```

ggplot(data = as.data.frame(as.table(matriz_correlacion_2)), aes(Var2, Var1, fill = Freq)) +
  geom_tile(color = "white") +
  geom_text(aes(label = round(Freq, 2)), vjust = 1) +
  scale_fill_gradient2(low = "blue", mid = "white", high = "red", midpoint = 0, limit = c(-1, 1)) +
  theme_minimal() +
  labs(title = "Heatmap de Correlación") +
  theme(axis.text.x = element_text(angle = 55, hjust = 1)) # Rotar etiquetas del eje x

```



```

formula <- as.formula(paste("WEIG ~", paste(vector_tres_estrellas_SINCG, collapse = "+")))

modelo_limpio = lm(formula, data = entrenamiento2)

summary(modelo_total) #- summary(modelo_limpio)$sigma**2

##
## Call:
## lm(formula = WEIG ~ ., data = entrenamiento)
##
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -7.3537 -1.2563 -0.0524  1.2437  8.6171 
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) -1.212e+02  3.024e+00 -40.061 < 2e-16 ***
## BIAC        8.889e-03  7.711e-02   0.115  0.908282  
## BIIL        1.324e-01  7.268e-02   1.821  0.069304 .

```

```

## BITRO      -1.427e-01  1.024e-01  -1.394  0.164190
## CHEST1     3.387e-01  7.712e-02   4.392  1.45e-05 ***
## CHEST2     1.223e-01  9.272e-02   1.319  0.188070
## ELBOW      2.326e-02  2.165e-01   0.107  0.914479
## WRIST      2.597e-01  2.505e-01   1.037  0.300525
## KNEE       4.609e-01  1.555e-01   2.964  0.003227 **
## ANKLE      9.814e-02  1.812e-01   0.542  0.588279
## SHOUL      7.694e-02  3.537e-02   2.175  0.030231 *
## CHESTG     1.518e-01  4.440e-02   3.420  0.000694 ***
## WAISTG     3.636e-01  3.102e-02  11.721  < 2e-16 ***
## NAVEL      -2.973e-02  2.789e-02  -1.066  0.287001
## HIP        2.588e-01  5.285e-02   4.898  1.42e-06 ***
## GLUTE      2.142e-01  5.946e-02   3.602  0.000357 ***
## BICEP       4.899e-02  9.180e-02   0.534  0.593904
## FLOREA     6.033e-01  1.559e-01   3.869  0.000128 ***
## KNEEG       2.165e-01  8.882e-02   2.437  0.015252 *
## CALF        3.712e-01  7.649e-02   4.854  1.76e-06 ***
## ANKLEG     1.871e-02  1.121e-01   0.167  0.867519
## WRISTG    -3.176e-01  2.374e-01  -1.338  0.181815
## AGE        -5.570e-02  1.444e-02  -3.858  0.000134 ***
## HEIG        3.005e-01  2.060e-02  14.586  < 2e-16 ***
## GENMale    -1.536e+00  5.847e-01  -2.627  0.008948 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.124 on 388 degrees of freedom
## Multiple R-squared:  0.9764, Adjusted R-squared:  0.9749
## F-statistic:  669 on 24 and 388 DF,  p-value: < 2.2e-16
summary(modelo_total)$r.squared - summary(modelo_limpio)$r.squared

## [1] 0.005806928
summary(modelo_limpio)$sigma**2 - summary(modelo_total)$sigma**2

## [1] 0.8880515
mean((testeo$WEIG - predict(modelo_limpio,newdata = testeo))^2)

## [1] 5.194494

```

R squared representa la proporción de variabilidad de la variable Y que es explicada por el modelo propuesto relativa a cuanta variabilidad tiene Y respecto de su media (es decir, a ser explicada sin aplicar el modelo).

El modelo total es un poco mejor pero la diferencia es poco significativa para justificar el uso de tantas variables extras.

Todos los coeficientes nos quedaron muy significativos mayores al 99%. (CONSULTAR SI ES 99% O ES 99.9%)

```
matriz_diseño_sin_intercept <- model.matrix(lm(WEIG ~ . + 0, data = entrenamiento)) #Solo lo usamos para el lasso

matriz_diseño_sin_intercept_test <- model.matrix(lm(WEIG ~ . + 0, data = testeo))

modelo_lasso = glmnet(x = matriz_diseño_sin_intercept , y = entrenamiento$WEIG, alpha = 1)

coef(modelo_lasso)

## 26 x 79 sparse Matrix of class "dgCMatrix"

##      [[ suppressing 79 column names 's0', 's1', 's2' ... ]]

##
## (Intercept) 69.37143 59.74210065 51.24055293 43.4878899 36.4312770 28.788554427
## BIAC          .
## BIIL          .
## BITRO         .
## CHEST1        .
## CHEST2        .
## ELBOW         .
## WRIST          .
## KNEE          .
## ANKLE          .
## SHOUL          .
## CHESTG        0.07752380 0.12920842 0.1765092 0.2193691 0.250679305
## WAISTG        0.03090824 0.07850605 0.1217063 0.1612632 0.192430309
## NAVEL         .
## HIP           .
## GLUTE         .
## BICEP          .
## FLOREA         .
## KNEEG          .
## CALF          .
## ANKLEG         .
## WRISTG         .
## AGE           .
## HEIG          .
## GENFemale     .
## GENMale        .
## 
```

```

## (Intercept) 19.68353485 11.41322707 3.8702754 -3.00239364 -10.05431326
## BIAC . . . . .
## BIIL . . . . .
## BITRO . . . . .
## CHEST1 . . . . .
## CHEST2 . . . . .
## ELBOW . . . . .
## WRIST . . . . .
## KNEE . . . . .
## ANKLE . . . . .
## SHOUL 0.01207651 0.02404502 0.0352395 0.04541149 0.05740292
## CHESTG 0.25485606 0.25201552 0.2498203 0.24785728 0.24225729
## WAISTG 0.21041295 0.22815426 0.2442609 0.25893312 0.26700727
## NAVEL . . . . .
## HIP . . . . .
## GLUTE . . . . .
## BICEP . . . . .
## FLOREA 0.06211135 0.12551728 0.1814249 0.23235284 0.28045679
## KNEEG 0.18550624 0.30215986 0.4082361 0.50488838 0.54130066
## CALF . . . . .
## ANKLEG . . . . .
## WRISTG . . . . .
## AGE . . . . .
## HEIG . . . . .
## GENFemale . . . . .
## GENMale . . . . .
##
## (Intercept) -19.04819860 -28.01779546 -36.20057745 -43.65716691 -50.44931853
## BIAC . . . . .
## BIIL . . . . .
## BITRO . . . . .
## CHEST1 . . . . .
## CHEST2 . . . . .
## ELBOW . . . . .
## WRIST . . . . .
## KNEE . . . . .
## ANKLE . . . . .
## SHOUL 0.06154317 0.06048891 0.06045741 0.06061090 0.06046041
## CHESTG 0.24006815 0.23986201 0.23884773 0.23787764 0.23707630
## WAISTG 0.26791753 0.26892872 0.26970652 0.27039659 0.27107955
## NAVEL . . . . .
## HIP 0.04989597 0.08326322 0.11403337 0.14208305 0.16751541
## GLUTE . . . . .
## BICEP . . . . .
## FLOREA 0.31594429 0.34296182 0.36782520 0.39015528 0.41087462
## KNEEG 0.53286056 0.51750467 0.50237645 0.48857704 0.47678564
## CALF 0.08628037 0.12535526 0.16104769 0.19353853 0.22290654
## ANKLEG . . . . .
## WRISTG . . . . .
## AGE . . . . .
## HEIG 0.01870320 0.04345753 0.06597634 0.08646871 0.10514177
## GENFemale . . . . .
## GENMale . . . . .
##
```

```

## (Intercept) -56.64031813 -62.299140777 -67.518146974 -72.15936908 -76.40934828
## BIAC . . . .
## BIIL . . . .
## BITRO . . . .
## CHEST1 . . . . 0.001235396 0.02808378 0.05024599
## CHEST2 . . . . .
## ELBOW . . . . .
## WRIST . . . . .
## KNEE . . 0.006605458 0.034869582 0.06644420 0.09007249
## ANKLE . . . .
## SHOUL 0.06060894 0.060800413 0.060323728 0.06102470 0.06135187
## CHESTG 0.23625712 0.235452692 0.235097144 0.23218432 0.22958385
## WAISTG 0.27164490 0.272089933 0.272409727 0.27105699 0.26968851
## NAVEL . . . .
## HIP 0.19081853 0.211887083 0.231096778 0.24736373 0.26245725
## GLUTE . . . .
## BICEP . . . .
## FLOREA 0.42941464 0.445248753 0.456672938 0.46476249 0.47380800
## KNEEG 0.46525302 0.453261838 0.438093131 0.42402530 0.41262118
## CALF 0.24990315 0.274613764 0.295360960 0.31268724 0.32896755
## ANKLEG . . . .
## WRISTG . . . .
## AGE .
## HEIG 0.12215720 0.137609610 0.151432786 0.16269875 0.17332986
## GENFemale .
## GENMale .
##
## (Intercept) -8.028072e+01 -83.788923815 -86.93778664 -89.82206616 -92.59765564
## BIAC . . . .
## BIIL . . . .
## BITRO . . . .
## CHEST1 6.852502e-02 0.086026781 0.10474206 0.11947871 0.13395995
## CHEST2 . . . .
## ELBOW . . . .
## WRIST . . . .
## KNEE 1.091478e-01 0.127577774 0.14631062 0.16222223 0.18840835
## ANKLE . . . .
## SHOUL 6.060873e-02 0.060490440 0.06142879 0.06128197 0.06452971
## CHESTG 2.275641e-01 0.225030579 0.22081088 0.21819842 0.21453158
## WAISTG 2.687169e-01 0.267563336 0.26608170 0.26519746 0.26682232
## NAVEL . . . .
## HIP 2.762315e-01 0.288688564 0.29938882 0.30930577 0.30453697
## GLUTE . . . .
## BICEP 1.583267e-04 0.004437745 0.01772932 0.02628056 0.03795522
## FLOREA 4.838685e-01 0.488344236 0.48077031 0.47810315 0.45989272
## KNEEG 4.025519e-01 0.393730744 0.38720829 0.38040074 0.36630386
## CALF 3.443864e-01 0.357919517 0.36972245 0.38115889 0.38403910
## ANKLEG . . . .
## WRISTG . . . .
## AGE .
## HEIG 1.833311e-01 0.192310585 0.20047660 0.20809630 0.21681950
## GENFemale .
## GENMale .
##

```

```

## (Intercept) -95.22973868 -97.61492375 -99.7751727 -1.017526e+02 -1.035242e+02
## BIAC .
## BIIL .
## BITRO .
## CHEST1 0.14043742 0.14895639 0.1572993 1.640174e-01 1.711141e-01
## CHEST2 .
## ELBOW .
## WRIST .
## KNEE 0.20797327 0.22498473 0.2410221 2.544778e-01 2.671427e-01
## ANKLE .
## SHOUL 0.06541846 0.06652134 0.0679871 6.959762e-02 7.014580e-02
## CHESTG 0.21450838 0.21260465 0.2103983 2.086937e-01 2.064215e-01
## WAISTG 0.27118494 0.27453389 0.2775216 2.801490e-01 2.829491e-01
## NAVEL .
## HIP 0.29685995 0.29080552 0.2850313 2.800781e-01 2.743982e-01
## GLUTE 0.05250705 0.07736291 0.1001090 1.205083e-01 1.402598e-01
## BICEP 0.02557286 0.01752596 0.0121344 7.152257e-03 3.666317e-03
## FLOREA 0.47205727 0.48303098 0.4901859 4.955165e-01 5.004549e-01
## KNEEG 0.35215645 0.34003786 0.3293454 3.196632e-01 3.111957e-01
## CALF 0.38363626 0.38347074 0.3831024 3.831462e-01 3.822941e-01
## ANKLEG .
## WRISTG .
## AGE .
## HEIG 0.22507795 0.23257875 0.2393696 2.455731e-01 2.511669e-01
## GENFemale .
## GENMale .
##
## (Intercept) -1.051197e+02 -1.065973e+02 -1.079459e+02 -1.091727e+02
## BIAC .
## BIIL .
## BITRO .
## CHEST1 1.795396e-01 1.867392e-01 1.929831e-01 1.984873e-01
## CHEST2 1.581522e-02 2.428455e-02 3.151201e-02 3.825518e-02
## ELBOW .
## WRIST .
## KNEE 2.788393e-01 2.882019e-01 2.967711e-01 3.041736e-01
## ANKLE .
## SHOUL 6.942744e-02 6.990979e-02 7.029146e-02 7.054232e-02
## CHESTG 2.025835e-01 1.992986e-01 1.966067e-01 1.941643e-01
## WAISTG 2.854810e-01 2.873833e-01 2.890696e-01 2.905879e-01
## NAVEL .
## HIP 2.691038e-01 2.651210e-01 2.615800e-01 2.584056e-01
## GLUTE 1.582041e-01 1.737948e-01 1.878712e-01 2.006776e-01
## BICEP 3.483636e-03 2.352904e-03 1.414639e-03 9.680604e-04
## FLOREA 5.038872e-01 5.064868e-01 5.084397e-01 5.101158e-01
## KNEEG 3.041923e-01 2.975403e-01 2.913507e-01 2.858194e-01
## CALF 3.806873e-01 3.798884e-01 3.793862e-01 3.788442e-01
## ANKLEG .
## WRISTG .
## AGE .
## HEIG 2.561383e-01 2.606336e-01 2.647661e-01 2.685369e-01
## GENFemale .
## GENMale .
##

```

```

## (Intercept) -1.102908e+02 -1.113036e+02 -1.122236e+02 -113.06257563
## BIAC .
## BIIL .
## BITRO .
## CHEST1 2.033848e-01 2.085454e-01 2.134166e-01 0.21701270
## CHEST2 4.399041e-02 5.032446e-02 5.630112e-02 0.06035319
## ELBOW .
## WRIST .
## KNEE 3.109580e-01 3.180049e-01 3.244493e-01 0.32924891
## ANKLE .
## SHOUL 7.082476e-02 7.085264e-02 7.095666e-02 0.07125797
## CHESTG 1.920270e-01 1.898097e-01 1.876236e-01 0.18586131
## WAISTG 2.919977e-01 2.933326e-01 2.945766e-01 0.29569062
## NAVEL .
## HIP 2.554869e-01 2.525653e-01 2.498127e-01 0.24757745
## GLUTE 2.122822e-01 2.231489e-01 2.331392e-01 0.24186450
## BICEP 3.725445e-04 1.536296e-04 3.154590e-05 .
## FLOREA 5.115421e-01 5.128612e-01 5.140394e-01 0.51497403
## KNEEG 2.808763e-01 2.762013e-01 2.720856e-01 0.26871823
## CALF 3.785030e-01 3.779042e-01 3.771983e-01 0.37681950
## ANKLEG .
## WRISTG .
## AGE .
## HEIG 2.719895e-01 2.751004e-01 2.779063e-01 0.28050177
## GENFemale .
## GENMale .
##
## (Intercept) -1.137636e+02 -1.142462e+02 -1.146590e+02 -1.150491e+02
## BIAC .
## BIIL 1.143001e-03 1.238784e-02 2.182412e-02 3.042722e-02
## BITRO .
## CHEST1 2.227271e-01 2.299197e-01 2.368307e-01 2.423340e-01
## CHEST2 6.557390e-02 6.660477e-02 6.993956e-02 7.208928e-02
## ELBOW .
## WRIST .
## KNEE 3.369937e-01 3.446411e-01 3.541431e-01 3.601986e-01
## ANKLE .
## SHOUL 7.131361e-02 7.219754e-02 7.157558e-02 7.134037e-02
## CHESTG 1.841699e-01 1.842284e-01 1.839430e-01 1.834242e-01
## WAISTG 2.971487e-01 2.992969e-01 3.021246e-01 3.044397e-01
## NAVEL .
## HIP 2.453020e-01 2.447963e-01 2.438116e-01 2.430666e-01
## GLUTE 2.492497e-01 2.499711e-01 2.513286e-01 2.525756e-01
## BICEP .
## FLOREA 5.140644e-01 5.091249e-01 5.060238e-01 5.044952e-01
## KNEEG 2.652437e-01 2.627046e-01 2.586443e-01 2.561096e-01
## CALF 3.758102e-01 3.758288e-01 3.760172e-01 3.760044e-01
## ANKLEG .
## WRISTG .
## AGE -1.495133e-03 -6.734086e-03 -1.159751e-02 -1.586361e-02
## HEIG 2.824733e-01 2.823321e-01 2.823605e-01 2.825150e-01
## GENFemale .
## GENMale .
##
```

```

## (Intercept) -1.154102e+02 -1.157254e+02 -1.160344e+02 -1.162919e+02
## BIAC .
## BIIL 3.818234e-02 4.501806e-02 5.159561e-02 5.717776e-02
## BITRO .
## CHEST1 2.471994e-01 2.526132e-01 2.564283e-01 2.604473e-01
## CHEST2 7.364326e-02 7.661273e-02 7.772621e-02 8.056144e-02
## ELBOW .
## WRIST .
## KNEE 3.651502e-01 3.714576e-01 3.746384e-01 3.780976e-01
## ANKLE .
## SHOUL 7.127948e-02 7.070563e-02 7.088825e-02 7.026136e-02
## CHESTG 1.829691e-01 1.822757e-01 1.818092e-01 1.808949e-01
## WAISTG 3.064097e-01 3.084990e-01 3.099596e-01 3.120067e-01
## NAVEL .
## HIP 2.426112e-01 2.421018e-01 2.417462e-01 2.412233e-01
## GLUTE 2.535052e-01 2.544141e-01 2.552597e-01 2.564476e-01
## BICEP 1.536335e-05 1.128947e-04 5.072700e-05 3.198899e-05
## FLOREA 5.031530e-01 5.020066e-01 5.011514e-01 5.001832e-01
## KNEEG 2.539488e-01 2.511806e-01 2.496199e-01 2.471585e-01
## CALF 3.761152e-01 3.761653e-01 3.762594e-01 3.760581e-01
## ANKLEG .
## WRISTG .
## AGE -1.972227e-02 -2.334186e-02 -2.647125e-02 -2.952067e-02
## HEIG 2.826808e-01 2.827729e-01 2.829480e-01 2.829240e-01
## GENFemale .
## GENMale .
##
## (Intercept) -1.165506e+02 -1.167948e+02 -1.170115e+02 -1.172074e+02
## BIAC .
## BIIL 6.213827e-02 6.708181e-02 7.171298e-02 7.603594e-02
## BITRO .
## CHEST1 2.655922e-01 2.694011e-01 2.731232e-01 2.765589e-01
## CHEST2 8.279749e-02 8.449544e-02 8.607865e-02 8.760446e-02
## ELBOW .
## WRIST 2.114445e-02 3.358912e-02 4.601224e-02 5.803031e-02
## KNEE 3.810862e-01 3.822652e-01 3.837837e-01 3.852135e-01
## ANKLE .
## SHOUL 6.950061e-02 6.937789e-02 6.927072e-02 6.919393e-02
## CHESTG 1.800568e-01 1.793263e-01 1.786295e-01 1.779413e-01
## WAISTG 3.133134e-01 3.144003e-01 3.155106e-01 3.165865e-01
## NAVEL .
## HIP 2.410667e-01 2.407945e-01 2.404340e-01 2.399785e-01
## GLUTE 2.574225e-01 2.585828e-01 2.597334e-01 2.609373e-01
## BICEP 5.333608e-05 1.040132e-06 5.824771e-05 1.251958e-04
## FLOREA 4.984871e-01 4.965574e-01 4.942597e-01 4.919824e-01
## KNEEG 2.452700e-01 2.436363e-01 2.419463e-01 2.403622e-01
## CALF 3.756121e-01 3.752711e-01 3.748592e-01 3.743847e-01
## ANKLEG .
## WRISTG .
## AGE -3.228849e-02 -3.473768e-02 -3.702731e-02 -3.912569e-02
## HEIG 2.828438e-01 2.827669e-01 2.826219e-01 2.824619e-01
## GENFemale .
## GENMale .
##
```

```

## (Intercept) -117.39394649 -1.175477e+02 -1.177007e+02 -1.178347e+02
## BIAC . . . .
## BIIL 0.08002267 8.365425e-02 8.700422e-02 9.004872e-02
## BITRO . . . .
## CHEST1 0.27889872 2.821397e-01 2.841509e-01 2.862470e-01
## CHEST2 0.08830731 8.965035e-02 9.041096e-02 9.105074e-02
## ELBOW . . . .
## WRIST 0.06784048 7.863663e-02 8.728922e-02 9.535407e-02
## KNEE 0.38553832 3.874646e-01 3.878015e-01 3.885861e-01
## ANKLE . . . .
## SHOUL 0.06940698 6.923132e-02 6.936604e-02 6.942489e-02
## CHESTG 0.17745016 1.768537e-01 1.763828e-01 1.759626e-01
## WAISTG 0.31745788 3.185104e-01 3.192980e-01 3.200834e-01
## NAVEL . . . .
## HIP 0.23953774 2.390764e-01 2.386277e-01 2.382334e-01
## GLUTE 0.26209690 2.631128e-01 2.641782e-01 2.650669e-01
## BICEP . 1.713371e-04 6.729683e-05 6.052232e-05
## FLOREA 0.49018455 4.878915e-01 4.863807e-01 4.847946e-01
## KNEEG 0.23935947 2.378526e-01 2.369746e-01 2.360774e-01
## CALF 0.37402656 3.735072e-01 3.731346e-01 3.727602e-01
## ANKLEG . . . .
## WRISTG . . . .
## AGE -0.04095273 -4.277110e-02 -4.429978e-02 -4.573941e-02
## HEIG 0.28239227 2.821870e-01 2.821118e-01 2.820042e-01
## GENFemale . . . .
## GENMale . . . .
##
## (Intercept) -1.179554e+02 -1.180649e+02 -1.181646e+02 -1.182642e+02
## BIAC . . . .
## BIIL 9.281105e-02 9.533201e-02 9.764052e-02 9.972085e-02
## BITRO . . . .
## CHEST1 2.882380e-01 2.901092e-01 2.918635e-01 2.927922e-01
## CHEST2 9.175471e-02 9.250888e-02 9.328341e-02 9.351132e-02
## ELBOW . . . .
## WRIST 1.028830e-01 1.099277e-01 1.165113e-01 1.215260e-01
## KNEE 3.894328e-01 3.902405e-01 3.909761e-01 3.907562e-01
## ANKLE . . . .
## SHOUL 6.942813e-02 6.940384e-02 6.937296e-02 6.957014e-02
## CHESTG 1.755565e-01 1.751519e-01 1.747477e-01 1.744729e-01
## WAISTG 3.208303e-01 3.215259e-01 3.221662e-01 3.226348e-01
## NAVEL . . . .
## HIP 2.378631e-01 2.375086e-01 2.371692e-01 2.368748e-01
## GLUTE 2.658869e-01 2.666618e-01 2.673953e-01 2.681061e-01
## BICEP 8.454298e-05 1.181910e-04 1.550397e-04 9.208757e-05
## FLOREA 4.833116e-01 4.819671e-01 4.807529e-01 4.799293e-01
## KNEEG 2.351943e-01 2.343578e-01 2.335849e-01 2.332314e-01
## CALF 3.724040e-01 3.720572e-01 3.717158e-01 3.714859e-01
## ANKLEG . . . .
## WRISTG . . . .
## AGE -4.706322e-02 -4.827294e-02 -4.937609e-02 -5.030248e-02
## HEIG 2.818977e-01 2.817960e-01 2.816979e-01 2.816722e-01
## GENFemale . . . .
## GENMale . . . .
##
```

```

## (Intercept) -1.186985e+02 -1.190516e+02 -1.193510e+02 -1.196469e+02
## BIAC      .
## BIIL      1.014816e-01  1.023740e-01  1.019640e-01  1.016573e-01
## BITRO     .
## CHEST1    2.975019e-01  2.990631e-01  3.011213e-01  3.031031e-01
## CHEST2    9.447729e-02  9.443524e-02  9.574614e-02  9.689368e-02
## ELBOW     .
## WRIST     1.342803e-01  1.427128e-01  1.435033e-01  1.455337e-01
## KNEE      3.968800e-01  4.014324e-01  4.025709e-01  4.045171e-01
## ANKLE     .
## SHOUL     7.165004e-02  7.281601e-02  7.343692e-02  7.423840e-02
## CHESTG    1.740699e-01  1.735517e-01  1.724163e-01  1.710572e-01
## WAISTG   3.246144e-01  3.272732e-01  3.300130e-01  3.326231e-01
## NAVEL     .
## HIP       2.340480e-01  2.309785e-01  2.284129e-01  2.263051e-01
## GLUTE    2.668748e-01  2.661439e-01  2.650541e-01  2.635753e-01
## BICEP     1.272899e-03  2.308643e-03  3.736165e-03  4.616833e-03
## FLOREA   4.801412e-01  4.830442e-01  4.869849e-01  4.920788e-01
## KNEEG     2.283960e-01  2.260409e-01  2.245831e-01  2.232207e-01
## CALF     3.712822e-01  3.704897e-01  3.694659e-01  3.688019e-01
## ANKLEG    .
## WRISTG   .
## AGE       -5.148416e-02 -5.248181e-02 -5.353484e-02 -5.441367e-02
## HEIG     2.824937e-01  2.834558e-01  2.840973e-01  2.849417e-01
## GENFemale 1.137196e-01  2.117496e-01  3.074799e-01  3.992971e-01
## GENMale   .
##
## (Intercept) -1.199168e+02 -1.202000e+02 -1.204043e+02 -1.206356e+02
## BIAC      .
## BIIL      1.048352e-01  1.067853e-01  1.086737e-01  1.105050e-01
## BITRO     -1.207023e-02 -2.264429e-02 -3.165063e-02 -4.111268e-02
## CHEST1    3.030619e-01  3.045029e-01  3.056717e-01  3.075225e-01
## CHEST2    9.815910e-02  1.006286e-01  1.027891e-01  1.059150e-01
## ELBOW     .
## WRIST     1.474848e-01  1.490323e-01  1.491399e-01  1.506621e-01
## KNEE      4.067081e-01  4.113062e-01  4.136757e-01  4.173373e-01
## ANKLE     2.335595e-02  2.617752e-02  3.090168e-02  3.335372e-02
## SHOUL     7.525610e-02  7.556410e-02  7.603420e-02  7.627898e-02
## CHESTG    1.697577e-01  1.677791e-01  1.663560e-01  1.643581e-01
## WAISTG   3.348952e-01  3.373022e-01  3.386743e-01  3.403014e-01
## NAVEL     .
## HIP       2.261054e-01  2.266182e-01  2.272999e-01  2.282259e-01
## GLUTE    2.620454e-01  2.592386e-01  2.576017e-01  2.549904e-01
## BICEP     4.676326e-03  5.506424e-03  5.809975e-03  7.346497e-03
## FLOREA   4.972312e-01  5.046507e-01  5.091598e-01  5.143840e-01
## KNEEG     2.222928e-01  2.195132e-01  2.182162e-01  2.156113e-01
## CALF     3.684615e-01  3.689437e-01  3.689896e-01  3.696708e-01
## ANKLEG    .
## WRISTG   .
## AGE       -5.499380e-02 -5.556389e-02 -5.600270e-02 -5.643725e-02
## HEIG     2.860520e-01  2.874657e-01  2.883482e-01  2.895078e-01
## GENFemale 4.803913e-01  5.678104e-01  6.270053e-01  6.969035e-01
## GENMale   .
##

```

```

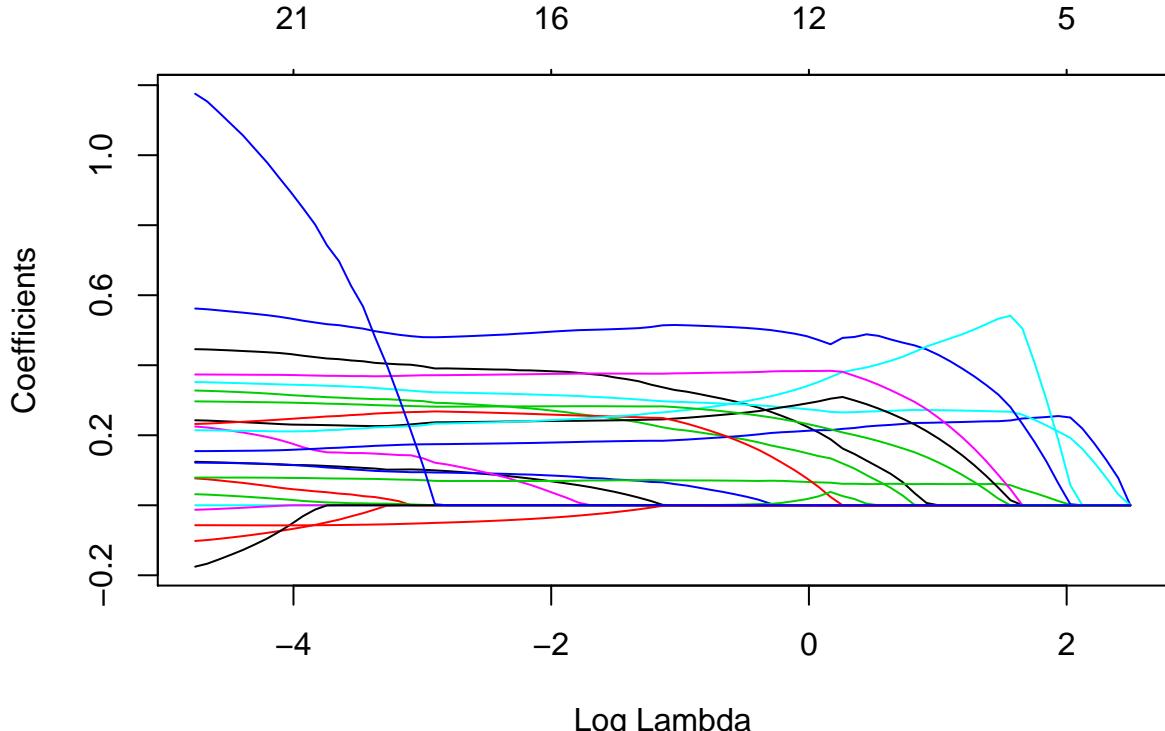
## (Intercept) -1.207937e+02 -120.98154185 -121.09911112 -1.211897e+02
## BIAC .
## BIIL 1.121641e-01 0.11377946 0.11475158 1.147856e-01
## BITRO -4.853972e-02 -0.05627551 -0.06238298 -6.830714e-02
## CHEST1 3.087250e-01 0.31083205 0.31257450 3.145122e-01
## CHEST2 1.079315e-01 0.11083514 0.11291320 1.155067e-01
## ELBOW .
## WRIST 1.511932e-01 0.15693052 0.16563831 1.770704e-01
## KNEE 4.192613e-01 0.42319848 0.42704995 4.317375e-01
## ANKLE 3.715537e-02 0.03985100 0.04319170 4.625026e-02
## SHOUL 7.667748e-02 0.07731069 0.07791455 7.805022e-02
## CHESTG 1.631184e-01 0.16125009 0.15983874 1.583198e-01
## WAISTG 3.412380e-01 0.34258157 0.34354395 3.444432e-01
## NAVEL .
## HIP 2.288940e-01 0.22958393 0.22990690 2.305125e-01
## GLUTE 2.536857e-01 0.25117088 0.24922142 2.468679e-01
## BICEP 8.050558e-03 0.01016659 0.01232423 1.541494e-02
## FLOREA 5.171649e-01 0.52216329 0.52720558 5.331667e-01
## KNEEG 2.146182e-01 0.21222774 0.21134086 2.108706e-01
## CALF 3.697186e-01 0.37055376 0.37121874 3.720773e-01
## ANKLEG .
## WRISTG -2.216236e-04 -0.01235943 -0.03098109 -5.416221e-02
## AGE -5.678484e-02 -0.05708288 -0.05731141 -5.748280e-02
## HEIG 2.901887e-01 0.29141474 0.29228235 2.931455e-01
## GENFemale 7.421887e-01 0.80218224 0.84856010 8.935220e-01
## GENMale .
##
## (Intercept) -1.212906e+02 -1.213961e+02 -121.48974033 -1.215839e+02
## BIAC .
## BIIL 1.160825e-01 1.176360e-01 0.11903146 1.203260e-01
## BITRO -7.371594e-02 -7.888516e-02 -0.08357331 -8.802088e-02
## CHEST1 3.169992e-01 3.192148e-01 0.32116530 3.227677e-01
## CHEST2 1.174076e-01 1.183897e-01 0.11931936 1.198703e-01
## ELBOW .
## WRIST 1.868297e-01 1.951237e-01 0.20173058 2.077323e-01
## KNEE 4.349941e-01 4.371252e-01 0.43894264 4.405801e-01
## ANKLE 5.031431e-02 5.524188e-02 0.05954187 6.388490e-02
## SHOUL 7.841647e-02 7.885938e-02 0.07898758 7.914153e-02
## CHESTG 1.571162e-01 1.563986e-01 0.15592983 1.555931e-01
## WAISTG 3.453356e-01 3.464178e-01 0.34742151 3.485225e-01
## NAVEL -1.438761e-03 -3.385048e-03 -0.00523013 -7.105805e-03
## HIP 2.320483e-01 2.338727e-01 0.23564573 2.373663e-01
## GLUTE 2.445984e-01 2.424354e-01 0.24045119 2.384820e-01
## BICEP 1.797908e-02 2.036715e-02 0.02269498 2.487885e-02
## FLOREA 5.385446e-01 5.429964e-01 0.54661970 5.501994e-01
## KNEEG 2.110222e-01 2.116378e-01 0.21217126 2.127405e-01
## CALF 3.724921e-01 3.726121e-01 0.37277857 3.729222e-01
## ANKLEG .
## WRISTG -7.556118e-02 -9.508485e-02 -0.11159545 -1.273717e-01
## AGE -5.746063e-02 -5.737966e-02 -0.05730794 -5.723089e-02
## HEIG 2.937306e-01 2.942204e-01 0.29469954 2.951745e-01
## GENFemale 9.358384e-01 9.790935e-01 1.01730061 1.055942e+00
## GENMale .
##

```

```

## (Intercept) -1.216636e+02 -121.74295454 -121.81857208 -1.218755e+02
## BIAC      .
## BIIL      1.214266e-01   0.12246540   0.12338457  1.242402e-01
## BITRO    -9.203550e-02  -0.09584186  -0.09935455 -1.019481e-01
## CHEST1    3.243522e-01   0.32568738   0.32683607  3.277323e-01
## CHEST2    1.206366e-01   0.12116788   0.12149878  1.217277e-01
## ELBOW     .
## WRIST     2.129237e-01   0.21779105   0.22216561  2.248165e-01
## KNEE      4.423255e-01   0.44379490   0.44511467  4.458599e-01
## ANKLE     6.733494e-02   0.07093389   0.07437509  7.677842e-02
## SHOUL     7.907880e-02   0.07910203   0.07913350  7.916909e-02
## CHESTG    1.552911e-01   0.15502139   0.15478934  1.546467e-01
## WAISTG    3.494457e-01   0.35040994   0.35138650  3.520331e-01
## NABEL     -8.723926e-03  -0.01032977  -0.01186971 -1.291920e-02
## HIP       2.389563e-01   0.24046644   0.24187623  2.428347e-01
## GLUTE     2.367092e-01   0.23497970   0.23331038  2.322389e-01
## BICEP     2.687992e-02   0.02875570   0.03047878  3.155107e-02
## FLOREA    5.534081e-01   0.55661587   0.55984042  5.619799e-01
## KNEEG     2.131080e-01   0.21356220   0.21403965  2.143610e-01
## CALF      3.731231e-01   0.37327623   0.37342152  3.735256e-01
## ANKLEG    .
## WRISTG    -1.411462e-01  -0.15444104  -0.16699076 -1.753507e-01
## AGE       -5.716997e-02  -0.05710245  -0.05703652 -5.701272e-02
## HEIG      2.956222e-01   0.29604606   0.29645385  2.967601e-01
## GENFemale 1.088764e+00   1.12167093   1.15354604  1.175456e+00
## GENMale   .
plot(modelo_lasso, xvar = "lambda")

```



podemos observar que a medida que aumenta el lambda tambien aumenta la cantidad de variables que pasan a valer a 0. # Aca

```

set.seed(2023) # Consultar, hacemos esto para poder realizar la interpretacion adecuadamente.

cv_lasso <- cv.glmnet(matriz_diseño_sin_intercept,entrenamiento$WEIG)

cv_lasso$lambda.min

## [1] 0.06025779

cv_lasso$lambda.1se

## [1] 0.3529311

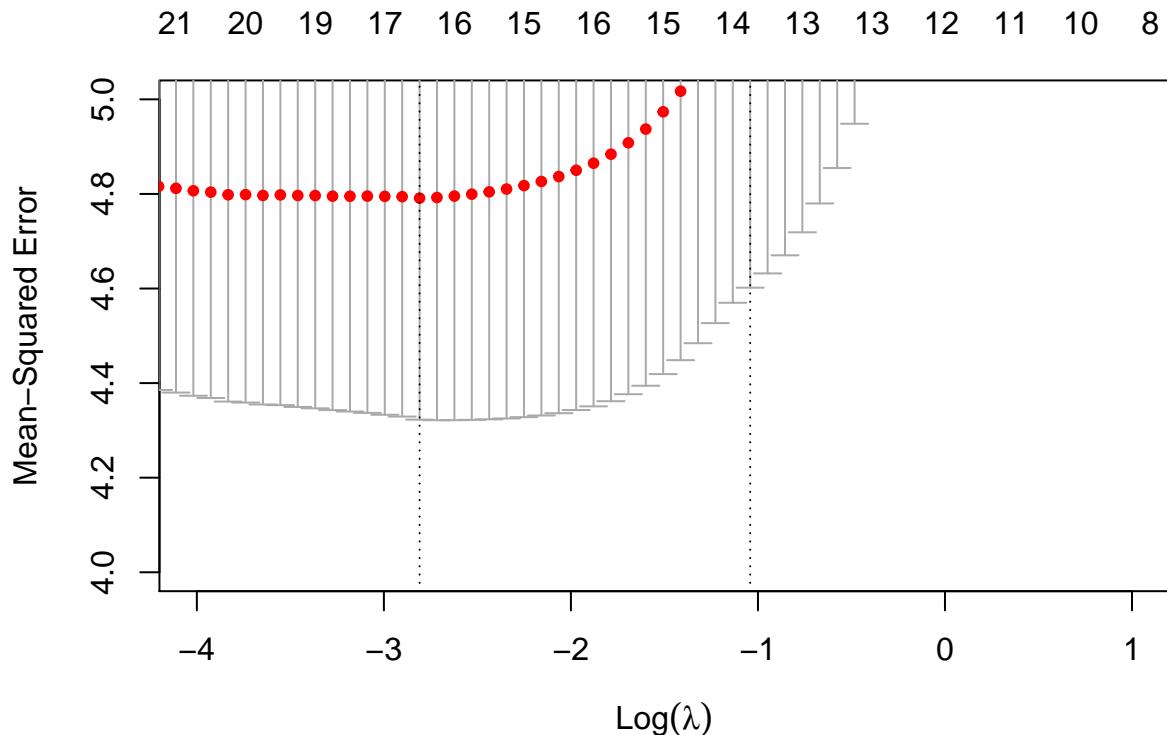
coef.glmnet(cv_lasso,s = "lambda.1se")

## 26 x 1 sparse Matrix of class "dgCMatrix"
##           lambda.1se
## (Intercept) -113.06257563
## BIAC          .
## BIIL          .
## BITRO         .
## CHEST1        0.21701270
## CHEST2        0.06035319
## ELBOW         .
## WRIST         .
## KNEE          0.32924891
## ANKLE         .
## SHOUL         0.07125797
## CHESTG        0.18586131
## WAISTG        0.29569062
## NAVEL         .
## HIP           0.24757745
## GLUTE         0.24186450
## BICEP          .
## FLOREA        0.51497403
## KNEEG         0.26871823
## CALF          0.37681950
## ANKLEG        .
## WRISTG        .
## AGE           .
## HEIG          0.28050177
## GENFemale     .
## GENMale       .

#predict(cv_lasso,newx = matriz_diseño_sin_intercept_test ,s = "lambda.1se")

plot(cv_lasso , xlim = c(-4,1) , ylim = c(4,5))

```



```
mean((testeo$WEIG-predict(cv_lasso,matriz_diseño_sin_intercept_test,s = "lambda.1se"))^2)
## [1] 4.317663
```

Las variables que se quedan este modelo no son las mas relevantes de modelo total si solo evaluamos la importancia de las variables por su p-valor.

Es Normal que nos de mejor pues estamos usando todas las