

Mínimos Cuadrados Ordinarios

Juan David Rincón Mora

2023

1. Introducción.

El objetivo principal de este cuaderno es estimar un modelo de interés económico mediante el método de **Mínimos Cuadrados Ordinarios (MCO, OLS)**. Se describirá detalladamente el proceso de construcción de matrices, la estimación de estadísticos y variables, y la interpretación de los resultados obtenidos.

Es importante señalar que este cuaderno se centra en el desarrollo de los diferentes estimadores y estadísticos necesarios para la estimación del modelo econométrico en términos de programación. No se profundizará en el desarrollo matemático detrás de los mismos, más allá de presentar las ecuaciones, fórmulas e interpretaciones correspondientes.

Por último, es pertinente mencionar que el contenido matemático, econométrico y estadístico presentado en este cuaderno se basa en gran medida en el enfoque y metodología propuestos por Judge, Hill, Griffiths, Lütkepohl y Lee (1988).

2. Modelo.

En este cuaderno se va a estimar el modelo **CAPM (Capital Asset Pricing Model)**, cuyo objetivo es explicar la rentabilidad de un activo financiero en función del riesgo de mercado que asume. Según Berk y DeMarzo (2020), el modelo matemático más sencillo del CAPM se expresa así:

$$r_i = r_f + \beta(E[R_{mkt}] - r_f)$$

Donde r_i representa la rentabilidad del activo a valorar, r_f es la tasa libre de riesgo, β es la sensibilidad de la rentabilidad del activo frente a cambios en la rentabilidad del portafolio de mercado, por lo tanto, representa una medida del riesgo sistémico, y $E[R_{mkt}]$ es el valor esperado de la rentabilidad del mercado, de donde se resta la tasa libre de riesgo, lo cual corresponde a la prima de riesgo del mercado.

Para estimar el modelo de forma econométrica, en primer lugar, se realiza una pequeña transformación a la expresión matemática. Se traslada el término de la tasa libre de riesgo r_f al lado izquierdo de la ecuación, de modo que la variable dependiente sea la prima de riesgo que genera el activo por sí mismo en relación con la tasa libre de riesgo. Esto da lugar a un nuevo parámetro α , que representa el intercepto y cuya interpretación es la prima de riesgo natural del activo. Además, se añade el término de perturbación ε_t , que recoge toda aquella información no observable que afecta el modelo. De esta forma, la ecuación del modelo econométrico queda como sigue:

$$(r_t^i - r_t^f) = \alpha + \beta(r_t^{mkt} - r_t^f) + \varepsilon_t$$

Es importante destacar que el modelo se construye a partir de datos tipo **serie de tiempo**, lo cual explica que cada variable esté subindexada a un momento t en particular. Los superíndices indican a qué tipo de rentabilidad corresponde cada componente según la notación inicial del modelo matemático.

2.1. Variables.

En este caso, se ha elegido a **Amazon** como el activo r^i a valorar, identificado por su ticker de cotización AMZN. Dado que esta acción cotiza principalmente en el índice **NASDAQ Composite**, se ha tomado este índice como la rentabilidad de mercado r^{mkt} . Para la tasa libre de riesgo r^f , se ha utilizado la tasa de mercado en periodicidad mensual de los **Bonos del Tesoro de Estados Unidos a 10 años**, representada por el indicador UST10.

Se han tomado los valores correspondientes a los periodos de **enero de 2015 hasta diciembre de 2022**. Es importante aclarar que en el caso de AMZN y NASDAQ, no se han considerado los precios directamente, sino la rentabilidad mensual, es decir, el cambio porcentual respecto al último precio de cada mes.

La base de datos utilizada para este ejercicio está disponible en el siguiente enlace: [*Amazon CAPM data*](#). A continuación se presentan los primeros valores de la base de datos:

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

# Importar datos.
archivo <- file.choose()
DATA <- as.data.frame(read_excel(archivo, sheet = 1))

# Datos (Primeros Valores).
head(DATA)
```

```
##      Date  AMZN  NASDAQ  UST10y  AMZN Rtrn  NASDAQ Rtrn  UST10 mv
## 1 2015-01-30 17.73 4635.24 0.01639 0.14239691 -0.021285671 0.001355679
## 2 2015-02-27 19.01 4963.53 0.01996 0.07219402 0.070824812 0.001648308
## 3 2015-03-31 18.61 4900.88 0.01927 -0.02104156 -0.012622065 0.001591823
## 4 2015-04-30 21.09 4941.42 0.02035 0.13326169 0.008271984 0.001680219
## 5 2015-05-29 21.46 5070.02 0.02128 0.01754386 0.026024908 0.001756269
## 6 2015-06-30 21.70 4986.87 0.02349 0.01118360 -0.016400330 0.001936736
```

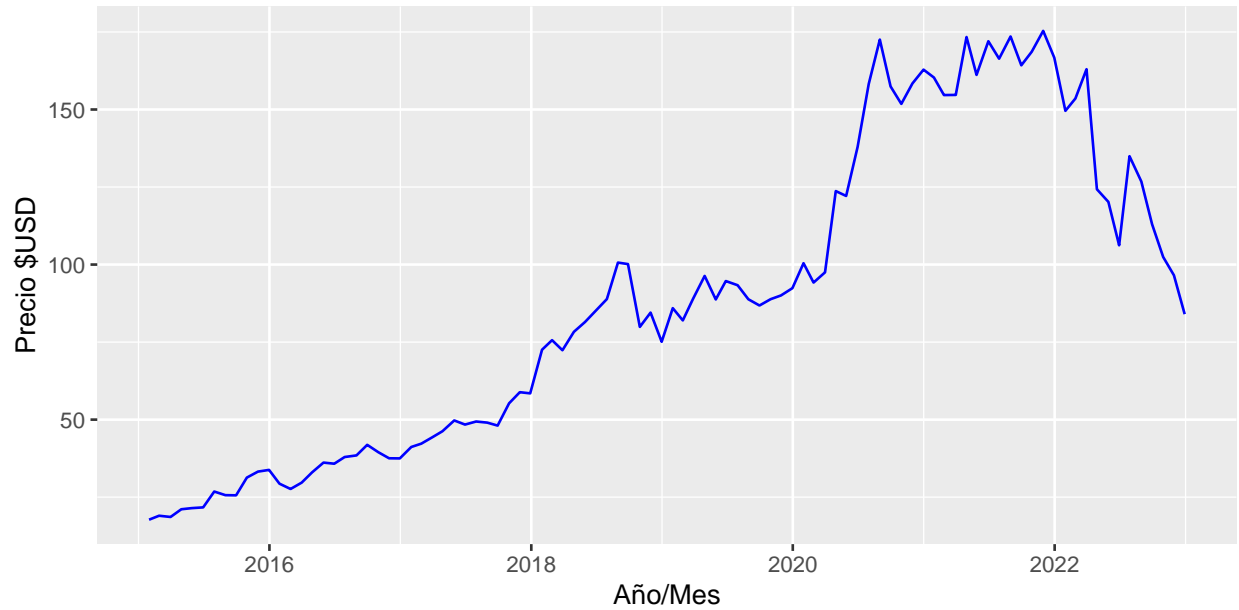
2.1.1. Representación gráfica.

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

# Gráfica Precios Amazon.
ggplot(data=DATA, aes(x = Date, y = AMZN)) +
  geom_line(color = "blue") +
  labs(x = "Año/Mes", y = "Precio $USD",
       title = "Último precio mensual Amazon (AMZN)",
       subtitle = "Enero 2015 - Diciembre 2022")
```

Último precio mensual Amazon (AMZN)

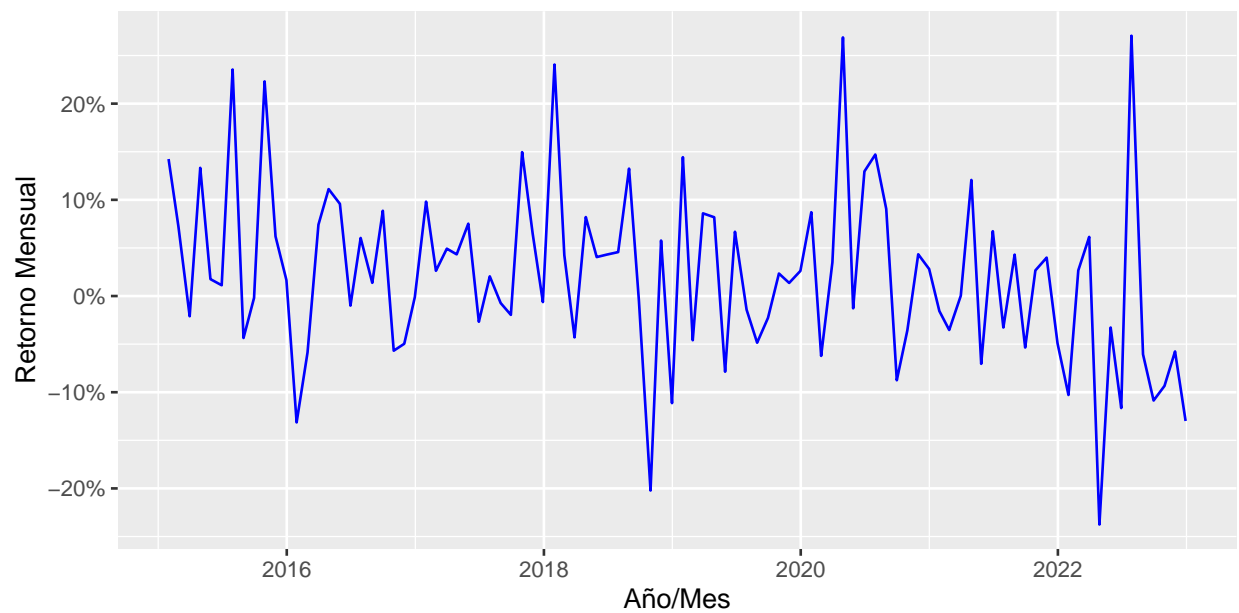
Enero 2015 – Diciembre 2022



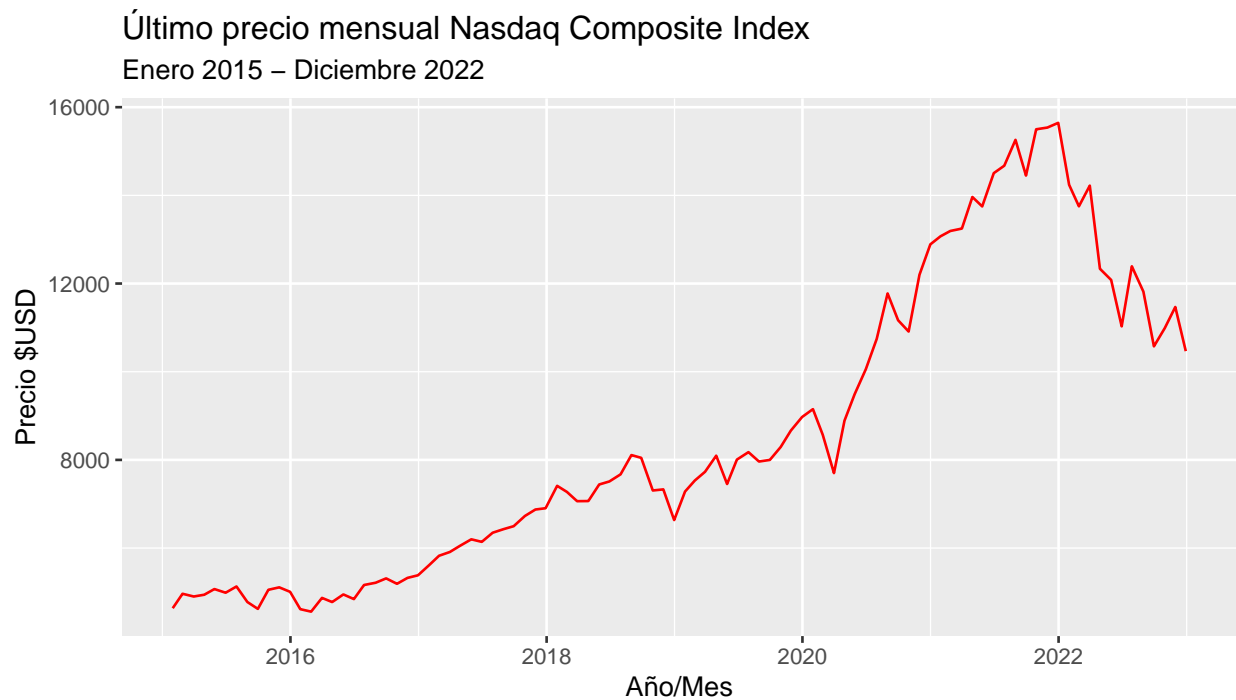
```
# Gráfica Retornos Amazon.
ggplot(data=DATA, aes(x = Date, y = `AMZN Rtrn`)) +
  geom_line(color = "blue") +
  labs(x = "Año/Mes", y = "Retorno Mensual",
       title = "Rentabilidad Mensual Amazon (AMZN)",
       subtitle = "Enero 2015 - Diciembre 2022") +
  scale_y_continuous(labels = scales::percent)
```

Rentabilidad Mensual Amazon (AMZN)

Enero 2015 – Diciembre 2022



```
# Gráfica Precios NASDAQ.
ggplot(data=DATA, aes(x = Date, y = NASDAQ)) +
  geom_line(color = "red") +
  labs(x = "Año/Mes", y = "Precio $USD",
       title = "Último precio mensual Nasdaq Composite Index",
       subtitle = "Enero 2015 - Diciembre 2022")
```



```
# Gráfica Retornos NASDAQ.
ggplot(data=DATA, aes(x = Date, y = `NASDAQ Rtrn`)) +
  geom_line(color = "red") +
  labs(x = "Año/Mes", y = "Retorno Mensual",
       title = "Rentabilidad Mensual Nasdaq Composite Index",
       subtitle = "Enero 2015 - Diciembre 2022") +
  scale_y_continuous(labels = scales::percent)
```

Rentabilidad Mensual Nasdaq Composite Index

Enero 2015 – Diciembre 2022



```
# Gráfica Bonos del Tesoro 10 años.  
ggplot(data=DATA, aes(x = Date, y = `UST10 mv`)) +  
  geom_line(color = "black") +  
  labs(x = "Año/Mes", y = "Retorno Mensual",  
       title = "Tasa mensual Bonos del Tesoro de Estados Unidos a 10 años",  
       subtitle = "Enero 2015 - Diciembre 2022") +  
  scale_y_continuous(labels = scales::percent)
```

Tasa mensual Bonos del Tesoro de Estados Unidos a 10 años

Enero 2015 – Diciembre 2022



2.1.2. Construcción de variables del modelo.

En primer lugar se extraen en vectores individuales cada variable:

```
# Amazon, Ri.
ri <- DATA$`AMZN Rtrn`

# Nasdaq, Rmkt.
rmkt <- DATA$`NASDAQ Rtrn`

# Bonos UST 10y, Rf.
rf <- DATA$`UST10 mv`
```

Luego se construyen las variables que efectivamente se van a utilizar en el modelo, es decir, la prima de riesgo propia del activo y la prima de riesgo del mercado:

```
# Prima del activo, Ri-Rf.
ri_rf <- ri - rf

# Prima del mercado, Rmkt-Rf.
rmkt_rf <- rmkt - rf
```

3. Estimación MCO, OLS.

Se realiza la estimación utilizando el método de Mínimos Cuadrados Ordinarios (MCO, OLS) del siguiente modelo en forma matricial:

$$Y = X\beta + e$$

Donde Y es una matriz de tamaño $T \times 1$ que contiene los valores de la variable dependiente, X es una matriz de diseño de tamaño $T \times K$ que contiene las variables independientes, β es una matriz de tamaño $K \times 1$ que contiene los parámetros que relacionan las variables explicativas con la variable explicada, y e es una matriz de tamaño $T \times 1$ que contiene las perturbaciones no observables.

3.1. Matrices de diseño.

De manera más específica, las matrices Y y X , de las cuales se dispone la información y los datos, se construyen de la siguiente manera:

$$Y = \begin{bmatrix} r_1^i - r_1^f \\ r_2^i - r_2^f \\ \vdots \\ r_T^i - r_T^f \end{bmatrix}_{T \times 1} \quad X = \begin{bmatrix} 1 & r_1^{mkt} - r_1^f \\ 1 & r_2^{mkt} - r_2^f \\ \vdots & \vdots \\ 1 & r_T^{mkt} - r_T^f \end{bmatrix}_{T \times 2}$$

```
# Matriz Y (Variable dependiente).
Y <- as.matrix(ri_rf)

# Matriz X (Variables independientes).
X <- cbind(1, rmkt_rf)

# Número de observaciones y variables.
N <- dim(X)[1]
K <- dim(X)[2]
```

3.2. Mínimos cuadrados ordinarios.

El estimador de Mínimos Cuadrados Ordinarios para los coeficientes (betas) del modelo, de forma matricial, es el siguiente:

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

```
# Función de estimación OLS.
Beta_OLS.f <- function(X,Y){

  # Fórmula.
  beta_gr <- solve(t(X)%*%X)%*%t(X)%*%Y

  # Resultado.
  return(beta_gr)
}

# Betas Estimados.
B_gr <- Beta_OLS.f(X,Y)
B_gr
```

```
##           [,1]
##      0.009957433
## rmkt_rf 1.268456732
```

Las interpretaciones de los betas estimados son las siguientes:

1. La prima de riesgo natural de Amazon, representada por el intercepto $\hat{\alpha}$, es de aproximadamente **0.9957%** mensual, manteniendo todo lo demás constante. Este coeficiente indica cuánta rentabilidad adicional genera el activo sin importar lo que suceda con la rentabilidad del mercado.
2. Dado que $\hat{\beta}$ es mayor a 1, esto indica que Amazon es más volátil que el mercado, es decir, reacciona de manera amplificada a los movimientos positivos o negativos del NASDAQ. En particular, el valor del coeficiente indica que ante un aumento de 1% (0.01 en la magnitud de los datos) en la prima de riesgo del mercado, la prima de riesgo de Amazon aumenta aproximadamente **1.2685%** (0.012685), manteniendo todo lo demás constante.

3.3. Estimación variable dependiente.

Para hallar la variable dependiente estimada se multiplica la matriz X de variables independientes por el vector $\hat{\beta}$ de coeficientes estimados por OLS.

$$\hat{Y} = X\hat{\beta}$$

```
# Variable dependiente estimada.
Y_gr <- X%*%B_gr
head(Y_gr)
```

```
##           [,1]
## [1,] -0.018762139
## [2,]  0.097704835
## [3,] -0.008072269
## [4,]  0.018318802
## [5,]  0.040741152
## [6,] -0.013302341
```

3.4. Estimación errores.

Para estimar las perturbaciones, que en principio no son observables, se restan los valores reales de la variable dependiente y los valores estimados por el modelo.

$$\hat{e} = Y - \hat{Y}$$

```
# Errores estimados.
e_gr <- Y - Y_gr
head(e_gr)
```

```
##           [,1]
## [1,]  0.15980337
## [2,] -0.02715912
## [3,] -0.01456111
## [4,]  0.11326267
## [5,] -0.02495356
## [6,]  0.02254920
```

3.5. Varianza estimada de los errores.

El estimador insesgado para la varianza de las perturbaciones del método OLS se calcula de la siguiente manera:

$$\hat{\sigma}^2 = \frac{\hat{e}' \cdot \hat{e}}{T - K}$$

```
# Función varianza estimada OLS.
sigma2_gr.f <- function(errors, T_obs, K_var){

  # Fórmula.
  sigma2_gr <- as.numeric((t(errors)%*%errors)/(T_obs-K_var))

  # Resultado.
  return(sigma2_gr)
}

# Varianza estimada.
sigma2_gr <- sigma2_gr.f(e_gr, N, K)
sigma2_gr
```

```
## [1] 0.003867785
```

```
# Desviación estándar estimada.
sigma_gr <- sqrt(sigma2_gr)
sigma_gr
```

```
## [1] 0.06219152
```


3.6. Matriz de Varianza-Covarianza de los betas.

La matriz de Varianza-Covarianza de los betas estimados por OLS se calcula de la siguiente manera:

$$\sum_{\hat{\beta}} = \text{VarCov}(\hat{\beta}) = \hat{\sigma}^2 (X'X)^{-1}$$

Para la cantidad de regresores de este modelo, la matriz contiene la siguiente información:

$$\sum_{\hat{\beta}} = \text{VarCov}(\hat{\beta}) = \begin{bmatrix} \text{Var}(\hat{\alpha}) & \text{Cov}(\hat{\alpha}, \hat{\beta}) \\ \text{Cov}(\hat{\beta}, \hat{\alpha}) & \text{Var}(\hat{\beta}) \end{bmatrix}_{2 \times 2}$$

```
# Matriz de Varianza-Covarianza.
varcov_beta_gr <- sigma2_gr*solve(t(X)%*%X)
varcov_beta_gr
```

```
##                                rmkt_rf
##          4.119061e-05 -0.0001122052
## rmkt_rf -1.122052e-04  0.0139704374
```

```
# Desviaciones estándar de los betas.
sd_beta_gr <- sqrt(diag(varcov_beta_gr))
sd_beta_gr
```

```
##                                rmkt_rf
## 0.006417991 0.118196605
```

3.7. Coeficiente de determinación.

Para medir la bondad de ajuste del modelo se utiliza el coeficiente de determinación R^2 , el cual representa el porcentaje de variabilidad de la variable dependiente que es explicada por la variabilidad de las variables explicativas del modelo. Existen dos maneras de calcular este estadístico:

$$R^2 = \frac{\text{SEC}}{\text{STC}} = \frac{\sum_{t=1}^T (\hat{Y}_t - \bar{Y})^2}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$$
$$R^2 = 1 - \frac{\text{SRC}}{\text{STC}} = 1 - \frac{\sum_{t=1}^T (\hat{e}_t)^2}{\sum_{t=1}^T (Y_t - \bar{Y})^2}$$

Donde SEC representa la suma explicada de cuadrados (SSR en inglés), SRC representa la suma de residuales al cuadrado (SSE en inglés), y STC representa la suma total de cuadrados (SST en inglés). Estos tres componentes se definen exactamente como se muestra en las fórmulas correspondientes.

```
# Función R2.
R2.f <- function(Y, Y_gr){

  # Suma explicada de cuadrados.
  ssr <- sum((Y_gr-mean(Y))^2)

  # Suma total de cuadrados.
  sst <- sum((Y-mean(Y))^2)
```

```

# Fórmula.
R2 <- as.numeric(ssr/sst)

# Resultado.
return(R2)
}

# Coeficiente de determinación R2.
R2 <- R2.f(Y, Y_gr)
R2

```

```
## [1] 0.5506059
```

Este resultado indica que aproximadamente el **55.1%** de la variabilidad de la prima de riesgo de Amazon es explicada por la variabilidad de la prima de riesgo del mercado NASDAQ.

3.7.1. Coeficiente de determinación ajustado.

Uno de los inconvenientes que tiene el R^2 como medida de bondad de ajuste es que su valor tiende a aumentar a medida que se añaden nuevas variables explicativas al modelo. Esto puede llevar a la interpretación errónea de que un modelo con más variables siempre es mejor. En realidad, algunas de estas variables adicionales pueden carecer de una relación económica coherente con el modelo o no ser estadísticamente significativas.

Para abordar este problema, se introduce el coeficiente de determinación ajustado R_{adj}^2 , que penaliza el aumento del R^2 a medida que se agregan más variables al modelo. Este coeficiente se define de la siguiente manera:

$$R_{adj}^2 = 1 - \frac{T-1}{T-k-1}(1 - R^2)$$

Donde R^2 es el coeficiente de determinación normal, T es el número de observaciones total, y k es el número de regresores sin el intercepto.

Cabe aclarar que en este caso particular, dado que el modelo solo tiene un regresor aparte del intercepto, no es necesario utilizar el coeficiente de determinación ajustado. Sin embargo, con el fin de proporcionar una comprensión más completa y con fines educativos, se presenta a continuación el resultado del cálculo del R_{adj}^2 .

```

# Función R2 ajustado.
R2adj.f <- function(R2, T_obs, K_var){

  # Fórmula.
  R2adj <- 1-((1-R2)*((T_obs-1)/(T_obs-(K_var-1)-1)))

  # Resultado.
  return(R2adj)
}

# Coeficiente de determinación R2 ajustado.
R2adj <- R2adj.f(R2, N, K)
R2adj

```

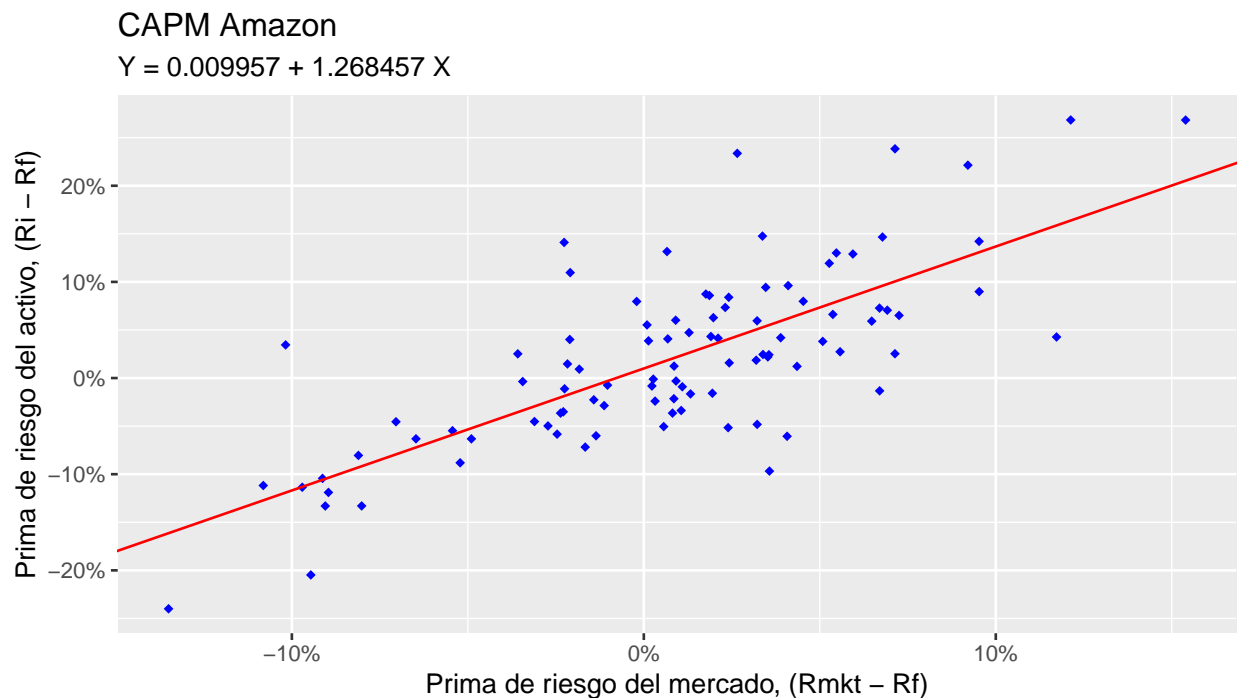
```
## [1] 0.5458251
```

Este resultado indica que, ajustado bajo el número de regresores, aproximadamente el **54.6%** de la variabilidad de la prima de riesgo de Amazon es explicada por la variabilidad de la prima de riesgo del mercado NASDAQ.

4. Uso del modelo.

```
# Data frame provisional.
df1 <- as.data.frame(cbind(ri_rf, rmkt_rf))

# Gráfica.
ggplot(df1, aes(x=rmkt_rf, y=ri_rf)) +
  geom_point(color="blue", pch=18) +
  geom_abline(color="red", intercept = B_gr[1], slope = B_gr[2]) +
  labs(x = "Prima de riesgo del mercado, (Rmkt - Rf)",
       y = "Prima de riesgo del activo, (Ri - Rf)",
       title = "CAPM Amazon",
       subtitle = paste("Y =", round(B_gr[1],6), "+", round(B_gr[2],6), "X")) +
  scale_y_continuous(labels = scales::percent) +
  scale_x_continuous(labels = scales::percent)
```



4.1. Pronóstico.

Una vez obtenidos los coeficientes de la regresión lineal, es posible utilizarlos para predecir y evaluar la rentabilidad futura de Amazon. Este proceso es sencillo y se logra mediante la sustitución de cualquier valor deseado de la prima de mercado en la forma funcional del modelo de regresión, o mediante la multiplicación de un vector transpuesto de información por la matriz de coeficientes estimados.

$$\hat{y}_t = \hat{\alpha} + x_t' \hat{\beta} \quad \text{ó} \quad \hat{y}_t = x_t' \hat{\beta}$$

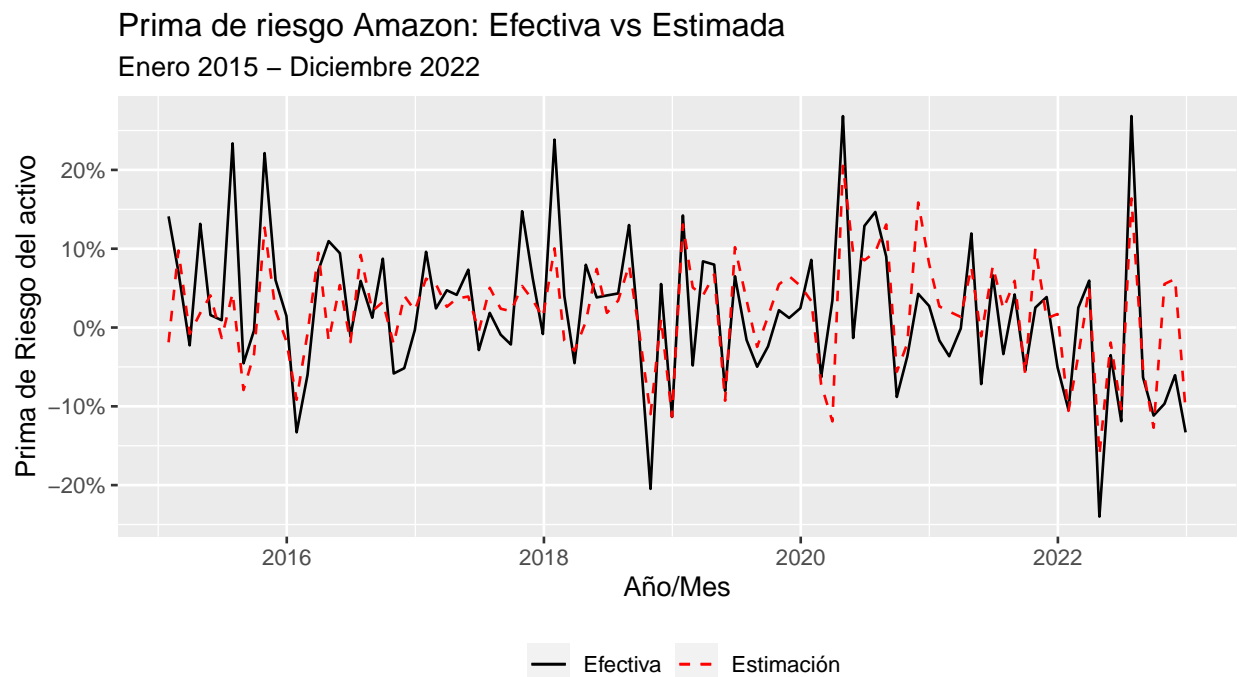
De esta manera, se puede obtener una estimación de la rentabilidad futura de Amazon utilizando los datos históricos y los patrones encontrados en el modelo de regresión lineal. Este enfoque permite a los inversores anticipar posibles resultados y tomar decisiones informadas sobre la inversión en acciones de Amazon en el futuro.

4.1.1. Pronóstico dentro de muestra.

A partir de las variables dependientes estimadas \hat{Y} , se pueden comparar con los valores efectivamente realizados, en este caso $y_t = r_t^i - r_t^f$.

```
# Data frame provisional.
df2 <- data.frame(Fecha = DATA$Date, Y_gr = Y_gr, ri_rf = ri_rf)

# Gráfica pronóstico dentro de muestra.
ggplot(df2, aes(x=Fecha)) +
  geom_line(aes(y=ri_rf, linetype="Efectiva", color="Efectiva")) +
  geom_line(aes(y=Y_gr, linetype="Estimación", color="Estimación")) +
  labs(title = "Prima de riesgo Amazon: Efectiva vs Estimada",
       subtitle = "Enero 2015 - Diciembre 2022",
       x = "Año/Mes", y = "Prima de Riesgo del activo",
       linetype = "", color = "") +
  scale_linetype_manual(values = c("Efectiva"="solid", "Estimación"="dashed")) +
  scale_color_manual(values = c("Efectiva"="black", "Estimación"="red")) +
  scale_y_continuous(labels = scales::percent) +
  theme(legend.position = "bottom")
```



Al analizar las series, se puede notar que en varios puntos la estimación se asemeja a la realización efectiva, en algunos puntos la estimación es muy cercana y en otros está alejada de la realización efectiva. Sin embargo, en algunos casos, la dirección de la estimación es completamente opuesta a la de la realización.

Para medir el error de estimación promedio dentro de la muestra, se calcula el **Error Absoluto Medio (EAM, MAE)**, el cual consiste en el promedio aritmético del valor absoluto de los errores estimados \hat{e} de la muestra de tamaño n que se quiera evaluar. No se toma el promedio de los residuales tal cual debido a que por propiedades de la estimación dicho promedio siempre será igual a cero. Por lo tanto, se utiliza el valor absoluto para evitar este problema y obtener una medida más precisa del error de la estimación. La

forma de calcular el MAE es la siguiente:

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} = \frac{\sum_{i=1}^n |\hat{e}_i|}{n}$$

```
# Error absoluto medio (Dentro de muestra).
```

```
mae <- sum(abs(e_gr))/N
```

```
mae
```

```
## [1] 0.04730879
```

El resultado obtenido indica que, considerando los datos de la muestra, **se espera que el modelo tenga un error promedio de aproximadamente el 4.73%**, ya sea positivo o negativo, al estimar la prima de riesgo de Amazon.

Para evaluar la dirección de la estimación, se va a calcular el porcentaje de coincidencias, es decir, se contarán aquellas estimaciones cuya dirección positiva o negativa coincida con la dirección de la realización de la variable. De esta manera, se podrá determinar con mayor precisión la capacidad del modelo para predecir la dirección adecuada de la prima de riesgo de Amazon.

```
# Vector de coincidencias (Dentro de muestra).
```

```
dMatch <- (Y > 0) == (Y_gr > 0)
```

```
# Porcentaje de aciertos (Dentro de muestra).
```

```
sum(dMatch)/N
```

```
## [1] 0.7604167
```

El resultado obtenido indica que el **76.04% de las estimaciones coinciden en dirección con la variable efectivamente realizada**, mientras que el 23.96% restante son estimaciones cuya dirección es contraria a la verdadera.

4.1.2. Pronóstico fuera de muestra.

Para evaluar el pronóstico del modelo fuera de muestra, teniendo en cuenta que la ventana de comparación posterior a la muestra es limitada, se realizarán dos tipos de análisis. Primero, se estimarán y compararán los datos anteriores a la muestra, específicamente de **enero de 2013 a diciembre de 2014**. Segundo, se compararán los datos más recientes posteriores a la muestra, que abarcan únicamente los meses de enero, febrero, marzo y abril de 2023. Es importante tener en cuenta que la muestra de datos finaliza en diciembre de 2022.

```
# Importar datos fuera de muestra.
```

```
DATA0 <- as.data.frame(read_excel(archivo, sheet = 2))
```

```
# Datos (Primeros Valores).
```

```
head(DATA0)
```

```
##      Date    AMZN    AMZN Rtrn  NASDAQ  NASDAQ Rtrn  UST10y    UST10 mv
## 1 2013-01-31 13.2750  0.058317057 3142.13  0.040609196 0.01985 0.001639305
## 2 2013-02-28 13.2135 -0.004632768 3160.19  0.005747712 0.01888 0.001559881
## 3 2013-03-28 13.3245  0.008400499 3267.52  0.033963173 0.01852 0.001530386
## 4 2013-04-30 12.6905 -0.047581523 3328.79  0.018751230 0.01675 0.001385231
## 5 2013-05-31 13.4600  0.060635909 3455.91  0.038188012 0.02164 0.001785691
## 6 2013-06-28 13.8845  0.031537890 3403.25 -0.015237640 0.02478 0.002041911
```

```

# Variables iniciales.
ri_0 <- DATA0$`AMZN Rtrn`
rmkt_0 <- DATA0$`NASDAQ Rtrn`
rf_0 <- DATA0$`UST10 mv`

# Variables del modelo.
ri_rf_0 <- ri_0 - rf_0
rmkt_rf_0 <- rmkt_0 - rf_0

# Matrices de diseño.
Y0 <- as.matrix(ri_rf_0)
X0 <- cbind(1, rmkt_rf_0)

# Pronóstico.
Y0_gr <- X0%*%B_gr

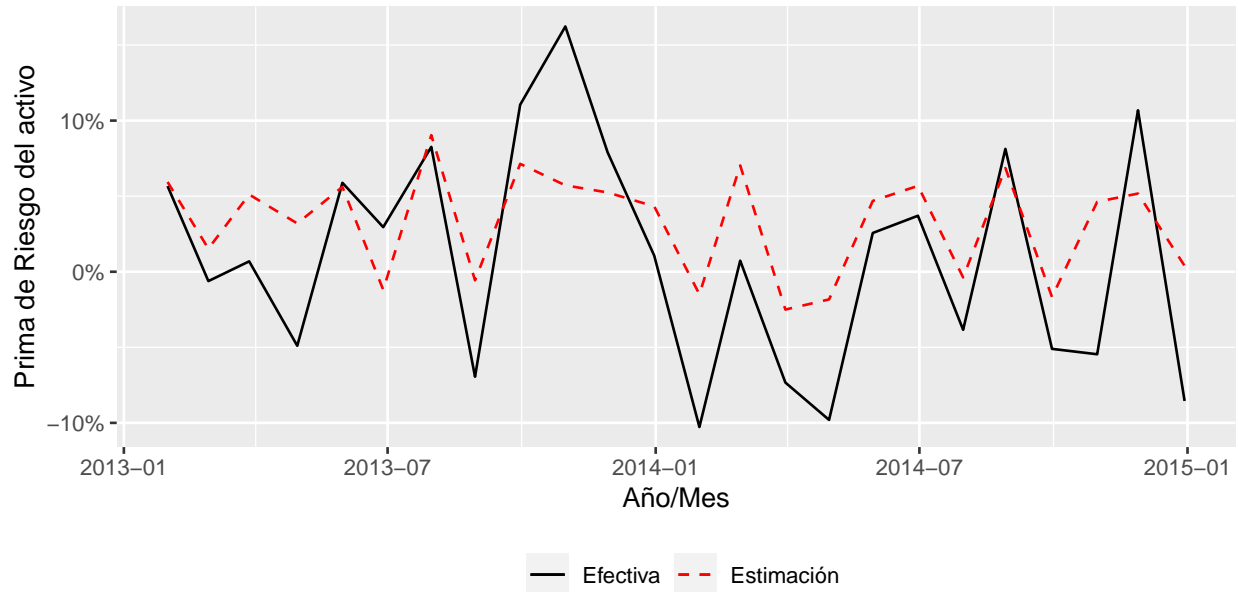
# Data frame provisional.
df3 <- data.frame(Fecha = DATA0$Date, Y0_gr = Y0_gr, ri_rf_0 = ri_rf_0)

# Gráfica pronóstico fuera de muestra.
ggplot(df3, aes(x=Fecha)) +
  geom_line(aes(y=ri_rf_0, linetype="Efectiva", color="Efectiva")) +
  geom_line(aes(y=Y0_gr, linetype="Estimación", color="Estimación")) +
  labs(title = "Prima de riesgo Amazon: Efectiva vs Estimada",
       subtitle = "Enero 2013 - Diciembre 2014",
       x = "Año/Mes", y = "Prima de Riesgo del activo",
       linetype = "", color = "") +
  scale_linetype_manual(values = c("Efectiva"="solid", "Estimación"="dashed")) +
  scale_color_manual(values = c("Efectiva"="black", "Estimación"="red")) +
  scale_y_continuous(labels = scales::percent) +
  theme(legend.position = "bottom")

```

Prima de riesgo Amazon: Efectiva vs Estimada

Enero 2013 – Diciembre 2014



```
# Error absoluto medio (Fuera de muestra).
```

```
e0_gr <- Y0 - Y0_gr
```

```
mae0 <- mean(abs(e0_gr))
```

```
mae0
```

```
## [1] 0.04647277
```

```
# Vector de coincidencias (Fuera de muestra).
```

```
dMatch0 <- (Y0 > 0) == (Y0_gr > 0)
```

```
# Porcentaje de aciertos (Fuera de muestra).
```

```
sum(dMatch0)/length(Y0)
```

```
## [1] 0.7916667
```

A partir de estos resultados se puede observar que la capacidad de pronóstico del modelo no difiere significativamente cuando se evalúa con datos fuera de la muestra. La serie de rentabilidad estimada de Amazon, graficada junto con la realización efectiva de la rentabilidad durante 2013 y 2014, muestra un comportamiento similar en la mayoría de los puntos. Específicamente, para este caso de comparación, se encontró que el error absoluto medio de estimación (positivo o negativo) fue de **4.65%**, y en cuanto a la estimación de la dirección, el porcentaje de aciertos fue de aproximadamente el **79.17%**.

A continuación, se evaluará la calidad de pronóstico del modelo utilizando los datos más recientes posteriores a la muestra. Para ello, se utilizarán los datos de rentabilidad mensual del NASDAQ y la tasa libre de riesgo correspondientes al período de enero a abril de 2023.

	Rentabilidad mensual NASDAQ	Tasa libre de riesgo mensual UST10Y
Enero	10.68%	0.2894%

	Rentabilidad mensual NASDAQ	Tasa libre de riesgo mensual UST10Y
Febrero	-1.11%	0.3206%
Marzo	6.69%	0.2866%
Abril	0.04%	0.2832%

Con esta información se construyen de manera individual los vectores con los datos correspondientes a cada mes y se realiza el pronóstico.

```
# Vectores de información.
xene <- c(1, 0.1068 - 0.002894)
xfeb <- c(1, -0.0111 - 0.003206)
xmar <- c(1, 0.0669 - 0.002866)
xabr <- c(1, 0.0004 - 0.002832)

# Pronóstico.
pronM <- data.frame(
  Mes = c("Enero", "Febrero", "Marzo", "Abril"),
  Estimación = round(c(xene**B_gr, xfeb**B_gr, xmar**B_gr, xabr**B_gr), 6)
)
pronM
```

```
##      Mes Estimación
## 1  Enero    0.141758
## 2 Febrero  -0.008189
## 3  Marzo    0.091182
## 4  Abril    0.006873
```

	Prima de riesgo efectiva de Amazon	Estimación prima de riesgo de Amazon	Error de estimación
Enero	22.48%	14.18%	8.3%
Febrero	-8.95%	-0.82%	-8.13%
Marzo	9.33%	9.12%	0.2%
Abril	2.09%	0.69%	1.40%

Los resultados indican que para estos cuatro meses el error absoluto medio de estimación es del **4.51%** y que en todos los pronósticos realizados, el modelo acertó en la dirección del cambio porcentual o rentabilidad.

5. Conclusiones.

En este documento se planteó el objetivo de construir un modelo CAPM para valorar y estimar la rentabilidad mensual de Amazon (AMZN), utilizando la rentabilidad mensual del mercado NASDAQ y la tasa libre de riesgo medida como la tasa de mercado convertida a mensual de los bonos del tesoro de Estados Unidos a 10 años. El ejercicio se realizó con fines pedagógicos para repasar paso a paso el proceso de estimación y su aplicación en términos de programación.

Respecto a los coeficientes estimados del modelo, se encontró que Amazon genera en promedio una prima de riesgo del 1%, manteniendo todo lo demás constante, y que su sensibilidad a los movimientos del mercado es amplificada, 1.27 veces los cambios en la prima de mercado del NASDAQ.

En cuanto a la calidad y utilidad del modelo en términos de pronóstico, se evaluó con datos dentro y fuera de muestra, y se concluye que el error de estimación se ubica entre 4.5% a 5%, y que el modelo predice con precisión la dirección del cambio porcentual de Amazon, con una tasa de acierto del 75% al 80%. Se considera que el modelo es muy útil para tomar decisiones de compra o venta de acciones de Amazon y se puede utilizar el valor pronosticado como un take profit o valor objetivo. Sin embargo, es importante tener en cuenta que cuando el pronóstico se encuentra alrededor de 0%, la varianza de las perturbaciones y el error promedio de estimación indican que el activo puede ir en cualquier dirección positiva o negativa.

Se destaca que este modelo está construido de manera sencilla y solo se basa en la información pasada de precios, lo que implica que puede no captar de manera completa otros aspectos que afecten la rentabilidad de Amazon. Además, en términos de utilización del modelo, se observa que la prima de riesgo mensual del NASDAQ y la tasa libre de riesgo, necesarias para pronosticar, no se conocen hasta finales de mes, justo cuando también se conoce el valor de rentabilidad mensual de Amazon. Por lo tanto, como complemento a este modelo, sería interesante construir otro que permita pronosticar mínimo a un mes adelante los precios del NASDAQ y la tasa libre de riesgo. Así se podría utilizar completamente el presente CAPM y pronosticar oportunamente la rentabilidad de la acción de Amazon para tomar decisiones de inversión que generen valor.

Referencias.

Berk, J. & DeMarzo, P. (2020). Optimal Portfolio Choice and the Capital Asset Pricing Model. *Corporate Finance* (pp. 393-443). Pearson Education.

Judge G., Hill R., Griffiths W., Lütkepohl H. & Lee T. (1988). Introduction to the Theory and Practice of Econometrics. John Wiley & Sons.