## UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO FACULTAD DE INGENIERÍA

### MAESTRÍA EN INGENIERÍA DE SISTEMAS

### Modelos ARMA & ARIMA

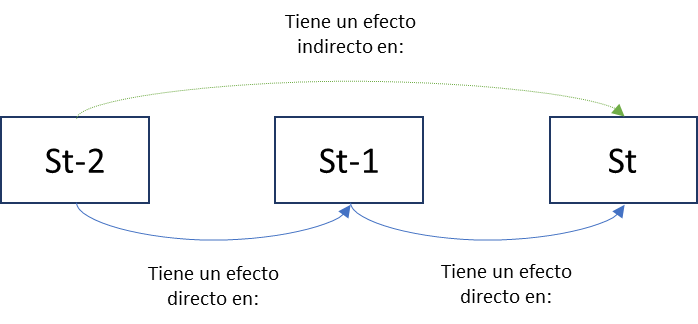
### Dr. Jair Camarena

### Sergio Ibarra Ramírez

## Series de Tiempo y Pronósticos

## Proceso autorregresivo (AR) y Proceso de Media Móvil (MA)

Los modelos AR y MA basan sus cálculos y parámetros en las relaciones qué existen entre los valores pasados de la variable estudiada y el valor al tiempo t así como un error aleatorio de cálculo asociado a pronóstico de cada modelo

Se puede asumir qué cada valor St-k tendrá un efecto directo en el siguiente valor y por lo tanto un efecto “indirecto” en el valor St de la siguiente manera:

Fuente: Elaboración propia Donde:

St: Medición al tiempo t

St-1: Medición al tiempo t -1 St-2: Medición al tiempo t -2

Para cuantificar dichos efectos y relaciones se utilizan las funciones de Autocorrelación ACF (qué cuantifica el efecto directo de cada medición en la siguiente) y la función de autocorrelación parcial PACF (qué cuantifica el efecto “indirecto” de una medición en el tiempo t-k con respecto a la medición en el tiempo t.

Serán estas relaciones (ACF & PACF) y la “forma” qué tengan sus gráficas a lo largo del tiempo las qué nos ayuden a determinar el mejor modelo a usar así como el número de parámetros a tomar en cuenta para el cálculo del pronóstico de la medición St

## Series estacionarias

Como preliminar, definimos un concepto importante, el de una serie estacionaria. Para que un ACF tenga sentido, la serie debe ser una serie débilmente estacionaria. Esto significa que la autocorrelación para cualquier retraso en particular es la misma independientemente de dónde nos encontremos en el tiempo.

Se dice que una serie es (débilmente) estacionaria si cumple las siguientes propiedades: La media E(Xt) es constante a lo largo de la serie .

La varianza Xt es constante a lo largo de la serie.

La covarianza (y también la correlación) entre Xt y Xt-h es la misma para todos en cada retraso h = 1, 2, 3, etc.

## Autocorrelation Function (ACF)

Fuentes: Olivas-Vasques. Apuntes sobre series de tiempo UNAM [1.2 Sample ACF and Properties of AR(1) Model | STAT 510 (psu.edu)](https://online.stat.psu.edu/stat510/lesson/1/1.2)

Es posible calcular la fuerza o intensidad de la dependencia de las variables aleatorias dentro de un proceso estocástico, ello mediante el uso de las autocovarianzas. Cuando las covarianzas son normalizadas respecto de la varianza, el resultado es un término que es independiente de las unidades de medida aplicada, y se conoce como la función de autocorrelación.

Denotemos el valor de una serie temporal en el tiempo como Xt . El ACF de la serie da correlaciones entre Xt y Xt-h para h = 1, 2, 3, etc. Teóricamente, la autocorrelación entre Xt y Xt-h es igual a:



El denominador en la segunda fórmula ocurre porque la desviación estándar de una serie estacionaria es la misma en todo momento.

La última propiedad de una serie débilmente estacionaria dice que el valor teórico de la autocorrelación de un retraso particular es el mismo en toda la serie. Una propiedad interesante de una serie estacionaria es que teóricamente tiene la misma estructura hacia adelante que hacia atrás.

### Notas sobre covarianza

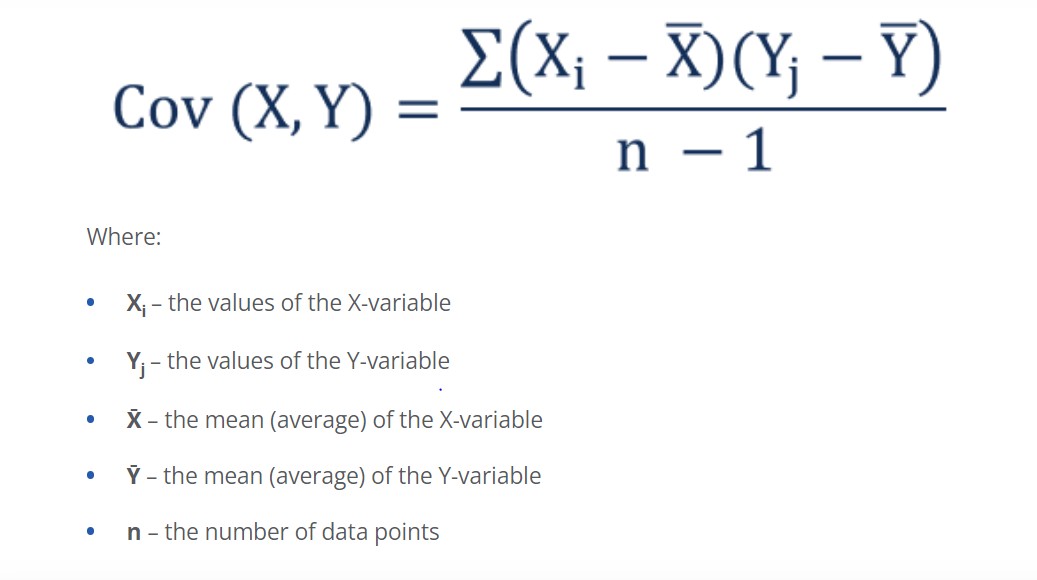
[Covariance - Definition, Formula, and Practical Example (corporatefinanceinstitute.com)](https://corporatefinanceinstitute.com/resources/knowledge/finance/covariance/)

La covarianza es una medida de la relación entre dos variables aleatorias. La métrica evalúa cuánto, en qué medida, las variables cambian juntas. En otras palabras, es esencialmente una medida de la varianza entre dos variables. Sin embargo, la métrica no evalúa la dependencia entre variables.

A diferencia del coeficiente de correlación, la covarianza se mide en unidades. Las unidades se calculan multiplicando las unidades de las dos variables. La varianza puede tomar cualquier valor positivo o negativo. Los valores se interpretan de la siguiente manera:

* Covarianza positiva: Indica que dos variables tienden a moverse en la misma dirección.
* Covarianza negativa: revela que dos variables tienden a moverse en direcciones inversas.

Para calcular la covarianza se usa la ecuación:

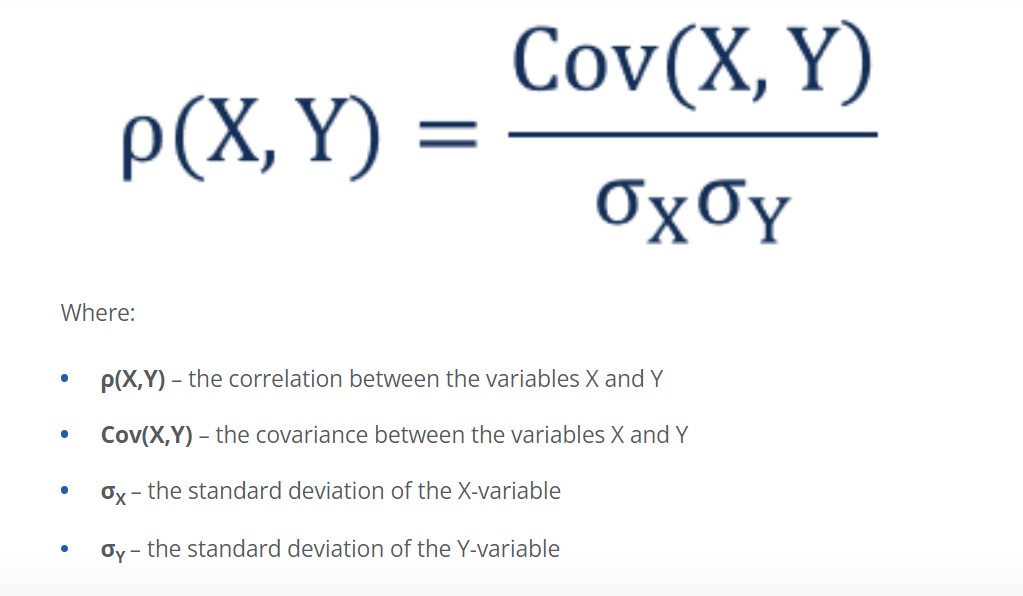


La covarianza mide la variación total de dos variables aleatorias de sus valores esperados. Usando la covarianza, solo podemos medir la dirección de la relación (si las variables tienden a moverse en

tándem o muestran una relación inversa). Sin embargo, no indica la fuerza de la relación, ni la dependencia entre las variables.

Por otro lado, la correlación mide la fuerza de la relación entre las variables. La correlación es la medida escalada de la covarianza. Es adimensional. En otras palabras, el coeficiente de correlación es siempre un valor puro y no se mide en ninguna unidad.

La relación entre los dos conceptos se puede expresar mediante la siguiente fórmula:



La función de autocorrelación simple tiene las siguientes propiedades:

## Partial Autocorrelation Function (PACF)

[**2.2 Partial Autocorrelation Function (PACF) | STAT 510 (psu.edu)**](https://online.stat.psu.edu/stat510/lesson/2/2.2)[**LinkClick.aspx (gobierno.pr)**](http://www.estadisticas.gobierno.pr/iepr/LinkClick.aspx?fileticket=4_BxecUaZmg%3D)

La autocorrelación parcial mide la correlación entre dos variables separadas por k periodos cuando no se considera la dependencia creada por los retardos intermedios existentes entre ambas.



Identification of an AR model order is often best done with the PACF

Tanto ACF como PACF comienzan con un retraso de 0, que es la correlación de la serie temporal consigo misma y, por lo tanto, da como resultado una correlación de 1.

La diferencia entre ACF y PACF es la inclusión o exclusión de correlaciones indirectas en el cálculo.

Además, puede ver un área azul en los gráficos ACF y PACF. Esta área azul representa el intervalo de confianza del 95 % y es un indicador del umbral de significación. Eso significa que cualquier cosa dentro del área azul es estadísticamente cercana a cero y cualquier cosa fuera del área azul es estadísticamente distinta de cero.

## Proceso autorregresivo (AR)

El modelo autorregresivo representa el valor actual de series de tiempo como combinación de uno o más valores anteriores de la serie, mostrando la dependencia de un valor con los valores anteriores más cercanos (Dickey y Fuller,1979).

El modelo autorregresivo es de orden p, el cual determina cuántos valores previos deben incluirse en la ecuación para estimar el valor actual (Dickey et al, 1979). Para el caso donde p=1, se tiene un AR (1) descrito por la ecuación:

𝑆𝑡 = Θ0 + Θ1 \* 𝑆𝑡 − 1 + ϵ

Donde:

St: Valor a ser estimado

St-1: Valor en el tiempo anterior t-1 conocido

Θ𝑝: : 𝑃𝑎𝑟á𝑚𝑒𝑡𝑟𝑜𝑠 𝑑𝑒𝑙 𝑚𝑜𝑑𝑒𝑙𝑜 𝐴𝑅 𝑎 𝑒𝑠𝑡𝑖𝑚𝑎𝑟

ϵ: es el error conformado por variables aleatorias no correlacionadas , las cuales tienen media cero y varianza constante, al que se le conoce como ruido blanco (2010, Gujarati)

Cuando se tiene un proceso autorregresivo (AR) de p orden, el modelo AR (p) tendrá la forma:

𝑆𝑡 = Θ0 + Θ1 \* 𝑆𝑡 − 1 + Θ2 \* 𝑆𝑡 − 2 + ... Θ𝑝 \* 𝑆𝑡 − 𝑝 + ϵ

Donde:

St: Valor a ser estimado

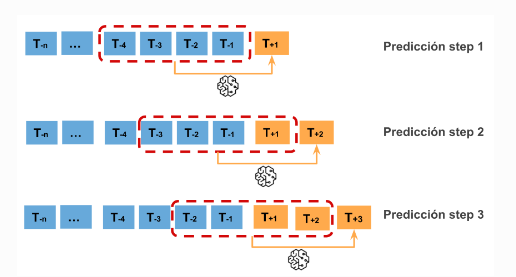
St-p: Valor en el tiempo anterior t-p conocido

Θ𝑝: 𝑃𝑎𝑟á𝑚𝑒𝑡𝑟𝑜𝑠 𝑑𝑒𝑙 𝑚𝑜𝑑𝑒𝑙𝑜 𝐴𝑅 𝑎 𝑒𝑠𝑡𝑖𝑚𝑎𝑟

ϵ: es el error conformado por variables aleatorias no correlacionadas , las cuales tienen media cero y varianza constante, al que se le conoce como ruido blanco (2010, Gujarati)

A continuación se presenta un diagrama del proceso de predicción recursivo para predecir 3 *steps*

a futuro utilizando los últimos 4 *lags* de la serie como predictores.

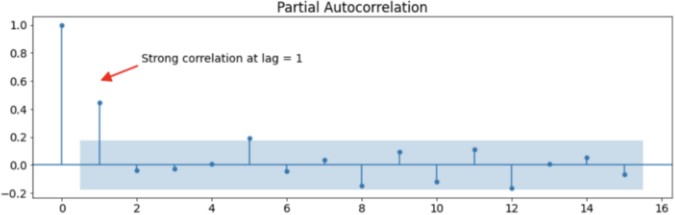


Fuente: Cienciadedatos.net ([Skforecast: forecasting series temporales con python y scikitlearn](https://www.cienciadedatos.net/documentos/py27-forecasting-series-temporales-python-scikitlearn.html) [(cienciadedatos.net)](https://www.cienciadedatos.net/documentos/py27-forecasting-series-temporales-python-scikitlearn.html)

Por lo que es posible observar que un proceso de Autoregresión (AR) es una combinación lineal de los términos de los valores St-p que impactan significativamente el valor de St

Para averiguar el orden de un modelo AR, se debe mirar la gráfica del PACF y observar qué correlaciones son estadísticamente significativas

De manera general se puede decir qué un procesos autorregresivo de orden p escrito como AR(p) es una regresión multilineal en el qué el valor qué se desea pronosticar es el valor de la serie de tiempo al tiempo t=t como una una función lineal de los valores pasados de la variable en estudio



Fuente: [Interpreting ACF and PACF Plots for Time Series Forecasting | by Leonie Monigatti |](https://towardsdatascience.com/interpreting-acf-and-pacf-plots-for-time-series-forecasting-af0d6db4061c) [Towards Data Science](https://towardsdatascience.com/interpreting-acf-and-pacf-plots-for-time-series-forecasting-af0d6db4061c)

## Proceso de media móvil (MA)

Una serie temporal además de estar influenciada por el/los valores en los tiempos anteriores t-p (conocidos) está influenciado por elementos/errores aleatorios, los cuáles inciden el valor de la variable dependiente en periodos posteriores al t. El modelo de media móvil (MA) asume que el valor actual (s\_t) depende de los términos de error, incluido el error actual (𝜖\_t, 𝜖\_(t-1),…). Debido a que los términos de error son aleatorios, no existe una relación lineal entre el valor actual y los términos de error.

La media móvil de primer orden o MA (1), es una serie de tiempo dada por:

𝑆𝑡 = λ + β0 \* ϵ 𝑡 + β1 \* ϵ 𝑡 − 1

Donde:

St: Valor a ser estimado

ϵt-1: Error obtenido en el pronóstico t-1

λ: Constante

β1: 𝑃𝑎𝑟á𝑚𝑒𝑡𝑟𝑜 𝑎 𝑒𝑠𝑡𝑖𝑚𝑎𝑟

#### Propiedades teóricas de una serie temporal con un modelo MA(1)

La media es E(xt)= μ

2

La varianza es Var(xt)= σ

β

2

\* (1 + β )

1

La función de autocorrelación (ACF) es: ρ1 = β

1

2

÷ 1 + β

1

Dónde:

ρt: Retraso de la autocorrelación en el tiempo t

([2.1 Moving Average Models (MA models) | STAT 510 (psu.edu)](https://online.stat.psu.edu/stat510/lesson/2/2.1#%3A~%3Atext%3DA%20moving%20average%20term%20in%2C0%20and%20the%20same%20variance)

Cuando se tiene un proceso de media móvil (MA)de p orden, el modelo MA (p) tendrá la forma:

𝑆𝑡 = λ + β0 \* ϵ 𝑡 + β1 \* ϵ 𝑡 − 1 + β2 \* ϵ 𝑡 − 2 + ... β𝑝 \* ϵ 𝑡 − 𝑝

Donde:

St: Valor a ser estimado

St-p: Valor en el tiempo anterior t-p conocido

λ: Constante

β𝑝: 𝑃𝑎𝑟á𝑚𝑒𝑡𝑟𝑜𝑠 𝑎 𝑒𝑠𝑡𝑖𝑚𝑎𝑟

#### Propiedades teóricas de una serie temporal con un modelo MA(2

La media es E(xt)= μ

2

La varianza es Var(xt)= σ

β

2

\* (1 + β

1

2

+ β )

2

La función de autocorrelación (ACF) es: ρ1 = (β

1

2

+ β1β2 ) ÷ (1 + β

1

2

+ β )

2

ρ2 = β

2

2

÷ (1 + β

1

2

+ β )

2

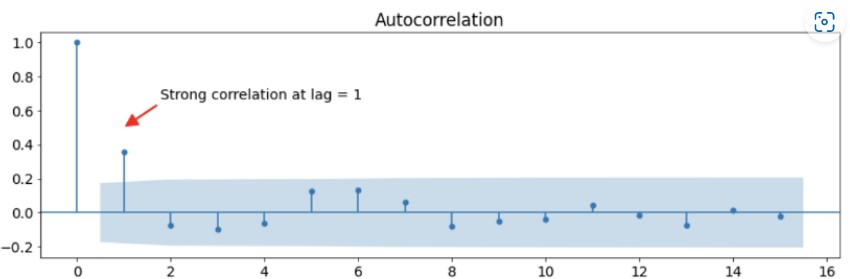
Dónde:

ρt: Retraso de la autocorrelación en el tiempo t

([2.1 Moving Average Models (MA models) | STAT 510 (psu.edu)](https://online.stat.psu.edu/stat510/lesson/2/2.1#%3A~%3Atext%3DA%20moving%20average%20term%20in%2C0%20and%20the%20same%20variance)

Por lo que es posible observar que un proceso de media móvil (MA) es una combinación lineal de los términos de error de los valores St-p que están involucrados en el modelo.

Para averiguar el orden de un modelo MA, se debe mirar la gráfica del ACF y observar qué correlaciones son estadísticamente significativas



Fuente: [Interpreting ACF and PACF Plots for Time Series Forecasting | by Leonie Monigatti |](https://towardsdatascience.com/interpreting-acf-and-pacf-plots-for-time-series-forecasting-af0d6db4061c) [Towards Data Science](https://towardsdatascience.com/interpreting-acf-and-pacf-plots-for-time-series-forecasting-af0d6db4061c)

## Recomendaciones de modelos a usar con base en los resultados de las gráficas de ACF y PACF para una serie de tiempo con t valores pasados

A continuación se presenta una tabla resumen de los métodos de pronóstico recomendados en series de tiempo con base en el resultado del análisis de sus respectivas gráficas de ACF y PACF

|  |  |  |  |
| --- | --- | --- | --- |
|  | AR(p) | MA(q) | ARMA (p,q) |
| ACF | Decaimiento geométrico | Valor significante en ciertos valores de lag p y posterior corte después de lag p | Decaimiento geométrico |
| PACF | Valor significante en ciertos valores de lag p y posterior corte después de lag p | Decaimiento geométrico | Decaimiento geométrico |

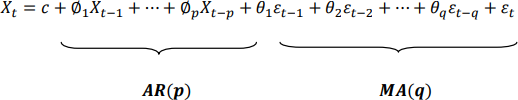
Fuentes:

S. Ali, “Reading the ACF and PACF Plots — The Missing Manual / Cheatsheet”. linkedin.com. <https://www.linkedin.com/pulse/reading-acf-pacf-plots-missing-manual-cheatsheet-saqib-ali/>

“Arauto”, “How to choose the parameters for the model”. arauto.readthedocs.io. <https://arauto.readthedocs.io/en/latest/how_to_choose_terms.html>

## Proceso Autoregresivo de media móvil (ARMA)

Es muy probable que una serie de tiempo, Xt, tenga características de AR y de MA a la vez y, por consiguiente, sea ARMA. Así, Xt sigue un proceso ARMA (p,q) , en este proceso habrá p términos autorregresivos y q términos de media móvil.



donde ϵt es un proceso de ruido blanco, y Θ𝑝 son los parámetros del modelo

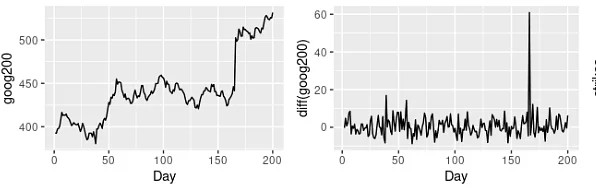
## Proceso Autoregresivo de media móvil Integrado (ARIMA)

Fuente:[How to Build ARIMA Model in Python for time series forecasting? (projectpro.io)](https://www.projectpro.io/article/how-to-build-arima-model-in-python/544#%3A~%3Atext%3DModel%20in%20Python%3F-%2CARIMA%20Model%2D%20Complete%20Guide%20to%20Time%20Series%20Forecasting%20in%20Python%2Cdata%20to%20predict%20future%20values)

Los modelos autorregresivos son conceptualmente similares a la regresión lineal, las suposiciones hechas por este último también son válidas aquí. Los datos de series de tiempo deben ser estacionarios para eliminar cualquier correlación y colinealidad obvia con los datos anteriores. En los datos de series de tiempo estacionarios, las propiedades o el valor de una observación de muestra no dependen de la marca de tiempo en la que se observa. Por ejemplo, dado un conjunto de datos hipotéticos de la población anual de un área, si se observa que la población aumenta al doble cada año o aumenta en una cantidad fija, entonces estos datos no son estacionarios.

Cualquier observación dada depende en gran medida del año, ya que el valor de la población dependería de qué tan lejos esté de un año pasado arbitrario. Esta dependencia puede inducir un sesgo incorrecto al entrenar un modelo con datos de series temporales.

Para eliminar esta correlación, ARIMA utiliza la diferenciación para hacer que los datos sean estacionarios. La diferenciación, en su forma más simple, implica tomar la diferencia de dos puntos de datos adyacentes.



Por ejemplo, el gráfico de la izquierda arriba muestra el precio de las acciones de Google durante 200 días. Mientras que el gráfico de la derecha es la versión diferenciada del primer gráfico, lo que significa que muestra el cambio en las acciones de Google de 200 días. Hay un patrón observable en el primer gráfico, y estas tendencias son un signo de datos de series de tiempo no estacionarios. Sin embargo, en la segunda figura no se observa ninguna tendencia ni estacionalidad, ni una varianza creciente. Así, podemos decir que la versión diferenciada es estacionaria.

Este cambio puede ser modelado como:



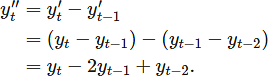


Donde B se denota como el operador “retroceso” definido como:



Sin embargo, la diferenciación para crear datos estacionarios puede no ser siempre tan sencilla. Múltiples iteraciones de diferenciación pueden ayudar más hasta cierto punto si es necesario. Diferenciar los datos d veces crea datos diferenciados de orden d.

Si d=2,



O,



Vemos que se establece aquí una generalidad. Por lo tanto, una serie diferenciada de orden d se definiría como:



El modelo final ARIMA se vería así:



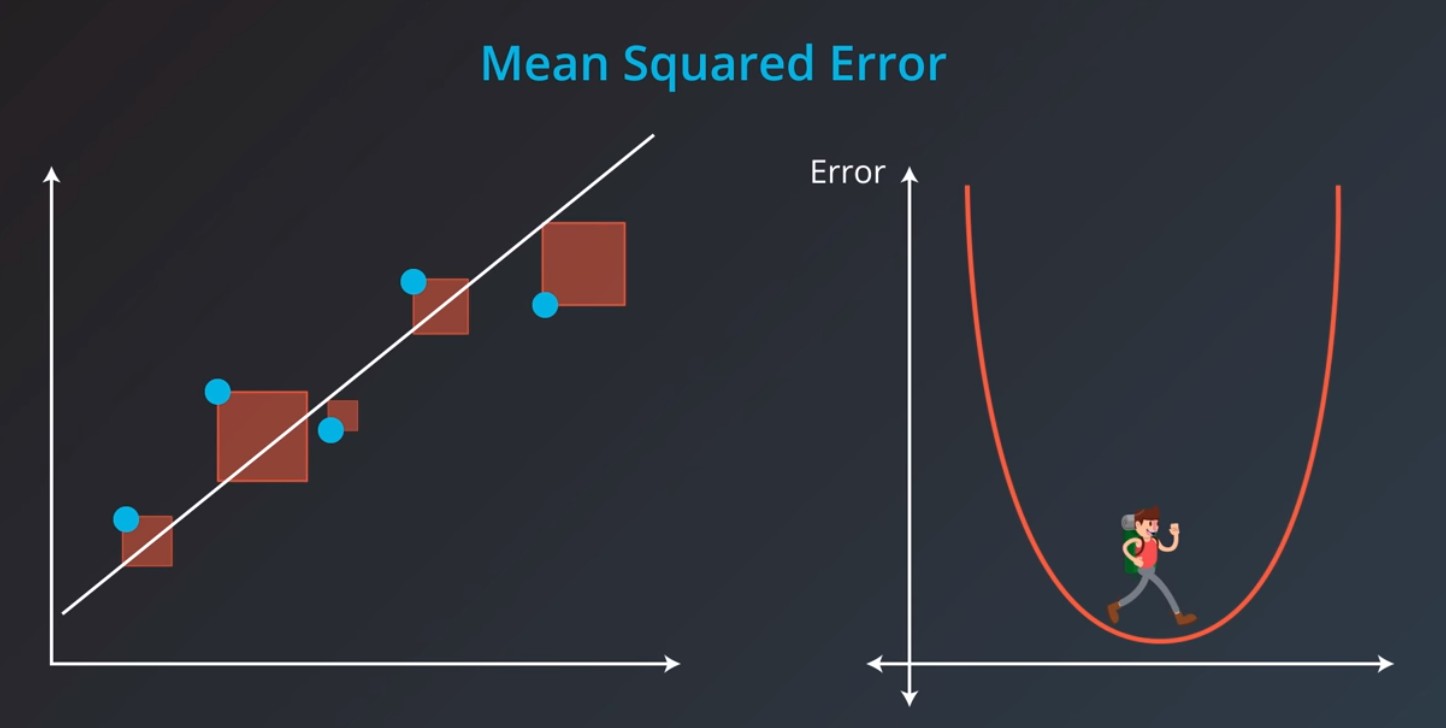
y′t denota una serie diferenciada de primer orden. Podría ser diferenciado d veces también.

En lo que respecta a los parámetros c, Ï•i y θi, se actualizan utilizando la estimación de máxima verosimilitud (MLE), al igual que en la regresión lineal. Sin embargo, a diferencia de la regresión lineal, al estimar y pronosticar una serie de tiempo usando ARIMA, podemos generar una amplia gama de predicciones futuras basadas en un solo modelo entrenado sin datos externos durante la inferencia.

**Series de tiempo como modelos de mínimos cuadrados**

Como se ha establecido anteriormente: Se puede decir qué un proceso autorregresivo de orden p escrito como AR(p) es una regresión multilineal en el qué el valor qué se desea pronosticar es el valor de la serie de tiempo al tiempo t=t como una una función lineal de los valores pasados de la variable en estudio

De manera más general se podría decir qué un proceso ARIMA es un modelo de regresión multilineal, qué intenta predecir el valor al tiempo t=t como una combinación lineal de los valores y errores previos a t-1, t-2,... t-k significativamente estadísticos. Y por lo tanto, como todo modelo de combinación lineal, ajusta sus parámetros con una minimización del error cuadrático.



**Fuente:** [**Error functions - JulienBeaulieu (gitbook.io)**](https://julienbeaulieu.gitbook.io/wiki/sciences/machine-learning/linear-regression/error-functions-1)