Universidad Nacional de Cuyo

Facultad de Ingeniería

Predicción de Precios de Criptomonedas con ARIMA y Prophet

Proyecto Final
Inteligencia Artificial I

MOLINA, Mauro
FLORES, Daniel Emiliano

Marzo 2022

Resumen

Índice

1. Introducción

1.1. Criptomonedas

Una criptomoneda es un activo digital que emplea un cifrado para garantizar integridad, autenticidad y confidencialidad de las transacciones, es decir, funcionan como moneda de cambio y solucionan los problemas que las monedas de cambio tradicionales presentan.

El dinero viene a solucionar los problemas del mercado, es decir, determinar cuánto de mi producto vale el tuyo. Por ejemplo: cuántas gallinas vale tu vaca (en palabras poco formales y en un contexto muy simple), o cuanto capital valen mis horas de servicio.

(...).El dinero es válido en la medida que otro lo acepte"(...)

Una vez que entendemos ese concepto, podemos abstraer a la moneda como cualquier objeto (real o no) que resuelva estos aspectos.

Pero para que nuestra monera sea válida, debe cumplir ciertos aspectos:

- Debe servir como moneda de cambio, es decir, quien aporta el capital recibe un bien o servicio a cambio.
- Debe servir de referencia de valor, es decir, poder determinar, a través de dicha moneda, cuánto vale un bien o servicio.
- Debe servir como reserva de valor, es decir, que su valor no se vea moficado con el paso del tiempo.

Estos 3 parametros son determinados por la sociedad, según el uso y confianza que le tengan a la moneda.

Las criptomonedas no son diferentes a las monedas corrientes, con algunas salvedades:

- No hay un ente central del que dependamos (como el banco) que regule las transacciones.
- Mantiene total privacidad, ya que toda la información está encriptada y nadie sabe la información del dueño del dinero.
- Evita problemas de "devaluación"de la moneda, ya que no depende de factores socio-económicos (como la emisión monetaria)

Este sistema funciona bajo el modelo Peer-To-Peer (P2P), donde cada usuario funciona como cliente o servidor de otro, según corresponda. Es decir, si queremos saber cuanto dinero tiene una persona, debemos mantener un registro global, que incluye a todas las transacciones que han ocurrido desde el origen de una determina criptomoneda. Este registro debe ser público, para que cualquier persona pueda consultarlo en cualquier momento.

Pero si existe este registro ¿Cómo puede existir la privacidad? La privacidad se sigue respetando, ya que realmente no se sabe a quien pertenece el dinero que figura en dichas transacciones. Esta cadena de transacciones se conoce como **BlockChain** (cadena de bloques). La idea aquí es formar una LinkedList, donde cada bloque apunta al anterior. Las personas encargadas de añadir los bloques al registro de transacciones se conocen como **mineros**.

Supongamos que Pepe le quiere transferir dinero a Pedro, entonces, la tarea del minero es crear un bloque con los datos (encriptados) de Pepe y Pedro y añadirlos a la cadena de bloque. Este registro esta almacenado en cada uno de los usuarios de la red. Una vez que se añade el bloque a la cadena, se avisa al resto de los mineros que añadan dicho en su registro. Nótese que un bloque puede contener muchas transacciones.

FirstLink

Para poder mantener esto de manera estable, las criptomonedas se rigen bajo ciertas reglas. Por ejemplo en el caso del Bitcoin: - Cada bloque está hecho de un fichero de texto, que contiene los datos de las transacciones. - Cada bloque tiene aproximadamente 1 MB - El minero que inserte el bloque recibirá una recompensa. Esta recompensa disminuye a medida que se inserten más bloques en la cadena. - El bloque debe ser aceptado por la mayoria de los mineros antes de insertarse en la cadena. - Por cada bloque se genera un Hash, que los identifica. Para hashear un bloque, se necesita: Hash anterior + Fecha y Hora de creación del bloque + Transaccion de recompensa para el minero + Todas las transacciones que quepan hasta completar 1 MB + Prueba de trabajo (detallaremos mas adelante).

LINK: SecondLink

Recordemos que los Hash tienen la propiedad de que si conocemos la formula, es muy facil, a partir de los datos, obtener ese hash, pero teniendo el hash es muy dificil saber cuales son los datos iniciales

Esta particularidad, hace que sea muy dificil o imposible la falsificación de nuevos bloques, ya que al cambiar algún dato de la transacción, cambiaría completamente su hash, el cual debe cumplir con ciertos requisitos.

Posibles problemas:

1- Supongamos que tenemos un amigo minero con el cual queremos hacer un pacto para que él se lleve la recompensa haciendo una transacción con un bitcoin y luego otra con el mismo bitcoin, realizando una estafa. Para evitar estos problemas, cada transacción debe estar aprobada por la comunidad de mineros, que no lo aprobarían. Como la recompensa es muy alta, si el bloque está mal, aún pueden llevarse ellos la recompensa.

2- Supongamos que un usuario malicioso tiene máquinas bots mineras para garantizar la mayor cantidad de votos a favor en sus transacciones. Este es el problema más grande que tenemos, ya que no se puede validar si son bots o votos autenticos. Para solucionar esto, se introduce el concepto de **_Proof of works_**. En este caso, la prueba de trabajo es que los hash inicien con cierta cantidad de 0's, el cual es modificado cada cierta cantidad de bloques. Como el hash cambia con cada minima modificación, la idea es que si se quiere falsificar una transacción, el minero deba buscar un número de forma tal que el hash empiece con la cantidad de 0's correspondientes. Esta cantidad de 0's se ajusta de acuerdo a la capacidad de cómputo de mineria que haya en la red.

Esto significa que para poder falsificar una transacción, se debe superar, en potencia informática, a más de la mitad de los mineros en toda la red.

¿De dónde salen las criptomonetas?

Las criptomonedas se generan cada vez que se crea un bloque, es decir, la recompensa de los mineros. En un principio, fue muy incentivada por especuladores y personas que evitaban mantener un registro de sus monedas.

¿Qué determina el precio de las criptomonedas?

El precio se determina por la oferta y la demanda. Cuando se incrementa la demanda, el precio sube, y cuando cae la demanda, el precio baja. Hay un número limitado de criptomonedas en circulación y las nuevas son creadas a una velocidad predecible y decreciente, esto significa que la demanda debe seguir este nivel de inflación para mantener un precio estable.

1.2. Series Temporales

Una serie de tiempo es una forma estructurada u ordenada de presentar los datos, en donde un parámetro temporal regular o irregular (fecha, hora, semana, día, mes, año, etc.) lleva asociado un valor.

Se usan para estudiar la relación causa-consecuencia entre diversas variables que cambian con el tiempo. Desde el punto de vista probabilístico, una serie temporal es una sucesión de variables aleatorias indexadas según parámetro creciente con el tiempo, que conforman un conjunto ordenado de datos y coexisten de forma dependiente entre ellas.

Con la ayuda del machine learning, puede servir para detectar patrones implicitos en el comportamiento de los fenómenos con el solo hecho de ordenarlos según un parámetro temporal.

Por ejemplo: Un caso trivial podría ser el de realizar una serie temporal con las ventas del año de una heladería. El resultado en este caso es trivial. Las ventas alcanzan un maximo relativo en verano y un mínimo relativo en invierno.

El instrumento de análisis que se suele utilizar es un modelo que permita reproducir el comportamiento de la variable de interés. Estos pueden ser:

- Univariantes: la serie temporal es analizada únicamente en función de su propio pasado;
- Multivariantes: son analizadas varias series temporales, en consecuencia a la suposición de dependencia o relación entre ellas.

Describimos matemáticamente una serie temporal univariante como un conjunto de observaciones, de tamaño T, sobre una variable Y:

$$Y_t, \quad t = 1, 2, ..., T \tag{1}$$

El pronóstico de la serie temporal implica extender los valores históricos hacia el futuro, tales como el periodo y el horizonte (cantidad de periodos proyectados). Si es un modelo univariante, se pronostica únicamente en términos de valores pasados Y_t, lo que comunmente llamamos extrapolación.

Si una serie temporal es extrapolable exactamente, decimos que ésta es determinista. Sin embargo, la gran parte de ellas no lo son y están condicionadas a una distribución de probabilidad dependiente de sus valores pasados.

Los modelos utilizados para caracterizar una serie temporal responden siempre a una misma fórmula:

$$Y_t = S_t + a_t \tag{2}$$

Donde S_t indica el comportamiento regular de la variable -conocida como parte sistémica- y a_t es el comportamiento aleatorio -tambien denominada *innovación*-.

En modelos de series univariantes, S_t se determina únicamente en función del pasado de la serie:

$$S_t = f(Y_t, Y_{t-1}, Y_{t-2}, ...) \tag{3}$$

1.2.1. Análisis de las Series Temporales

Las series temporales pueden ser descritas en función de sus componentes. En ciertos modelos, son resultado de cuatro componentes que actúan en conjunto:

- **Tendencia secular o regular.** Indica el crecimiento, decrecimiento o estacionalidad general y persistente del fenómeno observado, es la componente que refleja la evolución a largo plazo. La denotaremos como B_t.
- **Variación estacional o variación cíclica regular.** Está relacionada con el movimiento periódico de corto plazo. Se trata de una componente causal debida a la influencia de ciertos fenómenos que se repiten de manera periódica en un periodo de tiempo. Por ejemplo, si el periodo es un año, las estaciones son una variacion cíclica regular, si el periodo es una semana, los fines de semana son la variación estacional, si el periodo es un día, las horas son la variación estacional. Recopila las oscilaciones que se producen en esos períodos temporales regulares. La denotaremos como E_t.
- **Variación cíclica.** Recoge las oscilaciones periódicas de amplitud superior a un año. Movimientos normalmente irregulares alrededor de la tendencia, en las que, a diferencia de las variaciones estacionales, tiene un período y amplitud variables. En otras palabras, son variaciones estacionales de largo plazo y de eventos irregulares. Esta componente puede no existir, ya que la serie temporal puede no ser cíclica en ningún momento. La denotaremos como C_t.
- **Variación aleatoria o ruido.** Incluye todos aquellos factores que no muestran ninguna regularidad, debidos a fenómenos de carácter ocasional. Por ejemplo: Tormentas en relación con las cosechas de alguna verdura. La denotaremos como R_t.

Estas componentes generan los siguientes tipos de series:

- **Aditivas.** Sumando sus componentes: $X_t = B_t + E_t + C_t + R_t$
- **Multiplicativas.** Multiplicando sus componentes: $X_t = B_t * E_t * C_t * R_t$
- **Mixtas.** Sumando y multiplicando sus componentes, por lo que existen varias alternativas. Por ejemplo: $X_t = B_t + E_t * C_t * R_t$

Sin embargo, la clasificación arriba descrita no es utilizada para modelos ARIMA. En estos, se trata de obtener la representación de la serie en términos de la interrelación temporal entre sus elementos.

1.2.2. Coeficiente de autocorrelación

Uno de los instrumentos utilizados es el coeficiente de correlación ρ_{xy} entre dos variables x_t y y_t. Este coeficiente mide el grado de asociación lineal entre ellas:

$$\rho_{xy} = \frac{cov(x, y)}{\sqrt{V(x)V(y)}} \tag{4}$$

donde, por definición, $\rho_{xy} \in [-1, 1]$.

De este coeficiente de correlación poblacional podemos estimar el coeficiente de correlación muestral:

$$r_{xy} = \frac{\sum_{t=1}^{T}(x_t - \bar{x})(y_t - \bar{y})}{\sqrt{\sum_{t=1}^{T}(x_t - \bar{x})^2 \sum_{t=1}^{T}(y_t - \bar{y})^2}} \tag{5}$$

Si tenemos T observaciones $Y_1, ..., Y_T$, podemos crear $(T-1)$ pares de observaciones $(Y_1, Y_2), ..., (Y_{T-1}, Y_T)$. Es decir, consideramos $Y_1, ..., Y_{T-1}$ como una variable y $Y2, ..., Y_T$ como otra para definir la correlación entre ambas variables Y_t, Y_{t+1}:

$$r_{Y_t Y_{t+1}} = \frac{\sum_{t=1}^{T-1}(Y_t - \bar{Y}_{(1)})(Y_{t+1} - \bar{Y}_{(2)})}{\sqrt{\sum_{t=1}^{T-1}(Y_t - \bar{Y}_{(1)})^2 \sum_{t=1}^{T-1}(Y_{t+1} - \bar{Y}_{(2)})^2}} \tag{6}$$

Donde $\bar{Y}_{(1)}$ y $\bar{Y}_{(2)}$ es la media muestral de las primeras $T-1$ observaciones y la media muestral de las últimas $T-1$ observaciones, respectivamente.

La expresión anterior es aproximadamente:

$$r_{Y_t Y_{t+1}} \approx r_1 = \frac{\sum_{t=1}^{T-1}(Y_t - \bar{Y})(Y_{t+1} - \bar{Y})}{\sum_{t=1}^{T}(Y_t - \bar{Y})^2} \tag{7}$$

El subíndice indica el intervalo de correlación lineal de las observaciones analizadas. En el caso anterior descrito, r_1 analiza observaciones sucesivas y se denomina *coeficiente de autocorrelación de primer orden.*

Por tanto, el coeficiente de autocorrelación de orden k viene dado por:

$$r_k = \frac{\sum_{t=1}^{T-k}(Y_t - \bar{Y})(Y_{t+k} - \bar{Y})}{\sum_{t=1}^{T}(Y_t - \bar{Y})^2} \tag{8}$$

Para interpretar el conjunto de los coeficientes de autocorrelación de una serie temporal se utiliza un *correlograma*, que es un gráfico que agrupa los k coeficientes de autocorrelación.

Estos suelen comenzar en $k = 1$ puesto que un $k = 0$ en la Ecuación (8) siempre es 1.

1. Si una serie es puramente aleatoria, $r_k \approx 0$ para cualquier $k \neq 0$ siendo T suficientemente grande.

2. Series sin tendencia oscilantes en torno a una media constante:

a) Si contiene observaciones por encima de la media seguidas de una o más observaciones por encima de la media (lo mismo para observaciones por debajo de la media), entonces r_k decrece tendiendo a cero rápidamente a medida que aumenta k.

b) Si contiene observaciones alternantes por encima y debajo de la media, el correlograma presenta valores decrecientes que alternan su signo.

3. Si la serie contiene una tendencia, los valores de r_k no decrecerán a cero rápidamente. Esto se debe a que un valor observado tiene sucesivos valores que estarán por encima (o debajo) de la media. Este tipo de comportamiento nos da poca información.

4. Si una serie presenta algún tipo de ciclo, el correlograma también presentará una oscilación a la misma frecuencia. En las series con un comportamiento cıclico permanente, el correlograma da de nuevo poca información porque lo domina el comportamiento cıclico presente en los datos.

2. Procesos Estocásticos

Un proceso estocástico es una familia de variables aleatorias relacionadas entre sí y que siguen una ley de distribución conjunta. Lo denotaremos por:

$$..., Y_{t-2}, Y_{t-1}, Y_t, Y_{t+1}, Y_{t+2}, ... \text{ o } Y_t$$

A simple vista podemos entender que fijar $t = t_0$ nos dá una variable aleatoria de la secuencia *presumiblemente ordenada*[1]. Además, llamaremos *realización* del proceso a la asignación de un valor a cada una de las variables aleatorias del mismo.

En virtud de lo anterior, una serie temporal $Y_t, Y_{t+1}, Y_{t+2}, ...$ se puede interpretar como una realización muestral de un proceso estocástico para un número finito de periodos $t = 1, 2, ..., T$.

2.1. Características

Función de distribución. Incluye todas las funciones de distribución para cualquier subconjunto finito de variables aleatorias del proceso:

$$F[Y_{t_i}, Y_{t_{i+1}}, ..., Y_{t_n}] \text{ siendo } n \text{ finito.}$$

Momentos del proceso estocástico. El primer momento de un proceso estocástico es el conjunto de las medias de todas las variables aleatorias del proceso:

$$E(Y_t) = \mu_t < \infty, \quad t = 0, \pm 1, \pm 2, ... \tag{9}$$

El segundo momento viene dado por el conjunto de las varianzas de todas las variables aleatorias del proceso y por las covarianzas entre todo par de variables aleatorias:

[1] El orden en la sucesión de observaciones es único en una serie temporal. Alterarlo implicaría modificar las características de la serie.

$$
\begin{aligned}
V(Y_t) = E[Y_t - \mu_t]^2 = \sigma_t^2 < \infty, \quad t = 0, \pm 1, \pm 2, ... \\
cov(Y_t, Y_s) = E[Y_t - \mu_t][Y_s - \mu_s] = \gamma_{t,s} \quad \forall t, s \quad (t \neq s)
\end{aligned}
\tag{10}
$$

2.2. Procesos estocásticos estacionarios

Para hacer análisis y predicciones consistentes a una serie de tiempo, es necesario que la estructura probabilística que subyace en el proceso estocástico sea estable en el tiempo. Es decir, las regularidades del comportamiento pasado deben ser capaces de proyectarse a futuro. Estas características de un proceso estocástico se las conoce como *estacionariedad*.

Estacionariedad estricta. Un proceso estocástico es estrictamente estacionario si y solo si

$$F[Y_{t_i}, Y_{t_{i+1}}, ..., Y_{t_n}] = F[Y_{t_i+k}, Y_{t_{i+1}+k}, ..., Y_{t_n+k}] \tag{11}$$

en otras palabras, si la función de distribución de cualquier conjunto finito de variables aleatorias del proceso no se altera si se desplaza k periodos en el tiempo.

Estacionariedad en covarianza. Un proceso estocástico es estacionario en covarianza si y solo si:

1. Es estacionario en media, es decir, todas las variables aleatorias tienen la misma media y es finita:

$$E(Y_t) = \mu < \infty, \quad \forall t \tag{12}$$

2. Todas las variables aleatorias tienen la misma varianza y es finita:

$$V(Y_t) = E[Y_t - \mu]^2 = \sigma_Y^2 < \infty, \quad \forall t \tag{13}$$

3. La covarianza lineal entre dos variables aleatorias del proceso que disten k periodos de tiempo es la misma que existe entre cualesquiera otras dos variables que estén separadas también k periodos. También conocemos esta propiedad como *autocovarianza*:

$$cov(Y_t, Y_s) = E[Y_t - \mu][Y_s - \mu] = \gamma_{t,s} = \gamma_{|t-s|} = \gamma_k \quad \forall k \tag{14}$$

Finalmente, decimos que un proceso estocástico estacionario en covarianza está caracterizado si se conoce:

$\mu, \quad V(Y_t), \quad \gamma_k, \quad k = 0, \pm 1, \pm 2, ...$

Si bien las series económicas no presentan comportamientos estacionarios, particularmente las series temporales que reflejan el precio de una criptomoneda, se procederá a proponer modelos que para procesos estacionarios. A partir de estos se puede modelar para un proceso no estacionario.

2.3. Formalización de funciones

Función de autocovarianzas. Es una función de k (número de periodos de separación entre las variables) que recoge el conjunto de las autocovarianzas del proceso estocástico estacionario:

$$\gamma_k,\ k = 0, 1, 2, ...$$

Es una función simétrica:

$$\gamma_k = \gamma_{-k}$$

y, además, incluye la varianza del proceso para $k = 0$:

$$\gamma_0 = V(Y_t).$$

Función de autocorrelación. Al igual que la anterior, recoge toda la información de la estructura dinámica lineal del proceso estocástico con la diferencia de que es independiente de las unidades de la variable.

El coeficiente de autocorrelación de orden k de un proceso estocástico estacionario mide el grado de asociación lineal existente entre dos variables aleatorias del proceso separadas por k periodos:

$$\rho_k = \frac{cov(Y_t, Y_{t+k})}{\sqrt{V(Y_t) * V(Y_{t+k})}} = \frac{\gamma_k}{\sqrt{\gamma_0 * \gamma_0}} = \frac{\gamma_k}{\gamma_0} \tag{15}$$

donde

$$|\rho_k| \leq 1,\ \forall k$$

La función de autocorrelación de un proceso estocástico estacionario es una función de k que recoge el conjunto de los coeficientes de autocorrelación del proceso y se denota por ρ_k. La función de autocorrelación se suele representar gráficamente por medio de un correlograma.

El coeficiente de autocorrelación de orden 0 es, por definición, 1. Genera una función simétrica, por lo que un correlograma solo representa valores positivos de k. Esta función, además, tiende a cero rápidamente a medida que k tiende a infinito.

De estas premisas, al observar un correlograma podemos concluir que su correspondiente serie no es estacionaria si la función no decrece rápidamente.

2.3.1. Estacionaridad y estimación de momentos

Si la serie temporal $(Y_1, Y_2, ..., Y_T)$ no fuera estacionaria, ya no tendríamos una única media, ni una única varianza y las autocorrelaciones serían demasiadas. La estacionaridad es, desde luego, una restricción necesaria para estimar los momentos de un proceso estocástico.

2.4. Ruido blanco

El proceso estocástico ruido blanco, denotado w_t, lo definimos como:

$$\begin{aligned} E(w_t) &= 0, \forall t \\ V(w_t) &= \sigma^2, \forall t \\ Cov(w_t, w_s) &= 0, \forall t \neq s \end{aligned}$$

Un proceso ruido blanco w_t $RB(0, \sigma^2)$ es estacionario si la varianza σ^2 es finita, con función de autocovarianzas:

$$\gamma_k = \begin{cases} \sigma^2, & k = 0 \\ 0, & k > 0 \end{cases}$$

y función de autocorrelación:

$$\rho_k = \begin{cases} 1, & k = 0 \\ 0, & k > 0 \end{cases}$$

EJEMPLO DE RUIDO BLANCO

3. Modelos estacionarios

La estructura de dependencia temporal de una serie está recogida en la función de autocovarianzas y/o en la función de autocorrelación. Mediante un modelo ARMA, se tratará de reproducir el comportamiento general y posteriormente predecir con la ayuda del mismo.

3.1. Modelo Lineal General

Anteriormente descompusimos una serie temporal univariante en dos componentes (2). Ahora, nuestra parte sistémica contiene la información para construir el modelo mientras que la innovación contendrá valores sin dependencia entre sí, tanto en valores pasados como con su parte sistémica. La innovación, como se habrá podido intuir, es el ruido blanco.

Así, un modelo teórico capaz de describir un comportamiento de una serie temporal con media cero sería:

$$Y_t = f(Y_{t-1}, Y_{t-2}, ...) + w_t, \quad t = 1, 2, ... \tag{16}$$

El proceso estocástico puede responder a un proceso de Gauss, por lo que Y_t puede expresarse como una combinación lineal de sus valores pasados infinitos más una innovación:

$$Y_t = \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + ... + w_t \quad t = 1, 2, ... \tag{17}$$

donde Y_t es un proceso estacionario y π_i son constantes con $\phi_t \neq 0$.

Para que esto suceda, es preciso que el proceso sea *no anticipante*: cualquier valor Y_i no debe depender de valores futuros. En adición, el proceso debe ser *invertible*, es decir que el presente dependa de forma convergente con su propio pasado. A medida que nos alejemos en el tiempo, la influencia de Y_k debe disminuir. Esta restricción viene dada por:

ECUACION
Operador de retardos. Lo definimos como:

$$BY_t = Y_{t-1} \tag{18}$$

particularmente:

$$B^2Y_t = B(BY_t) = BY_{t-1} = Y_{t-2}$$

y generalmente:

$$B^kY_t = Y_{t-k} \tag{19}$$

Reescribimos (17) en función del operador de retardos:

$$\begin{gathered} Y_t = (\pi_1 B + \pi_2 B^2 + ...)Y_t + w_t \\ (1 - \pi_1 B - \pi_2 B^2 - ...)Y_t = w_t \end{gathered}$$

Si $(1 - \pi_1 B - \pi_2 B^2 - ...) = \pi_\infty(B)$:

$$\pi_\infty(B)Y_t = w_t \tag{20}$$

Teniendo presente que $\pi_\infty(B)$ es un polinomio de orden infinito, es necesario aproximarlo a uno de orden finito -los modelos a desarrollar deberán representar procesos estocásticos acotados en el tiempo:

$$\pi_\infty(B) \approx \frac{\phi_p(B)}{\theta_q(B)} \tag{21}$$

donde:

$$\begin{aligned} \phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p \\ \theta_p(B) = 1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q \end{aligned} \tag{22}$$

Sustituyendo (21) en (20):

$$\phi_p(B)Y_t = \theta_q(B)w_t \tag{23}$$

Por lo tanto, este modelo lineal general admite tres representaciones:

- *Puramente regresiva, $AR(\infty)$:* el valor presente de la variable se representa en función de su propio pasado más una innovación contemporánea.
- *Puramente de medias móviles, $MA(\infty)$:* el valor presente de la variable se representa en función de todas las innovaciones pasadas y presente.
- *Finita:*

$$\begin{gathered} (1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p)Y_t = (1 - \theta_1 B - \theta_2 B^2 - ... - \theta_q B^q)w_t \\ Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + w_t - \theta_1 w_{t-1} - \theta_2 w_{t-2} - ... - \theta_q w_{t-q} \end{gathered} \tag{24}$$

Este modelo se denomina *Autorregresivo de Medias Móviles* de orden (p, q). En él, el valor de Y_t depende del pasado de Y hasta el momento $t - p$ de la innovación contemporánea y su pasado hasta el momento $t - q$.

3.2. Procesos Autoregresivos de Medias Móviles

Como hemos señalado anteriormente, la ecuación 24 es una aproximación finita al modelo lineal general tanto en su forma $AR(\infty)$ como $MA(\infty)$.

De hecho, si es estacionario su representación $MA(\infty)$ es

$$Y_t = \frac{\theta_q(B)}{\phi_p(B)} w_t \quad \rightarrow \quad Y_t = \psi_\infty(B) w_t \quad \rightarrow \quad Y_t = w_t + \psi_1 w_{t-1} + \psi_2 w_{t-2} + ... \tag{25}$$

y si es invertible, su representación $AR(\infty)$ es

$$\frac{\phi_p(B)}{\theta_q(B)} Y_t = w_t \quad \rightarrow \quad \pi_\infty(B) Y_t = w_t \quad \rightarrow \quad Y_t = w_t + \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + ... \tag{26}$$

Teorema de estacionariedad. Un proceso autoregresivo de medias móviles finito $ARMA(p,q)$ es estacionario sí y solo sí el módulo de las raíces del polinomio autorregresivo $\phi(B)$ está fuera del círculo unidad.

Las condiciones de estacionariedad del modelo $ARMA(p,q)$ vienen impuestas por la parte autorregresiva, dado que la parte de medias móviles finita siempre es estacionaria.

Para comprobar si el modelo $ARMA(p,q)$ es no anticipante e invertible , se estudia su representación autorregresiva general.

Teorema de invertibilidad. Un proceso autorregresivo de medias móviles finito $ARMA(p,q)$ es invertible sí y solo sí el módulo de las raíces del polinomio de medias móviles $\theta_p(B)$ está fuera del círculo unidad.

Las condiciones de invertibilidad del modelo $ARMA(p,q)$ vienen impuestas por la parte de medias móviles, dado que la parte autorregresiva finita siempre es invertible porque está directamente escrita en forma autorregresiva

El modelo $ARMA(p,q)$ tiene media cero, varianza constante y finita y una función de autocovarianzas infinita. La función de autocorrelación es infinita decreciendo rápidamente hacia cero pero sin truncarse.

3.2.1. $ARMA(1,1)$

En este modelo, Y_t se determina en función de su pasado hasta el primer retardo, la innovación contemporánea y el pasado de la innovación hasta el retardo 1:

$$Y_t = \phi Y_{t-1} + w_t - \theta w_{t-1} \qquad w_t \ RB(0,\sigma^2) \qquad t = 1, 2, ... \tag{27}$$

La memoria de este proceso es larga debido a que presenta la estructura autorregresiva. Es decir, una perturbación w_t ingresa al sistema afectando a Y_t y, a través de él, al futuro. Sin embargo, por su estructura de medias móviles, la perturbación w_t afecta directamente a Y_t Y_{t-1}.

ARREGLAR GRAFICO

$$\begin{array}{cccccccccccc}
... & \rightarrow & Y_{t-2} & \rightarrow & Y_{t-1} & \rightarrow & Y_t & \rightarrow & Y_{t+1} & \rightarrow & Y_{t+2} \rightarrow ... & \\
 & & \uparrow & \nearrow & \uparrow & \nearrow & \uparrow & \nearrow & \uparrow & \nearrow & \uparrow & \nearrow
\end{array}$$

$$... \quad w_{t-2} \quad w_{t-1} \quad w_t \quad w_{t+1} \quad w_{t+2}$$

Por el teorema de estacionariedad, es necesario y suficiente comprobar que las raíces del polinomio autorregresivo estén fuera del círculo unidad:

$$\phi_1(B) = 1 - \phi B = 0 \quad \rightarrow \quad B = \frac{1}{\phi} \quad \rightarrow \quad |B| = |\frac{1}{\phi}| \quad \rightarrow \quad |\phi| < 1$$

Por el teorema de invertibilidad, es necesario y suficiente comprobar que las raíces del polinomio de medias móviles estén fuera del círculo unidad:

$$\theta_1(B) = 1 - \theta B = 0 \quad \rightarrow \quad B = \frac{1}{\theta} \quad \rightarrow \quad |B| = |\frac{1}{\theta}| \quad \rightarrow \quad |\theta| < 1$$

La media de este proceso es:

$$\begin{aligned} E(Y_t) = E(\phi Y_{t-1} + w_t - \theta w_{t-1}) &= \phi E(Y_t) \\ \rightarrow \quad (1-\phi)E(Y_t) &= 0 \\ \rightarrow \quad E(Y_t) &= 0 \end{aligned}$$

Para determinar la función de covarianzas, tenemos en cuenta que:

$$E(Y_{t-1}w_{t-1}) = E[(\phi Y_{t-2} + w_{t-1} - \theta w_{t-2})w_{t-1}] = E(w_{t-1})^2 = \sigma^2$$

En tanto, dicha función es:

$$\begin{aligned} \gamma_0 &= E(Y_t - E(Y_t))^2 = E(Y_t)^2 = E(\phi Y_{t-1} + w_t - \theta w_{t-1})^2 \\ &= \phi^2 E(Y_{t-1})^2 + E(w_t)^2 + \theta^2 E(w_{t-1})^2 + 2\phi E(Y_{t-1}w_t) - 2\phi\theta E(Y_{t-1}w_{t-1}) - 2\theta E(w_t w_{t-1}) \\ &= \phi^2\gamma_0 + \sigma^2 + \theta^2\sigma^2 - 2\phi\theta\sigma^2 \\ &= \phi^2\gamma_0 + (1+\theta^2 - 2\phi\theta)\sigma^2 \\ \rightarrow \quad \gamma_0 &= \frac{(1+\theta^2-2\phi\theta)\sigma^2}{1-\phi^2} \end{aligned}$$

$$\begin{aligned} \gamma_1 &= E(Y_t - E(Y_t))(Y_{t-1} - E(Y_{t-1})) = E(Y_tY_{t-1}) \\ &= E[(\phi Y_{t-1} + w_t - \theta w_{t-1})Y_{t-1}] \\ &= \phi E(Y_{t-1})^2 + E(Y_{t-1}w_t) - \theta E(Y_{t-1}w_{t-1}) \\ &= \phi\gamma_0 - \theta\sigma^2 \end{aligned}$$

$$\begin{aligned} \gamma_2 &= E(Y_t - E(Y_t))(Y_{t-2} - E(Y_{t-2})) = E(Y_tY_{t-2}) \\ &= E[(\phi Y_{t-1} + w_t - \theta w_{t-1})Y_{t-2}] \\ &= \phi E(Y_{t-1}Y_{t-2}) + E(Y_{t-2}w_t) - \theta E(Y_{t-2}w_{t-1}) \\ &= \phi\gamma_1 \end{aligned}$$

Es decir que la función de autocovarianzas para un $ARMA(1,1)$ es:

$$\gamma_k = \begin{cases} \frac{(1+\theta^2-2\phi\theta)\sigma^2}{1-\phi^2}, & k=0 \\ \phi\gamma_0 - \theta\sigma^2, & k=1 \\ \phi\gamma_{k-1}, & k>1 \end{cases}$$

La función γ_k nos brinda un formalismo a lo expresado previamente. La autocovarianza de orden 0 cuenta con términos provenientes de la parte autorregresiva y de medias móviles. De forma similar, en la autocovarianza de orden 1 tenemos términos de $AR(1)$ $MA(1)$. Finalmente, órdenes superiores dependen exclusivamente de la autorregresión.

La función de autocorrelación de un $ARMA(1,1)$ es:

$$\rho_k = \begin{cases} \phi - \frac{\theta\sigma^2}{\gamma_0}, & k=1 \\ \phi\rho_{k-1} \end{cases}$$

AGREGAR UNA DESCRIPCIÓN A AUTOCORRELACION

3.2.2. $ARMA(p,q)$

Los resultados obtenidos para $ARMA(1,1)$ son generalizables para este modelo. Observemos: para los primeros q coeficientes $\rho_1, ..., \rho_q$ dependerán de los parámetros autorregresivos y de medias móviles, mientras que los siguientes no dependerán del parámetro de medias móviles. En consecuencia, la función de autocorrelación decrecerá rápidamente a cero.

Es posible modificar $ARMA(p,q)$ de tal forma que su media no sea nula añadiendo una constante al proceso estacionario:

$$Y_t = \delta + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + w_t - \theta_1 w_{t-1} - \theta_2 w_{t-2} - ... - \theta_q w_{t-q} \qquad w_t \; RB(0,\sigma^2) \tag{28}$$

La única diferencia respecto al modelo sin la constante δ es que su media no es cero:

(CHEQUEAR FLECHA)

$$\begin{aligned} E(Y_t) &= E(\delta + \phi_1 Y_{t-1} + ... + \phi_p Y_{t-p} + w_t + \theta_1 w_{t-1} + ... + \theta_q w_{t-q}) \\ &= \delta + \phi_1 E(Y_t) + ... + \phi_p E(Y_t) \\ \rightarrow \quad \delta &= (1 - \phi_1 - ... - \phi_p) E(Y_t) \\ E(Y_t) &= \frac{\delta}{(1 - \phi_1 - ... - \phi_p)} \end{aligned}$$

4. Modelos no estacionarios

Lo visto hasta el momento responde a la estacionariedad en covarianza, es decir, media y varianza constantes y finitas junto a autocovarianzas independientes del tiempo. Sin embargo, en series económicas este comportamiento no se replica.

En este caso, quitaremos el supuesto de que tenemos una serie estacionaria en covarianza y mediante algún método transformaremos la serie para estabilizar la covarianza.

4.1. Modelos ARIMA(p,d,q)

Supongamos el modelo $ARMA(p,q)$:

$$\Phi_p(B)Y_t = \Theta_q(B)w_t$$

donde el polinomio AR se puede factorizar en función de sus p raíces $B_1, B_2, ...B_p$:

$$\Phi_p(B) = (1 - B_1^{-1}B)(1 - B_2^{-1}B)...(1 - B_p^{-1}B)$$

Si suponemos que $p-1$ raíces son estacionarias (con módulo fuera del círculo unidad) y una de ellas es unitaria, $B_i = 1$, entonces el polinomio AR se puede reescribir como sigue:

$$\Phi_p(B) = (1 - B_1^{-1}B)(1 - B_2^{-1}B)...(1 - B_p^{-1}B) = \varphi_{p-1}(B)(1 - (-1)^{-1}B)$$
$$\Phi_p(B) = \varphi_{p-1}(B)(1 - B)$$

donde el polinomio $\varphi_{p-1}(B)$ es el producto de los $p-1$ polinomios de orden 1 asociados a las raíces B_j asociadas al círculo unidad. Necesariamente se aclara que es un polinomio estacionario.

Sustituyendo en el modelo $ARMA(p,q)$ obtenemos:

$$\begin{aligned} \varphi_{p-1}(B)(1 - B)Y_t &= \Theta_q(B)w_t \\ \varphi_{p-1}(B)\Delta Y_t &= \Theta_q(B)w_t \end{aligned} \tag{29}$$

Nótese que Δ es el polinomio que recoge la raíz unitaria.

La ecuación 29 representa a un modelo con un comportamiento no estacionario ya que contiene una raíz unitaria. Un proceso Y_t con estas características se le denomina *proceso integrado de orden 1*

El proceso AR puede contener más de una raíz unitaria, por lo que se puede generalizar el modelo 29 como:

$$\varphi_{p-d}(B)\Delta^d Y_t = \Theta_q(B)w_t \tag{30}$$

Así mismo, el polinomio $\varphi_{p-d}(B)$ es estacionario porque sus $p-d$ raíces tienen módulo fuera del círculo unidad, y el polinomio Δ^d, de orden d, contiene las d raíces unitarias no estacionarias.

Este proceso Y_t con estas características se denomina *proceso integrado de orden d* y se denota por $Y_t \sim I(d)$.

Formalmente, un proceso Y_t es integrado de orden d, $Y_t \sim I(d)$, si Y_t no es estacionario, pero su diferencia de orden d, $\Delta^d Y_t$, sigue siendo un proceso $ARMA(p-d,q)$ estacionario e invertible.

5. Prophet

Este modelo surge por las limitaciones que presentan los modelos SARIMAX, como estacionariedad y valores igualmente espaciados en el tiempo (siendo que esto puede variar mucho en la realidad).

La idea de este algoritmo es plantear a la serie temporal como la suma de 4 componentes:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (31)$$

donde:

- $g(t)$. Describe la tendencia lineal o de crecimiento logístico de largo plazo, creando una función definida por partes en los puntos de cambio (simplificando la predicción).

  Los puntos de cambio son momentos en los datos donde los datos cambian de dirección.

- $s(t)$. Función que define la estacionalidad (anual, mensual, semanal, etc.)
- $h(t)$. Hechos concretos que pueden alterar los valores de la serie (vacaciones, feriados, paros, etc.), a definirse por el usuario
- ϵ_t. Término de error.

Prophet intenta ajustar las funciones (lineales o no) a los datos dando más o menos importancia a los distintos efectos, es decir, un modelo muy parametrizable:

1. **Tendencia lineal $g(t)$.** Para calcular el termino $g(t)$, Prophet ofrece 2 alternativas.

   - Crecimiento Logístico. Estos casos corresponden a situaciones en donde se da una saturación que no permite el crecimiento más allá de un límite determinado. Se aplica una sigmoide pero un poco cambiada.
     - C: Capacidad de Carga. Define la carga máxima que puede llegar a tomar la curva. Nótese que este valor puede no ser fijo y depender de otro parámetro externo C(t).
     - k: Es la tasa de crecimiento. Define qué tan rápido pasará de 0 a la capacidad de carga o viceversa.
     - m: Es un parámetro de compensación. Define el punto de inflexión de la función, es decir, cuando cambia de concavidad.

     $$g(t) = \frac{C}{1 + e^{-k(t-m)}} \quad (32)$$

   - Modelo Lineal por Partes: Para pronosticar problemas que no muestran un crecimiento saturado, existe una tasa de crecimiento constante por partes. δ: tiene los ajustes de tarifas

     k: es la tasa de crecimiento

     m: es el parámetro de compensación, en estos casos, los puntos de quiebre. Los puede definir el usuario o los puede calcular Prophet.

     $$g(t) = (k + a(t)^T\delta)t + (m + a(t)^T\gamma) \quad (33)$$

2. **Ajuste estacional $s(t)$.** Se hace utilizando series de Fourier. Fourier demostró que cualquier función periódica puede formarse a partir de una suma infinita de senos y cosenos.

   $$s(t) = \Sigma_{n=1}^{N}[a_n cos(\frac{2n\pi}{T}t) + b_n sen(\frac{2n\pi}{T}t)] \quad (34)$$

3. **Término h(t).** Queda definido por el usuario, y son aquellos valores que pueden llegar a alterar la serie temporal. Por ejemplo: La llegada del verano altera las ventas de una heladería, los hechos históricos importantes, los fin de semana largos en la venta de vuelos, etc.

4. **Error ϵ_t.** Se estima por diferencia.

6. Análisis de Resultados

7. Conclusión

Referencias

[1] Hyndman, R.J. y Athanasopoulos, G., *Forecasting: principles and practice*, 3ra edición, OTexts: Melbourne, Australia. 2021.

[2] Shumway, R. H. y Stoffer, D. S., *Time series analysis and its applications: With R examples.* Springer: New York. 2006.

[3] Porro Vargas, M. A., *Modelos de análisis de criptomonedas basado en aprendizaje automático y series temporales.* Universidad Central de Venezuela: Caracas. 2018.

[4] González Casimiro, M. P., *Análisis de series temporales: Modelos ARIMA.* Sarriko-On: País Vasco. 2009.

[5] Ask Python, *Crypto Price Prediction with Python.* 2021. Recuperado el 04-01-2022 de artículo.

[6] Benedict, N., *Ethereum Price Prediction with Python.* 2021. Recuperado el 04-01-2022 de artículo.

[7] Lertlumprasert, C., *CryptoCurrency Price Prediction with Python.* 2018. Recuperado el 04-01-2022 de artículo.

[8] Pierre, S., *A Guide to Time Series Forecasting in Python.* 2021. Recuperado el 04-01-2022 de artículo.

[9] Grigsby, J., *Multivariate Time Series Forecasting with Transformers.* 2021. Recuperado el 04-01-2022 de artículo.

[10] Lhessani, S., *How I Tripled My Return on Bitcoin Using Mathematics, Algorithms, and Python.* 2021. Recuperado el 04-01-2022 de artículo.

[11] Kharwal, A., *Cryptocurrency Price Prediction with Machine Learning.* 2021. Recuperado el 04-01-2022 de artículo.

[12] Kumar, S., *Cryptocurrency (Ethereum) Price Prediction Using Python.* 2021. Recuperado el 04-01-2022 de artículo.

[13] Vasulka, M., *Stock and cryptocurrency price prediction with python Prophet.* 2021. Recuperado el 07-01-2022 de artículo.

[14] Peixero, M., *The Complete Guide to Time Series Analysis and Forecasting*. 2019. Recuperado el 08-01-2022 de artículo .

[15] Rodó, P, *Modelo AR(1)*. 2021. Recuperado el 11-01-2022 de artículo .

[16] Rodó, P., *Modelo autorregresivo (AR)*. 2021. Recuperado el 11-01-2022 de artículo .

[17] Rodó, P., *Modelo ARMA*. 2021. Recuperado el 11-01-2022 de artículo .

[18] Wikipedia, *Media Móvil*. 2009. Recuperado el 11-01-2022 de artículo .

[19] Saenz, F., *Medias móvil simple, exponencial y ponderada: formulas y ejemplos*. 2020. Recuperado el 11-01-2022 de artículo .

[20] Lopez Briega, R. E., *Series de tiempo con Python*. 2016. Recuperado el 13-01-2022 de artículo .

[21] Ichi-Pro, *Análisis de series de tiempo con Facebook Prophet: cómo funciona y cómo usarlo*. 2021. Recuperado el 08-02-2022 de artículo .

[22] Lopez Saez, J. I., *Análisis de Series de Tiempo*. 2018. Recuperado el 08-02-2022 de artículo .

[23] Sagar A., *Cryptocurrency Price Prediction Using Deep Learning*. 2019. Recuperado el 12-03-2022 de artículo .