edix
Digital Workers

Fastbook 09

Analítica de Cliente & Predictive Analytics

Teoría general de series temporales (III)

Edix Educación

09. Teoría general de series temporales (III)

Entramos ahora en el último tramo de la asignatura, en el que seguiremos hablando de series temporales y dedicaremos nuestros últimos esfuerzos a estudiar los modelos avanzados de series temporales.

No obstante, llegados a este punto **debemos darnos una buena palmadita en la espalda y estar muy contentos por todo lo que sabemos ya sobre series temporales** (además de lo visto sobre análisis de segmentación).

Ahora daremos un paso más en los modelos que podemos utilizar para modelizar y predecir series temporales. Ya no nos queda nada para convertirnos en auténticos expertos en la materia. ¡Vamos allá!

Autor: Miguel Ángel Fernández

Lección 1 de 4

Introducción

Edix Educación

A modo de repaso, vamos a enumerar los grandes objetivos cubiertos hasta ahora para tener claro el punto en el que estamos antes de ver qué trataremos en este fastbook y en el siguiente. A grandes rasgos, hemos aprendido:

- Qué es una serie temporal.
- De dónde surgen las series temporales.
- Qué conceptos hay alrededor de una serie temporal: *granularidad, atípicos…*
- Cómo tratar y pintar series temporales en R.
- Qué elementos componen una serie temporal.
- Modelos básicos de series temporales con los que modelizar y predecir.

Con estos conocimientos que hemos adquirido estamos listos para convertirnos en auténticos expertos en series temporales. Esta disciplina, como cualquier otra, requiere de mucha experiencia y mucho más tiempo de estudio que el que tenemos en este y el cualquier curso que se pueda impartir.

Estamos en una situación similar a cuando uno termina un grado universitario: hemos aprendido muchas cosas, pero **es necesario tiempo y dedicación para dominar estos conceptos.**

El mejor consejo que te puedo dar es que intentes seguir aprendiendo con datos y problemas que realmente te resulten de interés, como pueden ser los datos de la empresa en la que estás trabajando.

Tratar correctamente las series temporales con R y/o Python, ajustar unos modelos y lanzar predicciones con los modelos aprendidos hasta este fastbook es el primer paso que nos permitirá dominar lo que hemos visto y levantará nuevas preguntas con las que seguir aprendiendo.

Dicho esto, aún nos quedan dos fastbooks para terminar la asignatura y, ahora, vamos a dar un paso más allá en los modelos que podemos utilizar para modelizar y predecir series temporales. Concretamente vamos a ver **la familia de modelos llamada ARIMA** en la literatura.

Estos modelos tienen un plus de complejidad comparado con los tres que ya hemos visto, por ello, como alumno, tienes que tener claro lo siguiente. **Dada la complejidad matemática que hay detrás de estos modelos, trataremos de adquirir una intuición básica de cómo funcionan y simplificaremos su explicación en este fastbook**. Los curiosos podrán aprovecharse de lo aprendido para poder entender en cualquier otra referencia todo el detalle que hay detrás de los modelos ARIMA con mayor facilidad.

Lección 2 de 4

ARIMA

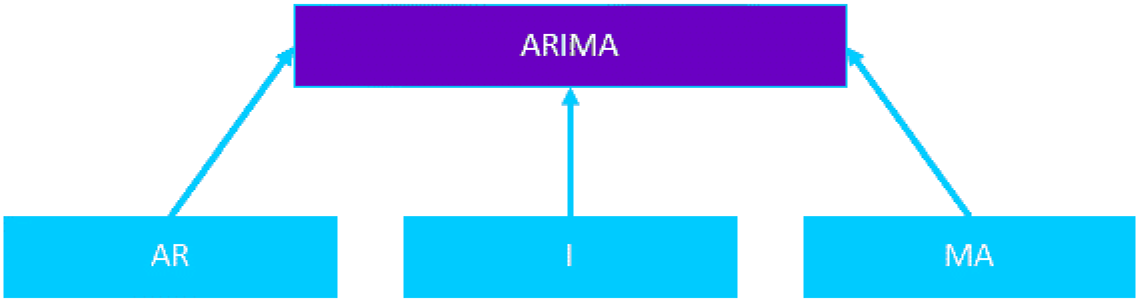
Edix Educación

El **modelo ARIMA** (AutoRegressive Integrated Moving Average) es uno de los modelos de series temporales más famosos y estudiados de toda la literatura. Se han escrito libros enteros tratando esta familia de modelos y, en este fastbook, vamos a entender las partes que lo componen.

Un modelo ARIMA cualquiera se compone de tres partes:

- La parte *AR*.
- La parte *I*.
- La parte *MA*.

Enseguida veremos qué son cada una de estas partes, pero simplemente debemos tener la idea de que un modelo *ARIMA* es una combinación de 3 modelos 'sencillitos':

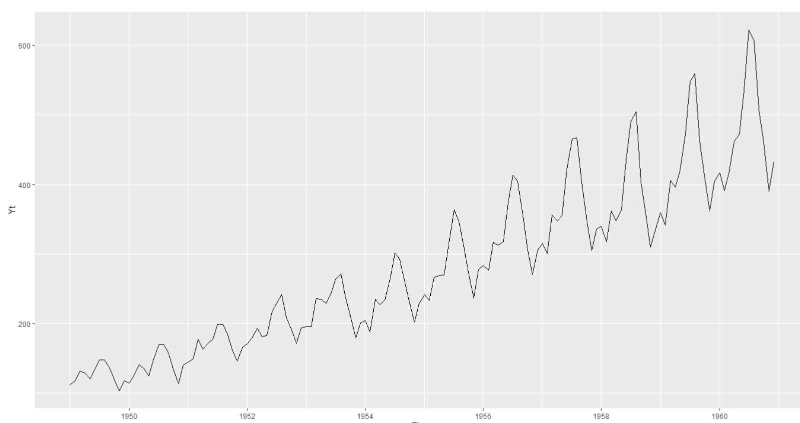


Nosotros vamos a ir directos a estudiar los modelos ARIMA, pero para comprender el modelo completo debemos ir en orden y estudiar cada una de sus partes. De ese modo **vamos a dedicar una sección a cada componente** (AR, I **y** MA) antes de ver la formulación completa del modelo ARIMA.

AR

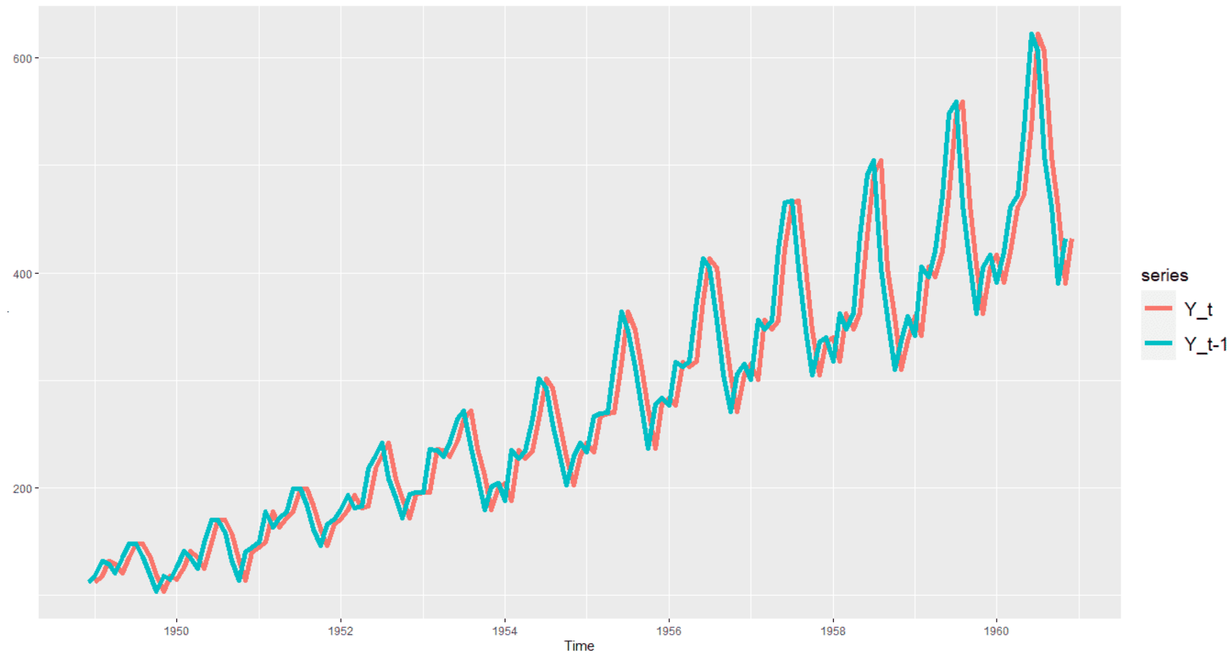
Las siglas AR quieren decir *autoregression*. Si tenemos una serie temporal Yt que queremos modelizar, la componente AR de un modelo ARIMA es la encargada de utilizar la información del pasado de nuestra serie (es decir, utilizar Y_{t-1}, Y_{t-2}, Y_{t-3}) para modelizar y poder predecir Y_t.

Supongamos que nuestra serie temporal Y_t que queremos modelizar es la siguiente (olvidémonos de qué puede estar midiendo):

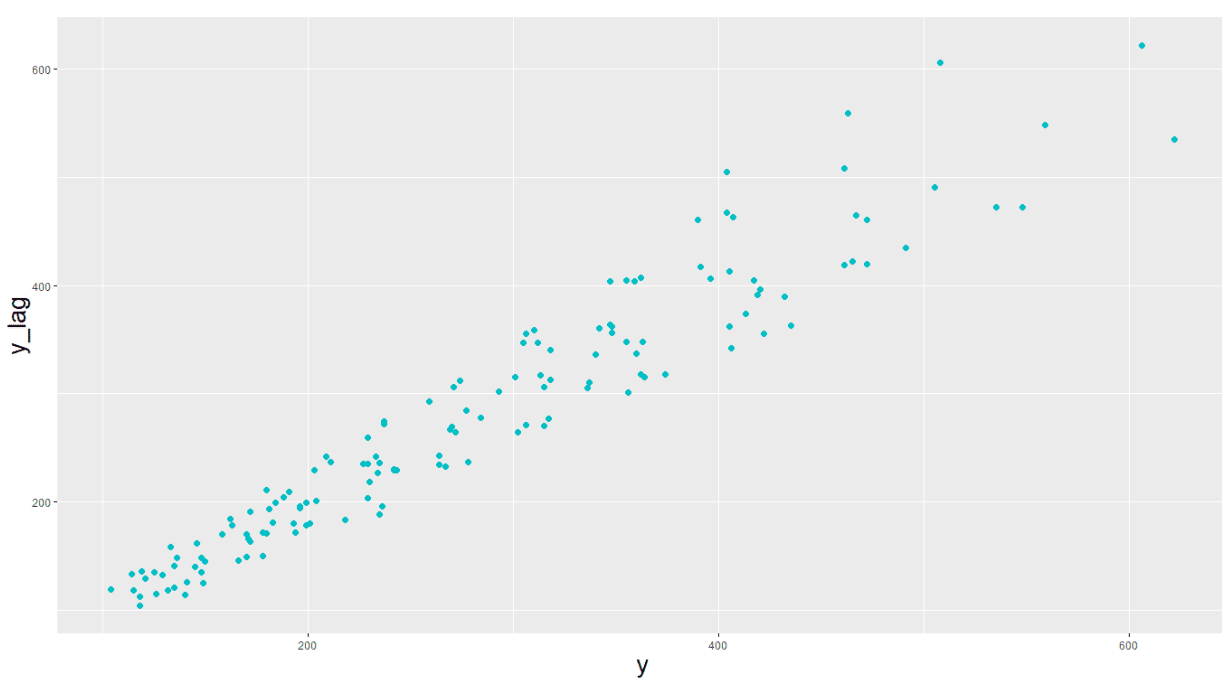


Si tuviéramos que predecir qué valor tiene la serie Y_t, una buena pista nos la da Y_{t-1}, es decir, justo el día de antes. En la mayoría de las series temporales es de esperar que el valor que toma la serie en un tiempo t es muy similar al que toma en un tiempo t-1.

Veamos cómo de cerca está Y_t de Y_{t-1}:

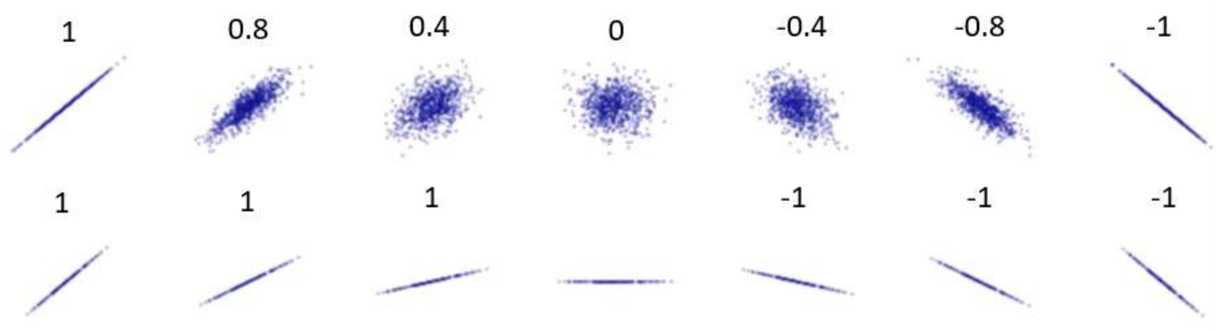


Como vemos, los valores de Y_{t-1} están muy cerca de Y_t. No obstante, este gráfico no es el más adecuado para ver si ambas series tienen relación una con otra. Cuando queremos analizar si dos variables tienen relación lineal una con la otra solemos recurrir a un gráfico de dispersión como el siguiente:



Si los puntos se distribuyen en línea recta indica que hay una fuerte relación lineal, y podemos cuantificar la fuerza de esta relación con el coeficiente de correlación lineal, que en este caso es **0.96**.

Cuanto más cerca esté este coeficiente de -1 o de 1, más fuerte es la relación lineal que existe entre dos variables (en nuestro caso, series temporales). Y cuando más cerca esté de 0, menos relación lineal existe. En la siguiente imagen ejemplificamos este punto:



Por este motivo, para modelizar una serie Y_t suele ser muy útil utilizar su **lag (o retardo) de orden 1:** Y_{t-1}. Matemáticamente expresamos esto con el siguiente modelo:

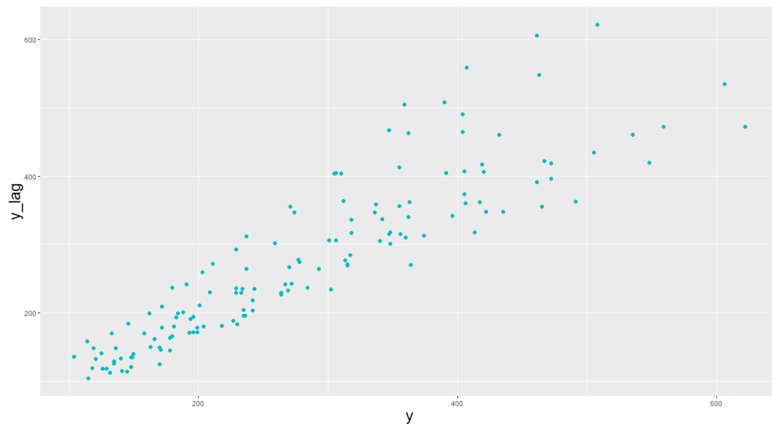
$$Y_t = \beta_1 \cdot Y_{t-1} + \varepsilon_t$$

Este modelo es lo que se conoce como un *AR* **de orden 1** (enseguida veremos a qué nos referimos por orden 1). De forma más general, si contemplamos un intercepto en la fórmula, tenemos:

$$Y_t = \beta_0 + \beta_1 \cdot Y_{t-1} + \varepsilon_t$$

Si recordamos bien el modelo de regresión lineal, este modelo *AR* **de orden 1 no es más que un modelo de regresión lineal, donde utilizamos como único regresor el retardo de orden 1 de la serie original.**

De la misma forma que Y_{t-1} suele ser útil para predecir valores de Y_t, también suele serlo Y_{t-2}. En el siguiente gráfico podemos ver la relación lineal que tienen Y_t y Y_{t-2}:



La correlación que tiene Y_t con su retardo de orden 2 (Y_{t-2}) es de: **0.89**. Entonces, ¿por qué no añadir esta variable a nuestro modelo? De la misma manera que hemos añadido Y_{t-1} a la ecuación, podemos añadir Y_{t-2} para conformar el que se denomina modelo *AR* **de orden 2**:

$$Y_t = \beta_0+\beta_1\cdot Y_{t-1}+\beta_2\cdot Y_{t-2}+ \varepsilon_t$$

Podríamos seguir añadiendo otros retardos: Y_{t-3}, Y_{t-4}, Y_{t-5}, Y_{t-6}... El problema es que llega un momento en el que los retardos dejan de tener correlación suficiente con la serie original y, por tanto, dejan de aportar valor predictivo al modelo. Por ello, como ocurre en muchas situaciones en la vida, hay que saber cuándo parar.

Por ejemplo, el retardo 100 de la serie que hemos visto tiene una correlación de 0.32 con la serie temporal original.

El orden de un modelo *AR* indica el máximo retardo hasta el que añadimos en el modelo de regresión propiamente dicho. **De forma más general describimos un modelo *AR* de orden p, también escrito como $AR(p)$, con la siguiente ecuación**:

$$Y_t = \beta_0+\beta_1\cdot Y_{t-1}+\beta_2\cdot Y_{t-2}+ \quad +\beta_p\cdot Y_{t-p}+ \varepsilon_t$$

Así nos iniciamos en los modelos ARIMA, aunque este modelo puede sofisticarse un poco más e incorporar la componente MA a la ecuación, algo que veremos a continuación.

MA

La motivación que surge tras el modelo *AR(p)* viene del término ε_t en la ecuación de regresión. Este término representa el residuo, es decir, la parte de Y_t que no somos capaces de entender (modelizar) utilizando los retardos: $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$.

En la práctica, cuando ajustamos un modelo *AR(p)* determinamos los mejores coeficientes β_x para explicar Y_t, estos valores fijos los solemos denotar por:

$$\widehat{\beta}_0, \widehat{\beta}_1, \ldots \widehat{\beta}_p$$

Y cuando queramos predecir el valor que va a tomar Y_t en el instante t+1, simplemente aplicaremos la fórmula que tenemos con los valores que conocemos. Es decir, si:

$$Y_t = \beta_0+\beta_1\cdot Y_{t-1}+\beta_2\cdot Y_{t-2}+ \quad +\beta_p\cdot Y_{t-p}+ \varepsilon_t$$

Entonces...

$$Y_{t+1} = \beta_0+\beta_1\cdot Y_t+\beta_2\cdot Y_{t-1}+ \quad +\beta_p\cdot Y_{t-p+1}+ \varepsilon_t$$

De esta fórmula en el momento de la predicción conocemos los valores de $Y_t, Y_{t-1}, \ldots, Y_{t-p+1}$, y el modelo de regresión nos dará los mejores valores de los coeficientes β_x.

Con estos ingredientes podemos realizar la predicción de Y_{t+1} que denotaremos por $\hat{Y}_{t+1}$ como:

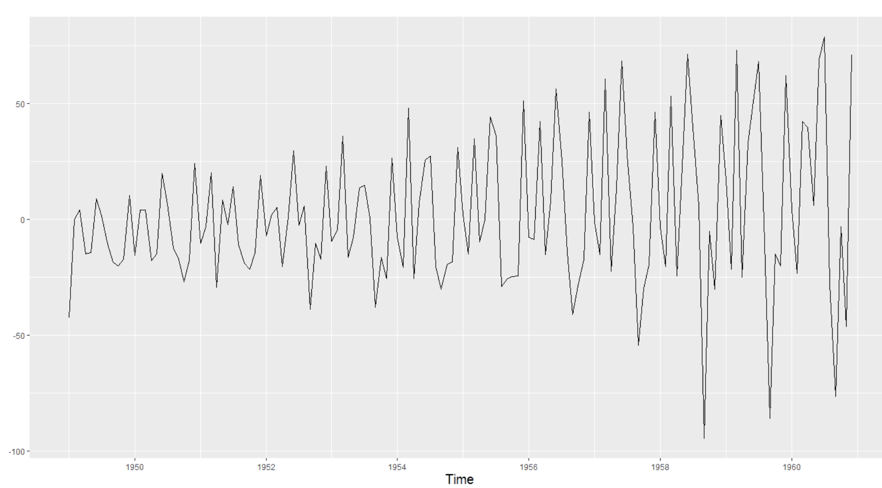
$$\hat{Y}_{t+1} = \hat{\beta}_0 + \hat{\beta}_1 \cdot Y_t + \hat{\beta}_2 \bullet Y_{t-1} + \cdots + \hat{\beta}_p \bullet Y_{t-p+1}$$

Esto ya es un valor fijo, lo que ocurre es que no será exactamente igual que el verdadero valor que tome nuestra serie en Y_{t+1}. La diferencia será el término de error ε_{t+1}.

De la misma manera, si aplicamos la fórmula con los coeficientes estimados en todos los instantes posibles de la serie que conocemos, podemos construir la serie del residuo ε_t completa.

Es igual a cuando veíamos el modelo de descomposición: hay una parte de la serie que en todo momento podemos modelizar y conocer (Tendencia + Estacionalidad), y el resto se irá al Residuo.

En el ejemplo anterior que estábamos viendo, si ajustamos un AR de orden 2, tenemos la siguiente serie de error ε_t:



Salta a la vista que la serie de error tiene un patrón repetitivo. No es una serie aleatoria que sea 100% imposible de predecir. Si cogiésemos esta serie del residuo, podríamos tratar de ajustar de nuevo un modelo AR sobre ella, por ejemplo. La idea de la componente MA sería: **tratar de modelizar el residuo**. Podemos ver esta componente siempre como la parte en la que el modelo aprende de sus errores.

Dado un modelo *AR(p)*:

$$Y_t = \beta_0+\beta_1 \cdot Y_{t-1}+\beta_2 \cdot Y_{t-2}+ \quad +\beta_p \cdot Y_{t-p}+ \varepsilon_t$$

La idea es que obtengamos la serie del residuo como:

$$Y_t - \beta_0+\beta_1 \cdot Y_{t-1}-\beta_2 \cdot Y_{t-2}- \quad -\beta_p \cdot Y_{t-p}+ \varepsilon_t$$

Y apliquemos de nuevo un modelo *AR(q)* sobre la serie del residuo:

$$\varepsilon_t = \quad _0+ \quad _1 \cdot \varepsilon_{t-1}- \quad _2 \cdot \varepsilon_{t-2}- \quad - \quad _p \cdot \varepsilon_{t-p}+ Y_t$$

Y al final acabamos con un residuo del residuo denotado por Y_t, y esta sería de nuevo la parte que nos queda sin explicar de nuestra serie Y_t. Por lo tanto, **un modelo MA de orden q, también escrito *MA(q)*, es un modelo *AR(q)* sobre los residuos de un modelo *AR(p)* ajustado.**

Ambos modelos, *AR(p)* y *MA(q)*, suelen ir casi siempre de la mano. Primero ajustamos un *AR(p)*, estudiamos el t e intentamos modelizar de nuevo el residuo aplicando un *MA(q)*, que sería como aplicar de nuevo un *AR(q)* sobre esta nueva serie.

Cuando empleamos las dos componentes, *AR(p)* y *MA(q)*, hablamos de un modelo *ARMA* de **órdenes** *p* **y** *q*, también escrito como *ARMA(p, q)* y que podemos describir con la siguiente ecuación:

$$Y_t - \beta_0 + \beta_1 \cdot Y_{t-1} + \quad + \beta_p \cdot Y_{t-p} + \quad {}_1 \cdot b_{t-1} + \ldots + \quad {}_q \cdot b_{t-q} + \varepsilon_t$$

Es de nuevo una regresión donde utilizamos los retardos de la serie Y_{t-1}, Y_{t-2}, $\ldots Y_{t-p}$ como regresores junto con los retardos del residuo que tenemos tras ajustar el primer *AR(p)* y que denotamos por b_{t-1}, $\quad b_{t-q}$.

Acabamos de construir el modelo ARMA *(p,q)* y ahora solo nos queda por conocer la componente I para terminar de construir el modelo ARIMA.

Diferencias

Los modelos ARIMA son, sin duda, los más famosos de toda la literatura de series temporales por todo el tiempo que han estado empleándose y por los buenos resultados predictivos que proporcionan. No obstante, esta familia de modelos, como cualquier otra, no puede predecir cualquier serie temporal (p. ej., las series de stocks financieros).

Estas series financieras tienen tantos efectos de variables exógenas a los propios mercados financieros que hace muy difícil predecir sus valores.

Para entender mejor este aspecto pensemos, por ejemplo, en el precio del petróleo: un tuit del presidente de los Estados Unidos puede provocar que el precio del petróleo suba o baje drásticamente por la confianza o desconfianza que genere este comentario en los brokers de los mercados (ni hablar ya si el tuit es de Donald Trump, ya expresidente).

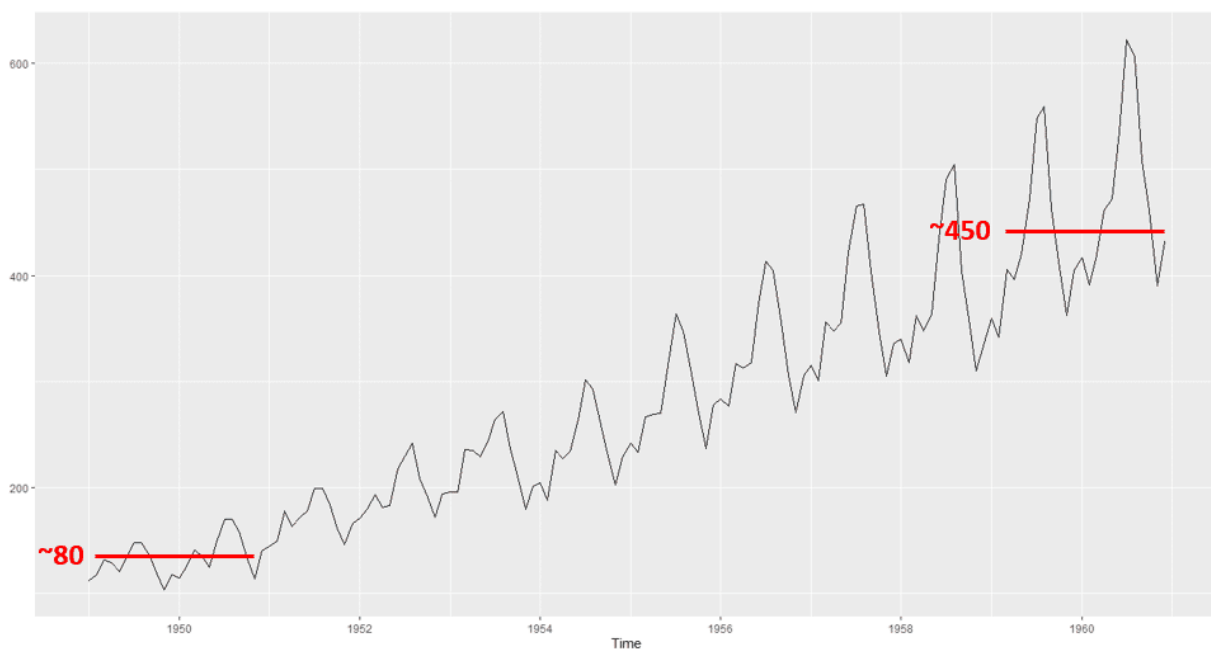
De modo que la variable 'hay tuit del presidente de los EE. UU.' puede afectar al precio del petróleo. Podríamos considerarla como variable regresora en nuestros modelos, pero ¿qué vamos a hacer? ¿Vamos a predecir cuándo va a hacer Donald Trump un tuit relevante al mercado de materias primas? Esto, como muchos otros efectos exógenos, se nos escapa de las manos.

Por esta razón, los modelos ARIMA requieren que, para ser efectivos, la serie que estemos tratando de modelizar sea lo que se denomina **estacionaria**.

Para que lo entendamos con facilidad, una serie se dice que es estacionaria cuando es **estacionaria en media** y **estacionaria en varianza**.

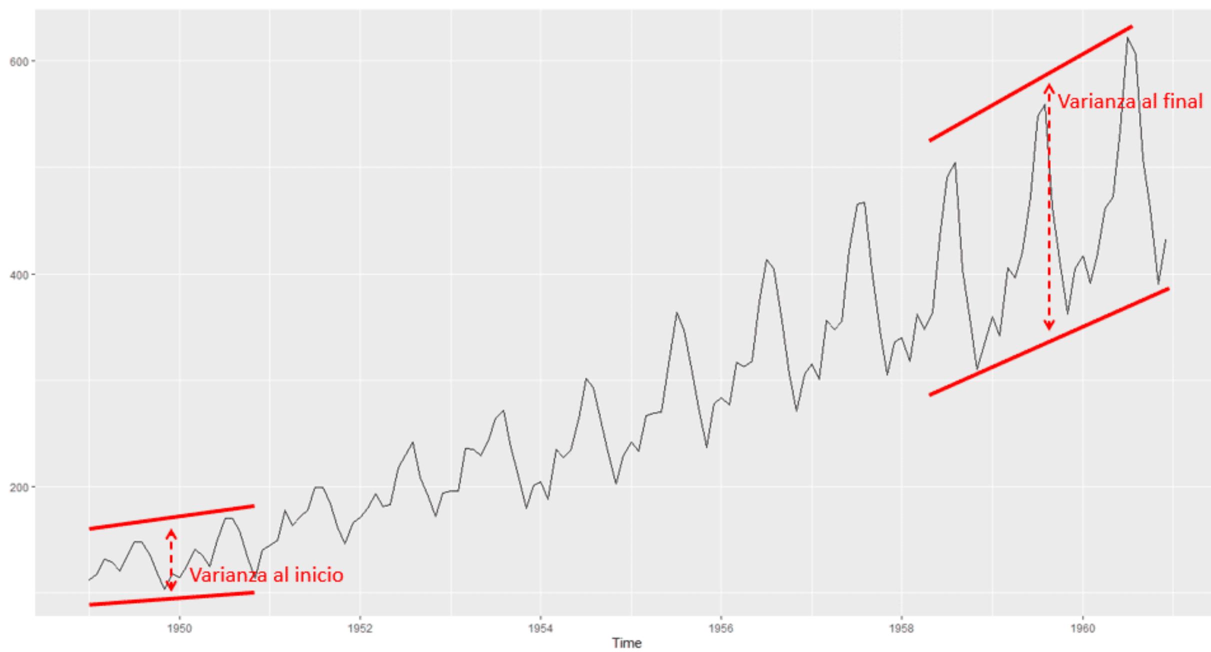
¿Cuándo una serie es estacionaria en media?

Cuando durante todo el histórico la serie mantiene una altura media que es estable todo el tiempo en torno a un mismo valor. Por ejemplo, la serie que hemos visto antes no es estacionaria: al principio del histórico la serie tiene un nivel medio cercano a 80, mientras que al final del histórico la serie tiene un nivel medio cercano a 450.
Para que la serie sea estacionaria en media durante todo el histórico debe tener un nivel medio similar.



¿Cuándo una serie es estacionaria en varianza?

Cuando durante todo el histórico la serie tiene una varianza similar, la serie está variando todo el tiempo entre los mismos valores. De nuevo, la serie de nuestros ejemplos tampoco es estacionaria en varianza. Al principio del histórico se mueve entre unos valores mucho más pequeños que al final del histórico:



El objetivo antes de ajustar un modelo ARIMA es asegurar que nuestra serie es estacionaria en media y en varianza, y para ello se realizan transformaciones sobre la serie.

Estas transformaciones pueden consistir en distintos reajustes:

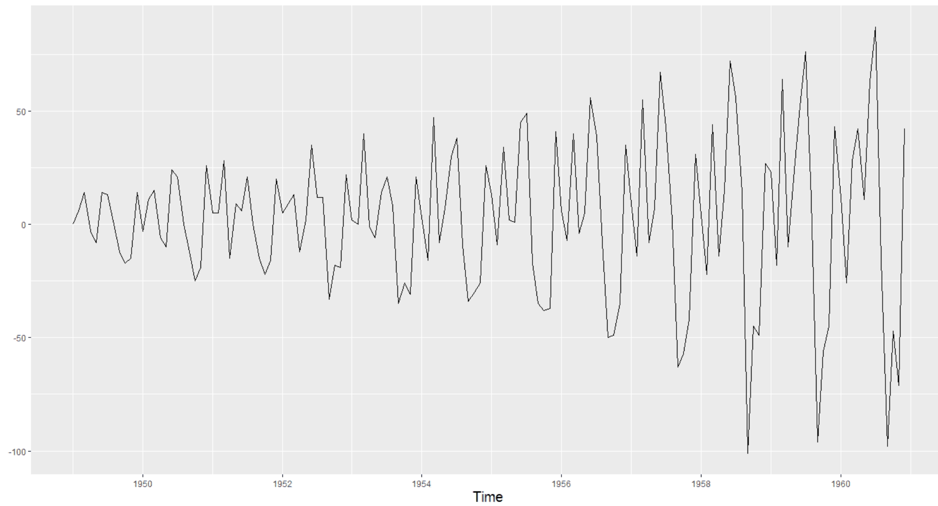
- Reajustes por calendario.
- Reajustes por población.
- Reajustes por efectos económicos.
- Reajustes por transformaciones matemáticas.

Pero en caso de no ser suficiente, en los modelos ARMA introducimos la componente **I** para dar lugar a los modelos ARIMA, donde esta componente indica **aplicar una diferencia a la serie**. Esto es hacer la diferencia de Y_t con un retardo suyo Y_{t-d}. Si aplicamos una diferencia, también escrito como $I(1)$, estaríamos restando a Y_t su retardo de orden 1: Y_{t-1}:

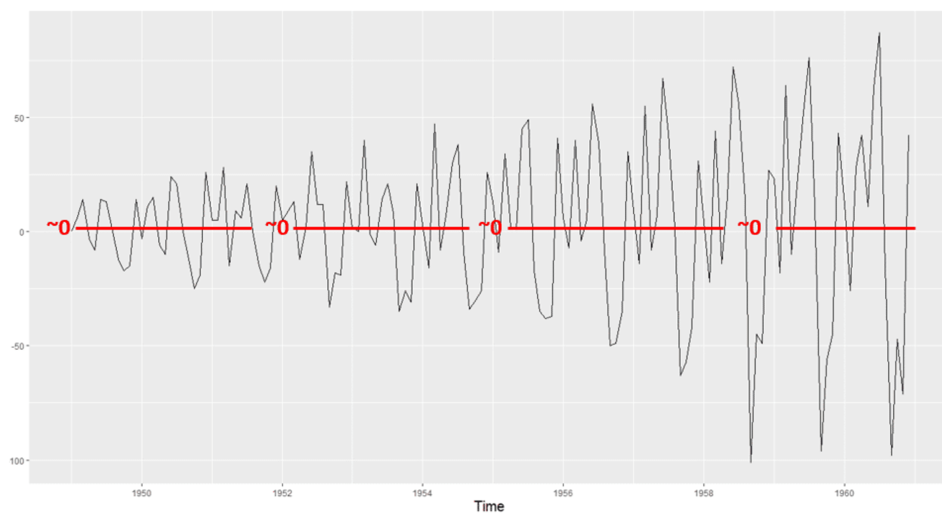
$$Z_t = Y_t - Y_{t-1}$$

Y pasaríamos a modelizar Z_t en el caso de ser estacionaria.

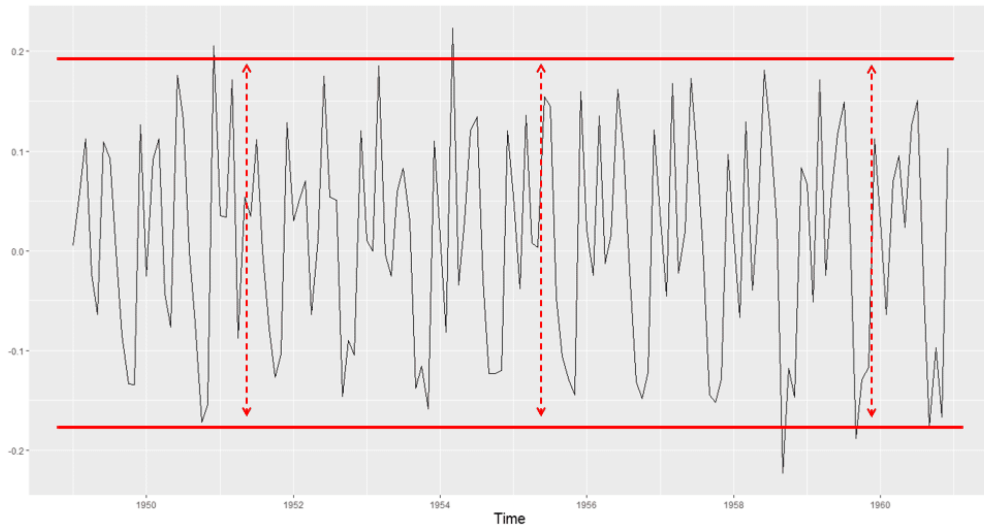
¿Qué pasa si en nuestro ejemplo aplicamos una diferencia de orden 1? Pues que la serie toma la siguiente forma:



Como vemos, la serie transformada ahora es estacionaria en media. En todos los periodos que escojamos la media es la misma:



Pero la serie sigue sin ser estacionaria en varianza, por lo que no podemos decir todavía que sea estacionaria del todo. ¿Qué tal si aplicamos el logaritmo a la serie antes de realizar la diferencia de orden 1? La serie que obtenemos como resultado es la siguiente:



Ahora ya tenemos una serie estacionaria, tanto en media como en varianza. Es sobre estas series sobre las que las componentes AR y MA de un modelo ARIMA tratan de extraer patrones con los que modelizar la serie.

Recapitulando

Para que las componentes AR y MA de un modelo ARMA puedan modelizar correctamente la serie que queremos, hay que **asegurar que esta sea estacionaria**, tanto en media como en varianza.
Si nuestra serie de partida ya es estacionaria por naturaleza, podemos aplicar un modelo ARMA sin más. Pero, en caso de que no sea estacionaria, hay que hacerla estacionaria, y para ello aplicamos transformaciones sobre la serie. Normalmente se cumple la siguiente regla:

- Si la serie no es estacionaria en media, aplicamos una diferencia de orden d.
- Si la serie no es estacionaria en varianza, aplicamos una transformación matemática como el logaritmo, la raíz cuadrada o Box-Cox.

La componente $I(d)$ en un modelo ARIMA indica cuántas veces aplicamos una diferencia de orden 1 a la serie Y_t. Matemáticamente esto lo describimos como:

$$B^d(Y_t) = Z_t^d$$

Y el resultado es una serie Z_t^d que es el resultado de repetir el siguiente proceso d veces:

$$Z_t^i = Z_t^{i-1} - Z_{t-1}^{i-1}$$

Donde

$$Z_t^1 = Y_t - Y_{t-1}$$

Con estos dos pasos nuestras series pasarán a ser estacionarias. En caso de no ser capaces de hacer nuestra serie estacionaria tras estas transformaciones, los modelos ARIMA no están diseñados para modelizar dicha serie y debemos esperar malos resultados en términos predictivos.

Modelo ARIMA

Una vez hemos estudiado las tres componentes que componen los modelos ARIMA, estamos en posición de **formular el modelo completo**. Un modelo ARIMA lo más general posible se escribe con la siguiente formulación:

$$Z_t^d = \beta_0 + \beta_1 \cdot Z_{t-1}^d + \ldots \beta_p \cdot Z_{t-p}^d + \alpha_1 \cdot b_{t-1} + \cdots + \alpha_q \bullet b_{t-q} + \varepsilon_t$$

Donde la serie original se ha diferenciado d veces:

$$Z_t^d = B^d(Y_t)$$

Que no nos asuste esta formulación, ya que es exactamente lo que hemos visto en las secciones anteriores, pero todo junto y con una nomenclatura matemática. Vamos a verlo parte por parte.

La base del modelo ARIMA es de nuevo una especie de regresión, solo que en primer lugar el modelo puede tener d diferencias aplicadas a la serie original Y_t, por tanto, la regresión la escribimos en términos de Z_t^d, no de Y_t, como vimos en el modelo ARMA anteriormente.

Por tanto, el primero paso es olvidarnos de Y_t y pensar en la serie diferenciada d veces Y_t.

El resto es el modelo ARMA que conocemos: tenemos en primer lugar los retardos de la serie Z_t^d, para que el modelo aprenda de lo que ha ocurrido en el pasado (Z_{t-1}^d, ... Z_{t-p}^d), y los retardos del residuo intermedio, que obtenemos tras ajustar la componente AR(p) (b_{t-1}, ...b_{t-q}) con los que el modelo aprende de sus errores.

Con todo esto, tenemos formulado un modelo ARIMA de órdenes p, d y q, que también se escribe como ARIMA $(p,\ d,\ q)$.

Lección 3 de 4

Ajuste de un modelo ARIMA

Edix Educación

Para ajustar un ARIMA nos apoyaremos en R o Python: podremos ajustar los modelos y predecir, como hemos hecho anteriormente con los otros modelos. La diferencia que tenemos ahora con los modelos ARIMA es que estos cuentan con varios hiperparámetros: p,d y q.

En ciencia de datos diferenciamos entre **parámetros** e **hiperparámetros**.

Un **parámetro** es un valor que los modelos son capaces de aprender de los datos.

En un modelo de regresión, cada uno de los β_x es un parámetro. El modelo es capaz de ajustar los valores que toma cada uno de ellos tras realizar el entrenamiento con los datos. Los valores *p, d* y *q* no se aprenden de los datos con los que vamos a modelizar, el modelo ARIMA necesita que nosotros le digamos qué valores van a tomar cada uno de ellos.

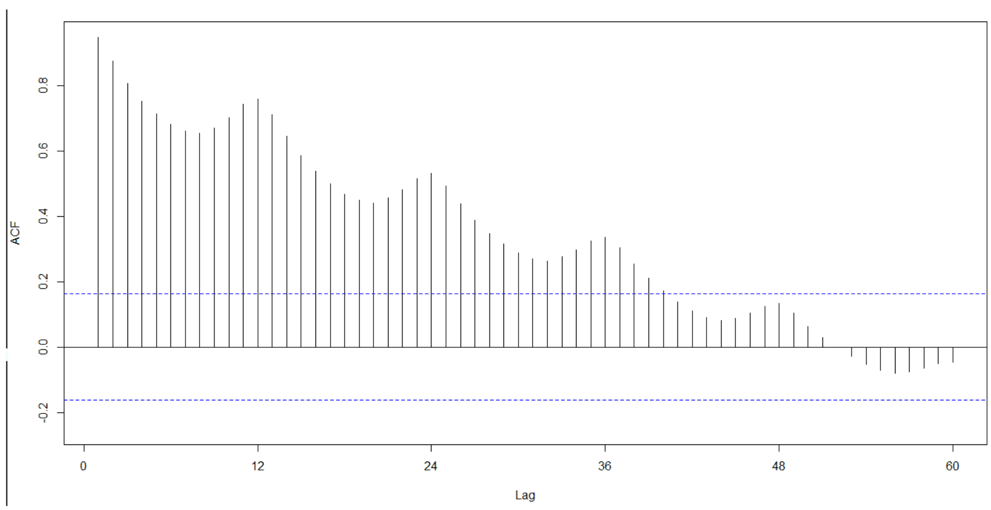
La elección de los mejores *p, d* y *q* no es una tarea fácil, especialmente porque no hay un camino directo con el que llegar al mejor resultado. Terminamos con distintas combinaciones de *p, d* y *q*, que acaba cada una de ellas en un modelo distinto, para después quedarnos con el mejor.

Independientemente, vamos a explicar las dos principales herramientas que se emplean para estudiar qué valores proponer de cada uno de estos hiperparámetros.

ACF

El ACF, **Auto-Correlation Function**, es un gráfico en el que estudiamos la correlación que tiene nuestra serie Y_t con cada uno de sus posibles retardos: Y_{t-1}, Y_{t-2}, Y_{t-3}...

En el eje X tenemos el orden del retardo que estamos utilizando y en el eje le indicamos la correlación. Nuestro ejemplo presenta el siguiente ACF:



Como vemos, la correlación con los retardos más cercanos es muy alta y va disminuyendo cuanto más alto es el retardo que estamos aplicando. Las líneas azules discontinuas indican si la correlación es estadísticamente significativa o no. Si la correlación se sale de la banda, es significativa, y no lo es en caso contrario. Como vemos, las correlaciones son significativas hasta el retardo 41. A partir de aquí ya no son significativas.

Interpretar este gráfico nos sirve para proponer posibles buenos valores del parámetro q. Veremos en detalle en el siguiente fastbook práctico cómo hacer esto en la práctica.

PACF

Por otro lado, tenemos la PACF, **Partial Auto-Correlation Function**, que es algo más compleja que la ACF. La idea que hay detrás de este gráfico es que cada retardo tenga un valor que indique por sí solo cuánto valor es capaz de informar sobre la serie Y_t sin recibir información de otros retardos. En otras palabras:

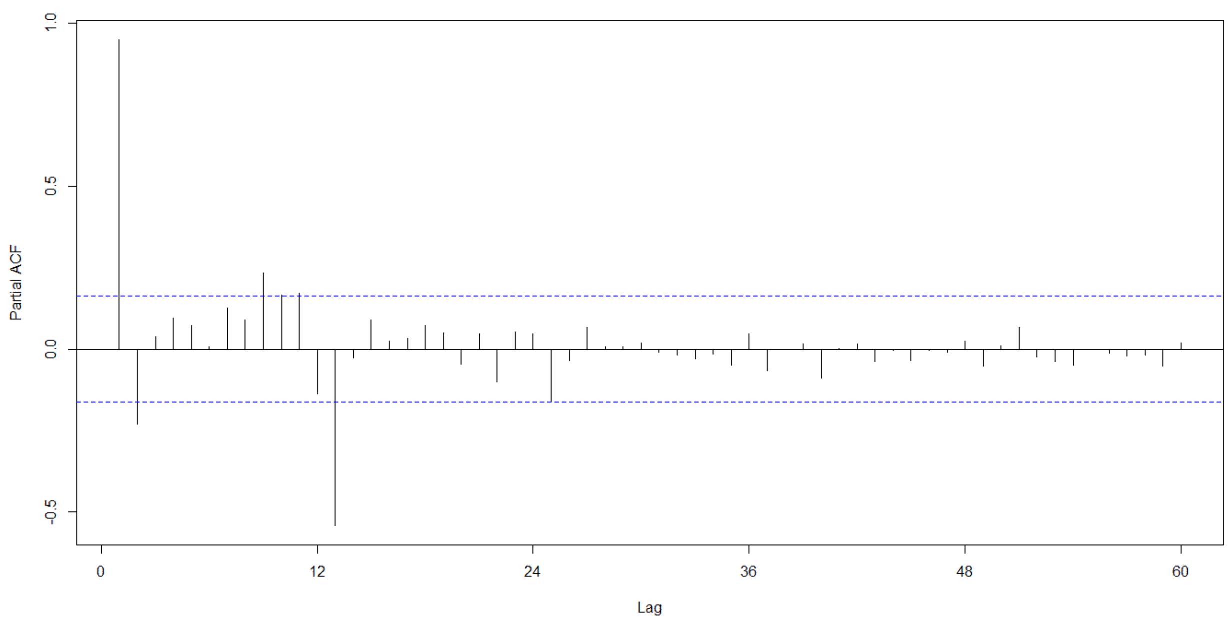
Si Y_t se correlaciona mucho con Y_{t-1}, esto hace que Y_{t-1} sea una buena variable para explicar Y_t. Si Y_{t-2} se correlaciona con Y_{t-1}, como Y_{t-1} se correlaciona con Y_t, esto va a hacer que Y_{t-2} también se correlacione con Y_t, pero lo hace gracias a Y_{t-1}, no porque ella tenga información sobre Y_t que no haya aportado Y_{t-1}.

Supongamos que Y_t es nuestro abuelo, Y_{t-1} es nuestra madre y Y_{t-2} somos nosotros. Nuestra madre conoce muy bien a nuestro abuelo (se correlacionan mucho), ha pasado muchos años más junto a él y sabe de qué equipo de fútbol es, cuál es su plato favorito, la película que más ha visto, etc.

Nosotros también conocemos muchas cosas de nuestro abuelo (nos correlacionamos mucho con él), algunas de ellas es posible que no las conozca nuestra madre, pero, sin duda, la mayoría de lo que sabemos es por lo que nos ha contado nuestra madre de él. Por tanto, no es 'justo' cuantificar nuestra correlación con el abuelo igual que lo hace nuestra madre.

De la misma manera ocurre con las variables, y el gráfico PACF trata de corregir la correlación que tiene cada retardo Y_{t-d} con la serie Y_t eliminando toda la información de Y_t que llega a cada retardo Y_{t-d} por que lo han ido filtrando los retardos $Y_{t-1}, \ldots Y_{t-d+1}$.

En nuestro ejemplo, la PACF es la siguiente:



Igual que con la PACF, las líneas azules discontinuas nos indican si las correlaciones parciales son estadísticamente significativas o no.

Este gráfico nos servirá para proponer valores para el hiperparámetro p y será en el siguiente fastbook donde veremos de forma práctica cómo hacerlo.

Resumen y conclusiones

Edix Educación

El modelo ARIMA es el modelo más avanzado que hemos estudiado. Es un modelo que ofrece muy buenos resultados predictivos sobre series estacionarias. Por tanto, nuestro primer cometido debe ser convertir nuestra serie no estacionaria en estacionaria, para que el modelo funcione correctamente.

Para hacer nuestra serie estacionaria, hemos visto dos tipos de transformaciones que nos suelen ayudar:

- Transformaciones matemáticas para hacer nuestra estacionaria en varianza.
- Diferencias para hacer nuestra serie estacionaria en media.

Del segundo tipo de transformaciones surge la componente *I* de los modelos ARIMA.

Una vez nuestra serie es estacionaria, el modelo ARIMA aplica las componentes AR y MA.

La primera para aprender patrones del pasado de la serie, y la segunda para aprender de los errores que se cometen al predecir en el pasado.

La gran pregunta es de qué orden hacer las diferencias (¿*d*?), hasta qué orden meter retardos de la serie (¿*p*?) y hasta qué orden meter retardos del residuo (¿*q*?).

Dedicaremos el último fastbook a responder estas preguntas y cerrar los modelos ARIMA, que será de forma práctica con R.

¡Enhorabuena! Fastbook superado

edix

Creamos Digital Workers