BOURBAKI

COLEGIO DE MATEMÁTICAS

BOURBAKI
COLEGIO DE MATEMÁTICAS

Índice

01 Introducción

Las siguientes notas son la bitácora de un curso de 8 horas que impartimos junto a Ana Isabel Ascencio. Además de este documento los invitamos a consultar el Github del curso en este link.

El curso es una invitación al forecasting utilizando series de tiempo y el cálculo de riesgo, nuestros ejemplos principales serán los algoritmos ARMA, ARIMA y ARCH, la organización de las clases es la siguiente:

1. Introducción al problema de forecasting y descripción del conjunto de datos (treinta minutos)
2. Procesos estocásticos y momentos (noventa minutos).
3. Visualización y simulación de procesos estocásticos en R (una hora).
4. ARCH, VaR y el cálculo del riesgo (una hora).
5. Implementación de ARIMA y (ARIMA+ARCH) (dos horas).
6. Dudas y complementos (dos horas).

Series de tiempo y el cálculo del riesgo

Cuando buscamos hacer un forecast a lo largo del tiempo, los modelos de series temporales son más adecuados que otros como por ejemplo las regresiones lineales, existen dos razones fundamentales para que esto suceda, la primera es porque incluyen componentes adicionales por ejemplo para la temporalidad de un modelo. La segunda razón y quizás la más importante es porque permiten un ruido considerablemente más complicado que el usual, a saber independiente e idénticamente distribuido. En este curso haremos énfasis en las ventajas que esto significa para una predicción, especialmente cuando lo que buscamos predecir es el riesgo.

Preludio sobre procesos estocásticos

En esta sección hablaremos únicamente de procesos estocásticos, por el momento sin hacer mención a las series de tiempo. Más adelante incluiremos el estudio de estos procesos en el de las series del tiempo, por el momento recomendamos al estudiante que piense en estos procesos estocásticos como los errores al hacer una predicción de una serie de tiempo.

Definition 02.1. Un proceso estocástico es una familia de variables aleatorias $\{\dots, X_{-2}, X_{-1}, X_0, X_1, X_2, \dots, X_T, \dots\}$.

Example 02.1. *Sea $(X_i)_{i\in\mathbb{Z}}$ un proceso estocástico,*

- *Un proceso independiente e idénticamente distribuido es un proceso estocástico tal que $X_i \sim \mathbb{P}$ para cualquier $i \in \mathbb{Z}$ y para cualesquiera $i \neq j \in \mathbb{Z}$, las variables aleatorias X_i, X_j son independientes.*

- *Un proceso es un ruido blanco si para cualquier* $i \in \mathbb{N}$ $Var[X_i] = \sigma^2$, $\mathbb{E}[X_i] = 0$ *y para cualesquiera* $i \neq j \in \mathbb{Z}$, $Cov(X_i, X_j) = 0$.
- *Si* $(X_i)_{i\in\mathbb{Z}}$ *es un proceso independiente e idénticamente distribuido, entonces* $(S_i)_{i\in\mathbb{Z}}$ *definido de la siguiente manera* $S_i = X_1 + X_2 + \ldots + X_i$ *es una marcha aleatoria (y un ejemplo de una cadena de markov).*
- *Un proceso estacionario es un proceso tal que:*
  1. *Para todo* $t \in \mathbb{Z}$, $\mathbb{E}\left[X_i^2\right] < \infty$,
  2. *Existe un* $\mu_t \in \mathbb{R}$ *tal que* $\mathbb{E}[X_i] = \mu_X$,
  3. *Existe una función* $\gamma_X : \mathbb{Z} \to \mathbb{R}$ *tal que para todo* $i, j \in \mathbb{Z}$, $Cov(X_i, X_{i+j}) = \gamma_X(j)$.

Remark 02.2. *Notemos que un proceso estacionario tiene varianza constante al igual que un ruido blanco.*

Example 02.3. *Es importante notar la diferencia que existe entre los procesos independientes e idénticamente distribuidos y el ruido blanco. Si bien es cierto que todo proceso i.i.d. es un ruido blanco, lo contrario no es cierto tal y como lo muestra los siguientes ejemplos:*

1. *Sea* (X_t) *un proceso i.i.d. e* Y *una variable aleatoria de Bernoulli con parámetro* $p = \frac{1}{2}$. *Entonces la variable aleatoria* $Z_t = X_t Y - X_t (Y-1)^2$ *es un ruido blanco.*
2. *Sea* (X_t) *un proceso i.i.d. el proceso estocástico definido por* $Z_t = X_t$ *cuando* t *es par y* $Z = \frac{1}{\sqrt{2}}\left(X_{t-1}^2 - 1\right)$ *cuando* t *es impar, es un ruido blanco.*

02 Instalación de R, R Studio y Paquetes requeridos ARIMA y ARCH

Para este curso utilizaremos el lenguaje de programación R a través del entorno R Studio, así como un conjunto de paquetes que nos permitirán realizar el manejo, preprocesamiento y modelado de los datos.

Instalación de R y Rstudio

01.1 Para Windows

1. Descargar R desde el repositorio CRAN[1], haciendo click aquí. Cuando termine la descarga ejecutar el archivo R-4.0.0-win.exe, aceptar todas las opciones por defecto hasta que se complete la instalación.

2. Descargar el archivo .exe desde Rstudio dando click aquí. Igualmente cuando termine la descarga, ejecutar el archivo .exe y aceptar las opciones por defecto hasta completar la instalación

[1] Administrado por R fundation. CRAN es el más grande repositorio de paquetes de R, sus siglas provienen del inglés: the comprehensive R archive network.

01.2 Para Mac

1. Descargar R desde el repositorio CRAN, haciendo click aquí. Cuando termine la descarga ejecutar el archivo R-4.0.0.pkg, aceptar todas las opciones por defecto hasta que se complete la instalación.
2. Descargar el archivo .dmg desde Rstudio, dando click aquí. Igualmente cuando termine la descarga, ejecutar el archivo .dmg y aceptar las opciones por defecto hasta completar la instalación

01.3 Para Ubuntu

1. Previo a la instalación, en la terminal, actualiza los paquetes instalados mediante: sudo apt update sudo apt -y update
2. Enseguida instala R mediante sudo apt -y install r-base
3. Descargar el archivo .deb desde Rstudio , correspondiente a la versión de Ubuntu.
4. Cuando termine la descarga, desde la terminal ubicate en la carpeta de tus descargas y corre la siguiente línea: sudo dpkg -i rstudio-1.2.5033-amd64.deb
5. Si encuentras problemas de dependencias corre la siguiente linea y después vuelve al paso anterior sudo apt -f install

01.4 Para otras versiones de Linux

1. Descargar R desde el repositorio CRAN, siguiendo las instrucciones de acuerdo a la distribución que de Lunux que se tenga.
2. Descargar el archivo correspondientes según la distribución de Linux desde Rstudio

03 Series de tiempo y el cálculo de riesgo

Desde el punto de vista de las aplicaciones las series de tiempo son modelos matemáticos más flexibles que las regresiones usuales pues incluyen diversos componentes tales como la temporalidad.

Definition 00.1. Una serie de tiempo uni-variada es una familia finita de observaciones $S = \{y_I, y_{I+1}, \ldots, y_0, y_1, \ldots, y_T\} \subseteq \mathbb{R}$ tales que:

1. Existe una función llamada la tendencia $f : \mathbb{Z} \to \mathbb{R}$
2. Una función llamada la temporalidad $s : \mathbb{Z} \to \mathbb{R}$ llamada de temporalidad con periodo $r \in \mathbb{N}$ que satisface $s(t) = s(t+r)$ para cualquier $t \in \mathbb{Z}$.
3. Un proceso estocástico $(X_t)_{t\in\mathbb{Z}}$

Tal que para todo $t \in [I, T]$,

$$y_t = f(t) + s(t) + X_t$$

Desde un punto de vista matemático también son súmamete interesantes pues involucran procesos estocásticos más interesantes que los independientes e idénticamente distribuidos que utilizamos para las regresiones, tales como los hemos visto en la sección 02.

Eliminación de la tendencia y temporalidad utilizando diferenciación

A pesar de que una serie de tiempo en toda su generalidad puede contener una componente de tendencia y otra de temporalidad, en este curso pronto nos deshaceremos de ellas por razones provenientes de una intuición financiera, tal como lo explicaremos más adelante.

Definition 01.1. Sea $S = \{y_I, y_{I+1}, \dots, y_0, y_1, \dots, y_T\} \subseteq \mathbb{R}$ una serie de tiempo, definimos inductivamente el operador diferencial de orden $d \in \mathbb{N}$ aplicado a la serie S como una nueva serie $\nabla^d(S)$ definida de la siguiente manera:

$$\nabla(y_t) = y_t - y_{t-1}$$

$$\nabla^d(y_t) = \nabla\left(\nabla^{d-1}(y_t)\right)$$

Remark 01.1. *En finanzas se define el returno diario simple de una serie de tiempo S como*

$$\frac{\nabla(y_t)}{y_{t-1}}$$

La siguiente proposición nos permitirá reducir el estudio de nuestras series de tiempo S únicamente a su proceso estocástico X_t.

Proposition 01.2. *Sea $S = \{y_I, y_{I+1}, \dots, y_0, y_1, \dots, y_T\} \subseteq \mathbb{R}$ una serie de tiempo tal que su función de tendencia es polinomial de grado k y su temporalidad es* *periódica de periodo* r*, si* $d = max(k, r)$*, entonces la serie de tiempo* $\nabla^{d-1}(S)$ *no tiene ni temporalidad ni tendencia.*

Remark 01.3. *Si bien es cierto que no todas las tendencias y temporalidades requieren ser polinomiales, existen diversos resultados que garantizan la aproximación de funciones mediante polinomios y series periódicas.*

Geometría de las series de tiempo y funciones de autocorrelación

---

Definition 02.1. Sea $(X_i)_{i\leq N}$ un proceso estacionario, definimos la función de auto-covarianza a la función γ_X. La función de auto-correlación se define como $\rho_X(t) = \frac{\gamma_X(t)}{\gamma_X(0)}$.

Exercise 02.1. *Si* (X_t) *es un ruido blanco, entonces* $\rho_X(t) = 0$ *si* $t \neq 0$ *y* $\rho_X(0) = \sigma^2$.

Además de la función de auto-correlación, en este curso utilizaremos la auto-correlación parcial para obtener información de una serie de tiempo, a continuación haremos algunas observaciones importantes sobre la geometría de las series de tiempo.

Definition 02.2. 1. Sean X, Y dos variables aleatorias, diremos que son ortogonales cuando $\mathbb{E}[XY] = 0$.

2. Sean X, Y dos variables aleatorias, definimos su distancia como $d(X,Y) = \sqrt{\mathbb{E}\left[(X-Y)^2\right]}$

3. Sean $X_1, X_2, \ldots, X_T, Y$ variables aleatorias, definimos la proyección ortogonal de Y en $(X_1, \ldots, X_T)$ como la variable aleatoria

$$\Pi_{X_1,\ldots,X_T}(Y) = \lambda_1 X_1 + \ldots + \lambda_T X_T$$

que para cualesquiera $\lambda'_1, \ldots, \lambda'_T$ satisface:

$\mathbb{E}\left[\Pi_{X_1,\ldots,X_T}(Y)\right] \leq \mathbb{E}\left[\lambda'_1 X_1 + \ldots + \lambda'_T X_T\right]$

Proposition 02.2. *Sean $X_1, X_2, \ldots, X_T, Y$ variables aleatorias, entonces todas las variables X_i son ortogonales a $Y - \Pi_{X_1,\ldots,X_T}(Y)$.*

Definition 02.3. Sean $X_1, X_2, \ldots, X_T, Y$ variables aleatorias, el coeficiente de correlación parcial entre X_1 y X_T se define como

$$r_{X_2,\ldots,X_{T-1}}(X_1, X_T) = \rho\left(X_1 - \Pi_{X_2,\ldots,X_{T-1}}(X_1), X_T - \Pi_{X_2,\ldots,X_{T-1}}(X_T)\right)$$

Proposition 02.3. *Si $X_1, \ldots, X_T$, entonces la función de auto-correlación parcial de orden t se define como $r(t) = r_{X_2,\ldots,X_t}(X_1, X_{t+1})$ y satisface*

- *Solo depende de t*
- *Para $t \geq 2$, $\rho(1) = r(1)$*
- $r(t) = r(-t)$

Estimación de momentos

Hasta el momento hemos estudiado a las series de tiempo utilizando exclusivamente variables aleatorias, sin embargo en la práctica solo tendremos acceso a simulaciones de las variables aleatorias, es decir a un conjunto de observaciones $\{x_1, x_2, \ldots, x_T\}$. Gracias a la ley de los grandes números y a la hipótesis de estacionaridad es posible aproximar la mayoría de los momentos de las variables aleatorias que generan nuestras observaciones.

Definition 03.1. Sea $\{x_1, x_2, \ldots, x_T\}$ un conjunto de observaciones generadas por las variables aleatorias $X_1, \ldots, X_T$, definimos las siguientes cantidades:

- El promedio empírico $\hat{x} = \frac{1}{T} \sum_{i \le T} x_i$
- La varianza empírica $\hat{\sigma}^2 = \frac{1}{T} \sum_{i \le T} (x_i - \hat{x})^2$
- La autocovarianza de orden $t \le T$, $\hat{\sigma}^2(t) = \frac{1}{T} \sum_{i \le T} (x_i - \hat{x})(x_{i+t} - \hat{x})$
- La autocorrelación de orden $t \le T$, $\hat{\rho}(t) = \frac{\hat{\sigma}^2(t)}{\hat{\sigma}^2}$

Remark 03.1. *Para calcular las funciones de auto-correlación parcial es necesario utilizar el algoritmo de Durbin-Watson pues las versiones empíricas de las proyecciones ortogonales no es sencillo.*

Procesos auto-regresivos

El estudio de los procesos auto-regresivos está muy cercano al de las caminatas aleatorias sin embargo en esta sección analizaremos por qué no son equivalentes.

Definition 04.1. ($AR(1)$) Decimos que un proceso estocástico $(X_t)_t$ es auto-regresivo de orden uno cuando existe algún $\phi \in \mathbb{R}$ tal que para todo $t \in \mathbb{Z}$ se tiene:

$$X_t = \phi \cdot X_{t-1} + \epsilon_t$$

Donde $(\epsilon_t)_{t\in\mathbb{Z}}$ es un ruido blanco tal que además para cualquier $t' < t$, $Cov(X_{t'}, \epsilon_t) = 0$.

A continuación resumimos las principales propiedades del proceso anterior en una proposición.

Proposition 04.1. *Sea $X = (X_t)_t$ un proceso auto-regresivo de orden uno, entonces:*

1. *Si $|\phi| < 1$ entonces X es un proceso estacionario tal que su función de auto-covarianza satisface $\rho(t) = \phi^t$.*
2. *Si $|\phi| = 1$ entonces X no es un proceso estacionario.*

Es posible definir los procesos estocásticos auto-regresivos de orden arbitrario de la siguiente manera:

Definition 04.2. $(AR(p))$ Decimos que un proceso estocástico $(X_t)_t$ es auto-regresivo de orden p cuando existen $\phi_1, \ldots, \phi_p \in \mathbb{R}$ tales que para todo $t \in \mathbb{Z}$ se tiene:

$$X_t = \phi_1 \cdot X_{t-1} + \phi_2 \cdot X_{t-2} + \ldots + \phi_p \cdot X_{t-p} + \epsilon_t$$

Donde $(\epsilon_t)_{t\in\mathbb{Z}}$ es un ruido blanco tal que además para cualquier $t' < t$, $Cov(X_{t'}, \epsilon_t) = 0$.

Promedios móviles

El problema de los procesos auto-regresivos es que no admiten una descomposición única, esto se debe a la naturaleza infinita de su definición inductiva. El proceso estocástico de promedios móviles no tiene este problema.

Definition 05.1. $(MA(q))$ Decimos que un proceso estocástico $(X_t)_t$ es un promedio móvil de orden q cuando existen $\theta_1, \ldots, \theta_q \in \mathbb{R}$ tales que para todo $t \in \mathbb{Z}$ se tiene:

$$X_t = \theta_1 \cdot_{t-1} + \theta_2 \cdot \epsilon_{t-2} + \ldots + \theta_q \cdot \epsilon_{t-q} + \epsilon_t$$

Donde $(\epsilon_t)_{t\in\mathbb{Z}}$ es un ruido blanco.

Proposition 05.1. *Sea $X = (X_t)_t$ un promedio movil orden q, entonces X es un proceso estacionario tal que su función de auto-covarianza satisface lo siguiente:*

- *Si $t > q$, $\rho(t) = 0$,*
- *Si $t \leq q$, $\rho(t) = \sigma^2 \left(\sum_{0 \leq j \leq q-t} \theta_j \theta_{j+t} \right)$,*

ARMA

A continuación vamos a introducir la clase de procesos estocásticos ARMA los cuales son una combinación entre los modelos auto-regresivos y los promedios móviles. Este tipo de procesos son muy importantes pues es posible demostrar que cualquier proceso estacionario puede aproximarse utilizando modelos ARMA.

Definition 06.1. ($ARMA(p,q)$) Decimos que un proceso estocástico $(X_t)_t$ es un proceso auto-regresivo con promedios móviles de orden (p,q) cuando existen $\theta_1, \ldots, \theta_q \in \mathbb{R}$ y $\phi_1, \ldots, \phi_p \in \mathbb{R}$ tales que para todo $t \in \mathbb{Z}$ se tiene:

$$X_t = \phi_1 \cdot X_{t-1} + \phi_2 \cdot X_{t-2} + \ldots + \phi_p \cdot X_{t-p} + \theta_1 \cdot_{t-1} + \theta_2 \cdot \epsilon_{t-2} + \ldots + \theta_q \cdot \epsilon_{t-q} + \epsilon_t$$

Donde $(\epsilon_t)_{t \in \mathbb{Z}}$ es un ruido blanco.

Proposition 06.1. *Un modelo ARMA es estacionario.*

06.1 ARIMA

Los modelos $ARIMA(p,d,q)$ son una combinación de lo que hicimos en la sección 01 y un modelo ARMA a saber, dada una serie de tiempo S, primero diferenciamos mediante d-veces mediante el operador ∇^d y después aplicamos un modelo $ARMA(p,q)$, esta es una primera aproximación a los procesos estocásticos no-estacionarios. En la siguiente sección introduciremos otra clase de modelos que buscan remediar la poca cantidad de procesos estacionarios presentes en los problemas cotidianos.

ARCH

Vale la pena mencionar que hasta este momento hemos estudiado series de tiempo lineales en el sentido que las componentes de nuestros procesos estocásticos estocásticos son lineales, los modelos tipo ARCH son un ejemplo de modelos en los que esto no se cumple, comenzamos con un ejemplo:

Definition 07.1. ($ARCH(p)$) Sea $(X_t)_t$ un proceso estocástico, diremos que este proceso es autorregresivo con heterocedasticidad condicional de orden p cuando existan $\alpha_0, \alpha_1, \ldots, \alpha_p$ tales que para todo $t \in \mathbb{Z}$ se tiene:

$$X_t = \epsilon_t \sqrt{\left(\alpha_0 + \alpha_1 \cdot X_{t-1}^2 + \ldots + \alpha_p \cdot X_{t-p}^2\right)}$$

Donde $(\epsilon_t)_{t\in\mathbb{Z}}$ es un ruido blanco .

Proposition 07.1. *Si $X = (X_t)_t$ es un modelo $ARCH(p)$, entonces*

$$Var(\epsilon_t | X_{t-1}, X_{t-2}, \ldots) = \alpha_0 + \alpha_1 \cdot X_{t-1}^2 + \ldots + \alpha_p \cdot X_{t-p}^2$$

04 Predicción del riesgo: un estudio comparado entre ARIMA y (ARIMA+ARCH)

El conjunto de datos fue obtenido de [SP 500 stock data] (https://www.kaggle.com/camnugent/sandp500) y consta de 619,040 observaciones diarias de precios durante cinco años (2013-2018) para 506 empresas.

El estudio se centra en el análisis de riesgo dada de los retornos de la empresa Citigroup Inc utilizando los modelos ARIMA, GARCH y una combinación de ambos. Para ello se realiza una transformación logarítmica, para después estacionalizar la serie de tiempo a partir de las 1264 mediciones del cierre diario.

Para la selección de los modelos se hace uso de la metodología Box-Jenkins, así como la comparación del Criterio de Información Akaike (AIC) y Log-Likelihood Se utilizan métodos visuales y test estadísticos para probar la estacionalidad de la serie, la elección de los modelos y la distribución de los residuos, antes de la predicción de riesgo.

05 Complementos de Series de tiempo y cálculo de riesgo

Esta sección incluye un compendio de temas que buscan complementar el conocimiento de los estudiantes sobre el forecasting utilizando series de tiempo así como la predicción del riesgo

Value at Risk

La medida de Value at Risk es una de las cantidades más utilizadas para modelar la volatilidad de los precios de un activo, desde los años 80's JP. Morgan la ha utilizado en su negocio. Vale la pena mencionar que a pesar de su uso extendido en las matemáticas financieras, el matemático francés N. Taleb es un severo crítico de estos métodos por considerarlos insuficientes y poco rigurosos.

Definition 01.1. Sea X una variable aleatoria y $\mathbb{P}_X$ ley de probabilidad asociada. Para cada valor $p \in (0,1)$ definimos el quantil de orden p de la variabble aleatoria X de la siguiente manera:

$$q_X(p) = inf\{x \in \mathbb{R} : p = \mathbb{P}(X \leq x)\}$$

Definition 01.2. Sean $X = (X_t)_{t\in\mathbb{Z}}$ un proceso estocástico, $t \in \mathbb{Z}$ y $\alpha \in (0,)$

un grado de confianza, definimos el valor en riesgo de X en el tiempo t con confianza α de la siguiente manera:

$$VaR_{\alpha}(X_t) = q_{-X_t}(\alpha)$$

Remark 01.1. *Si $Var(X_t) = \sigma_t^2$, en este curso utilizaremos la siguiente ecuación del valor en riesgo:*

$$VaR_{\alpha}(X_t) = \sigma_t^2 \cdot q_{-X_t}(\alpha)$$

Box-Jenkins

El método de Box-Jenkins consta de tres fases fundamentales:

- Identificación de los parámetros p, q, d: esta fase normalmente hace uso de las funciones de autocorrelación definidas al inicio de este texto.
- Aproximación de los parámetros: una vez definidos los parámetros p, q, d es necesario aproximar los coeficientes de los modelos $MA(q), AR(p)$ con el objetivo de encontrar nuestra serie de tiempo adecuada, esta parte normalmente se hace con un algoritmo de máxima verosimilitud.

- Test de adecuación: el test más conocido para hacer este proceso es el de Ljung-Box test.

Test de ruido blanco

Proposition 03.1. *Sean* $\{x_1, x_2, x_3, \ldots, x_T,\}$ *observaciones centradas generadas por un ruido blanco* $(X_t)_{t\in\mathbb{Z}}$. *Entonces*

$$\sqrt{T}\hat{\rho}(t) \to \mathbb{N}(0,1)$$

Definition 03.1. Sean $\{x_1, x_2, x_3, \ldots, x_T,\}$ observaciones centradas de un proceso estocástico, si $\alpha = .95$ es un nivel de confianza y definimos nuestra hipótesis cero:

H_0 : Nuestras observaciones son generadas por un ruido blanco débil.

Entonces si $\hat{\rho}(1) \in \left[\frac{-1.96}{T}, \frac{1.96}{T}\right]$ diremos que nuestra H_0 es cierta, de lo contrario la rechazaremos.

06 Referencias

[1] «Conda — conda 4.8.3.post5+125413ca documentation». https://docs.conda.io/projects/conda/en/latest/ (accedido mar. 26, 2020).

[2] Anaconda is Bloated - Set up a Lean, Robust Data Science Environment with Miniconda and Conda-Forge.

[3] «Miniconda — Conda documentation». https://docs.conda.io/en/latest/miniconda.html (accedido mar. 26, 2020).

[4] «A brief introduction — conda-forge 2019.01 documentation». https://conda-forge.org/docs/user/introduction.html (accedido mar. 26, 2020).

BOURBAKI

COLEGIO DE MATEMÁTICAS

escuela-bourbaki.com