Universitat de Barcelona

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

SÈRIES TEMPORALS I ALGUNS EXEMPLES

Autor: Ramón Rosselló Cifre

Director: Dr. Carles Rovira Escofet
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 20 de juny de 2021

Abstract

Any phenomenon that evolves in time, such as the average daily temperature from a region, or monthly sales of a company, from a probabilistic point of view, can be attempted to model using a time series. During this project, we focus on those series that can be decomposed classically, explaining the main characteristics of the most important classic models associated with them. To illustrate the theoretical results, we start from three different sets of observations: the average daily temperature of Madrid, the monthly sales of a company and the number of monthly admissions to ICU (Intensive Care Unit) of a hospital. This way, in parallel with the theoretical concepts explained, we are trying to adjust a probabilistic model suitable for each of the examples, in order to explain accurately how each phenomena behaves. Once justified, that the models found are correct using certain diagnostics, our proposal is to predict how each set of data will evolve in the future.

Resum

Qualsevol fenomen que evoluciona al llarg del temps, com la temperatura mitjana diària d'una determinada regió, o les vendes mensuals d'una empresa, des del punt de vista probabilístic, es pot intentar modelitzar mitjançant una determinada sèrie temporal. Al llarg d'aquest treball, ens centrem amb aquelles sèries que es poden descomposar de forma clàssica, explicant les principals característiques dels models clàssics més importants associats a elles. Amb l'objectiu d'il·lustrar els resultats teòrics, partim de tres conjunt d'observacions diferents: la temperatura mitjana diària de Madrid, les vendes mensuals d'una determinada empresa i el nombre de trasllats mensuals a la UCI (Unitat de Cures Intensives) d'un hospital. D'aquesta forma, paral·lelament als conceptes teòrics explicats, intentem ajustar un model probabilístic adient a cadascun dels exemples, amb la finalitat que aquests expliquin de forma adequada com es comporta cada fenomen. Finalment, un cop justificats, mitjançant uns determinats diagnòstics, que els models trobats són correctes, el nostre propòsit és realitzar prediccions per a saber com evolucionarà cada conjunt de dades en el futur.

2020 Mathematics Subject Classification. 62M10, 62M20

Agraïments

Vull agrair principalment al meu tutor, Carles Rovira Escofet, per totes les indicacions que m'ha fet arribar des del primer moment, el guiament a l'hora d'escollir el tema i la seva disponibilitat durant aquests mesos de treball.

Índex

1 Introducció

El projecte

En aquest projecte veiem com ajustar un model probabilístic (una determinada sèrie temporal) a un conjunt d'observacions que ha estat recollit durant un cert període temporal, amb els objectius d'explicar com es comporta el fenomen i de realitzar prediccions de futures observacions. En concret, es treballa amb 3 conjunts de dades diferents, els quals s'han mencionat en el resum inicial i detallarem després. El treball està dividit en 5 seccions.

1. En la primera, la idea és estacionar les nostres dades. És a dir, realitzar una sèrie de transformacions a les observacions, de forma que les puguem modelitzar mitjançant una sèrie temporal estacionària.
2. En la següent secció ens centrem en l'estudi de les sèries estacionàries, concretament, dels processos lineals, explicant les principals característiques i com es realitzen les prediccions.
3. En la tercera part, estudiem els processos ARMA (un dels processos linals més importants), i veiem com ajustar un model ARMA a les nostres dades "estacionades", de forma que siguem capaços de realitzar prediccions. Posteriorment, mitjançant uns diagnòstics, es valida el model trobat.
4. En la secció 4 s'explica com és el model que ajusta les dades originals, i a traslladar les prediccions de les dades "estacionades"a les dades originals.
5. Finalment, es detallen les conclusions finals de tot el projecte.

Observació 1.1. De totes les referències mencionades al final del treball, les dues primeres (la [1] i la [2]) són les més importants, és a dir, aquelles de les quals, principalment, s'ha extret gairebé tota la informació necessària per dur a terme aquest projecte.

1.1 Nocions bàsiques

En primer lloc, detallem un conjunt de nocions bàsiques per a saber tractar les sèries temporals.

Definició 1.2. *Sigui* $(\Omega, \mathbb{F}, \mathbb{P})$ *un espai de probabilitat. Definim una sèrie temporal com un conjunt ordenat* $\{X_t\}_{t\in T\subseteq \mathbb{R}}$ *de variables aleatòries* $X_t : \Omega \longrightarrow \mathbb{R}$ *indexades en un conjunt* T*.*

Notació 1. *Al llarg del treball, en fer referència a una sèrie temporal* $\{X_t\}_{t\in\mathbb{Z}}$*, obviem que aquesta està definida sobre un espai de probabilitat* $(\Omega, \mathbb{F}, \mathbb{P})$*.*

Observacions:

1. T s'anomena l'espai de paràmetres.
   (a) Si T és un conjunt numerable, com $\mathbb{N}$ o $\mathbb{Z}$, la sèrie temporal és discreta.
   (b) Si T és un conjunt no numerable, com $\mathbb{R}$, la sèrie temporal és contínua.

2. El subíndex t representa l'instant del temps de la sèrie i X_t l'estat d'aquesta al temps t.

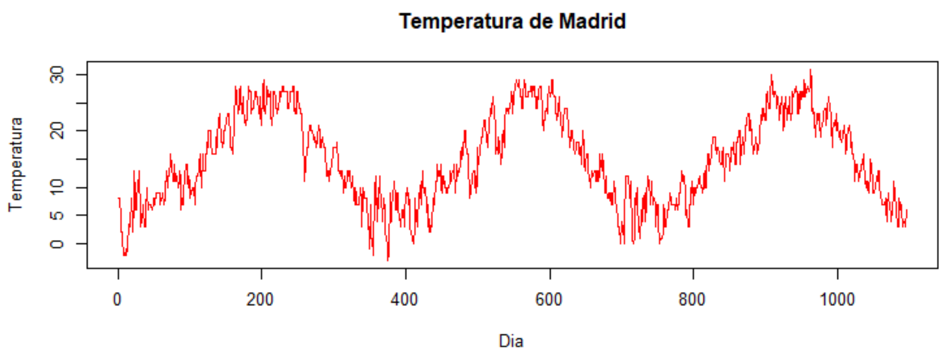
Definició 1.3. *Un model complet d'una sèrie temporal* $\{X_t\}_{t \in T}$ *és l'especificació de totes les distribucions conjuntes en dimensió finita dels vectors aleatoris* $(X_{t_1}, ..., X_{t_m})$*, per tot* $\{t_1, \ldots, t_m\} \subseteq T$ *i per tot* $m \geq 1$ *.*

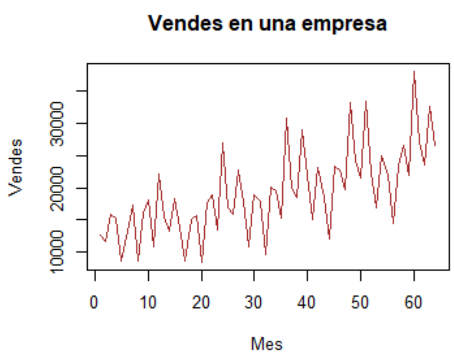
Nosaltres partim d'un nombre finit de dades ordenades cronològicament $\{x_1, ..., x_n\}$ sobre un determinat fenomen. El nostre objectiu és determinar la sèrie temporal $\{X_t\}_{t \in \mathbb{Z}}$ (suposarem que l'espai de paràmetres és $\mathbb{Z}$) de la que provenen les dades. És a dir, que sigui justificable que

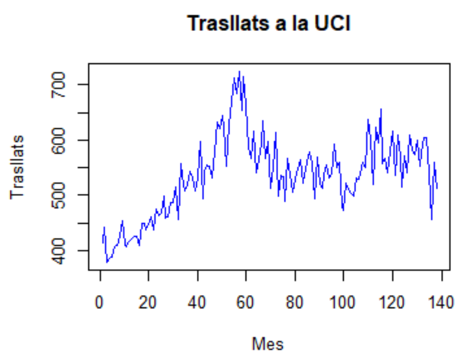
$$\{x_1, ..., x_n\} = \{X_1(w_0), ..., X_n(w_0)\} \text{ per a un cert } w_0 \in \Omega.$$

Un cop l'haguem trobada (havent-la validada amb una sèrie de diagnòstics), no en coneixarem el seu model complet, només tindrem accés a alguna informació com les seves propietats de segon ordre, per exemple $E(X_t), Var(X_t)$ o $Cov(X_{t_1}, X_{t_2})$, on $\{t, t_1, t_2\} \subseteq \mathbb{Z}$. En aquesta situació, el nostre objectiu serà intentar predir com es comportarà el fenomen en el futur, és a dir, estimar numèricament $X_{n+h}(w_0)$, per a $h \in \mathbb{N}$.

Exemple 1.4. Per a representar les n dades, el més còmode és graficar-les en un pla cartesià on l'eix d'abscisses representa l'instant del temps i l'eix d'ordenades el valor de l'observació. Al llag del treball, intentarem determinar tres sèries temporals que ajustin els tres conjunts d'observacions següents:







- El primer gràfic, el de color vermell, recull la temperatura mitjana diària, en graus Celsius, a l'aeroport de Barajas de Madrid des de l'1 de gener de 2009 fins al 31 de desembre de 2011 (1095 observacions). Denotem per $\{X_t^1\}_{t \in \mathbb{Z}}$, on $X_t^1 : \Omega_1 \longrightarrow \mathbb{R}$, són variables aleatòries definides sobre un espai de probabilitat $(\Omega_1, \mathbb{F}_1, \mathbb{P}_1)$, la sèrie temporal que modelitza la temperatura diària a l'aeroport de Barajas de Madrid. En particular, de la qual prové aquesta mostra, és a dir, que $\exists w_0^1 \in \Omega_1$, tal que les observacions de la temperatura corresponen a

$$\{X_1^1(w_0^1), ..., X_{1095}^1(w_0^1)\}.$$

- El segon gràfic, el de color marró, indica les vendes mensuals, en euros, d'una determinada empresa des del gener de 2015 fins a l'abril de 2020 (64 dades). Denotem per $\{X_t^2\}_{t\in\mathbb{Z}}$, on $X_t^2 : \Omega_2 \longrightarrow \mathbb{R}$, són variables aleatòries definides sobre un espai de probabilitat $(\Omega_2, \mathbb{F}_2, \mathbb{P}_2)$, la sèrie temporal que modelitza les vendes mensuals de l'empresa. En concret, de la qual prové aquestes dades, és a dir, que $\exists w_0^2 \in \Omega_2$, tal que la mostra de les vendes correspon a

$$\{X_1^2(w_0^2), ..., X_{64}^2(w_0^2)\}.$$

- L'últim gràfic, el de color blau, mostra el nombre mensual de trasllats de pacients a la UCI, a l'Hospital General d'Horton (Kansas), des del novembre de 1999 fins a l'abril de 2011 (138 observacions). Denotem per $\{X_t^3\}_{t\in\mathbb{Z}}$ la sèrie temporal, on $X_t^3 : \Omega_3 \longrightarrow \mathbb{R}$ són variables aleatòries definides sobre un espai de probabilitat $(\Omega_3, \mathbb{F}_3, \mathbb{P}_3)$, que modelitza el nombre mensual de trasllats de pacients a la UCI. En particular, de la qual prové aquesta mostra, és a dir, que $\exists w_0^3 \in \Omega_3$, tal que les dades dels trasllats corresponen a

$$\{X_1^3(w_0^3), ..., X_{138}^3(w_0^3)\}.$$

El nostre objectiu és estudiar les sèries temporals $\{X_t^k\}_{t\in\mathbb{Z}}$, per a $k \in \{1, 2, 3\}$.

A continuació introduirem una sèrie de nocions per ajudar a tractar les propietats de segon ordre d'una sèrie temporal qualsevol amb espai de paràmetres $\mathbb{Z}$.

Definició 1.5. *Sigui $\{X_t\}_{t\in\mathbb{Z}}$ una sèrie temporal tal que $E(X_t^2) < \infty$, $\forall t \in \mathbb{Z}$. Definim la funció mitjana μ_X:*

$$\begin{aligned} \mu_X : \mathbb{Z} &\longrightarrow \mathbb{R} \\ t &\longmapsto E(X_t), \end{aligned}$$

i la funció d'autocovariància (ACVF) γ_X:

$$\begin{aligned} \gamma_X : \mathbb{Z} \times \mathbb{Z} &\longrightarrow \mathbb{R} \\ (t, s) &\longmapsto Cov(X_t, X_s). \end{aligned}$$

Ara bé, com podem trobar la sèrie temporal $\{X_t\}_{t\in\mathbb{Z}}$ que s'ajusti a la nostra mostra? Clarament és una pregunta molt general i cal fer algunes consideracions sobre com és aquesta sèrie. El que suposarem, és que la sèrie es pot descomposar de forma clàssica. A continuació veurem en què consisteix tal descomposició.

1.2 Descomposició clàssica

Per entendre com és la descomposició clàssica d'una sèrie temporal, cal introduir què entenem per una sèrie temporal estacionària.

Definició 1.6. *Sigui* $\{X_t\}_{t\in\mathbb{Z}}$ *una sèrie temporal tal que* $E(X_t^2) < \infty$, $\forall t \in \mathbb{Z}$. *Diem que* $\{X_t\}_{t\in\mathbb{Z}}$ *és una sèrie temporal estacionària si:*

- $\mu_X(t) = \mu$. *És a dir,* $\mu_X(t)$ *és independent de* t, $\forall t \in \mathbb{Z}$.
- $\gamma_X(t+h,t)$ *és independent de* t *per a tot* h. *És a dir,*

$$\gamma_X(t+h,t) = \gamma_X(h,0), \;\; \forall t,h \in \mathbb{Z}.$$

En tal cas definim la funció d'autocovariància (ACVF) com a :

$$\gamma_X(h) = \gamma_X(h,0), \;\; \forall h \in \mathbb{Z}.$$

Definició 1.7. *Sigui* $\{X_t\}_{t\in\mathbb{Z}}$ *una sèrie temporal estacionària. Definim la funció d'autocorrelació (ACF)* p_X*:*

$$\begin{aligned} p_X : \mathbb{Z} &\longrightarrow \mathbb{R} \\ h &\longmapsto \frac{\gamma_X(h)}{\gamma_X(0)}. \end{aligned}$$

Un cop explicat el concepte d'estacionarietat podem definir com ve donada la descomposició clàssica d'una sèrie temporal.

Sigui $\{X_t\}_{t\in\mathbb{Z}}$ una sèrie temporal. Diem que aquesta admet una descomposició clàssica si:

$$X_t = m_t + s_t + Y_t, \;\; \forall t \in \mathbb{Z},$$

on :

1. **Component tendència $\mathbf{m_t}$:**

   És determinista, és a dir, $m_t \in \mathbb{R}$, $\forall t \in \mathbb{Z}$. Representa el comportament a llarg termini de la sèrie temporal. En la gran majoria de casos és una recta, o una funció polinomial que evolciona al llarg del temps.

2. **Component d'estacionalitat $\mathbf{s_t}$:**

   És determinista, és a dir, $s_t \in \mathbb{R}$, $\forall t \in \mathbb{Z}$. A més, $\exists d \in \mathbb{Z}$, no negatiu, tal que $s_t = s_{t+d}$, $\forall t \in \mathbb{Z}$. Representa el comportament periòdic de la sèrie. Moltes sèries temporals presenen oscil·lacions regulars cada cert període d que han de ser recollides en aquesta component. Sol estar relacionada amb la freqüència temporal amb la qual està definida la sèrie. Per exemple, si cada instant "t" representa un mes diferent, llavors la sèrie sol presentar una component d'estacionalitat de període 12 (anual). De la mateixa forma, si l'instant "t" representa un dia, llavors la sèrie sol presentar una component d'estacionalitat de períodes 7, 30 o 365, és a dir, un període setmanal, mensual o anual respectivament.

3. **Component irregular o soroll $\mathbf{Y_t}$:**

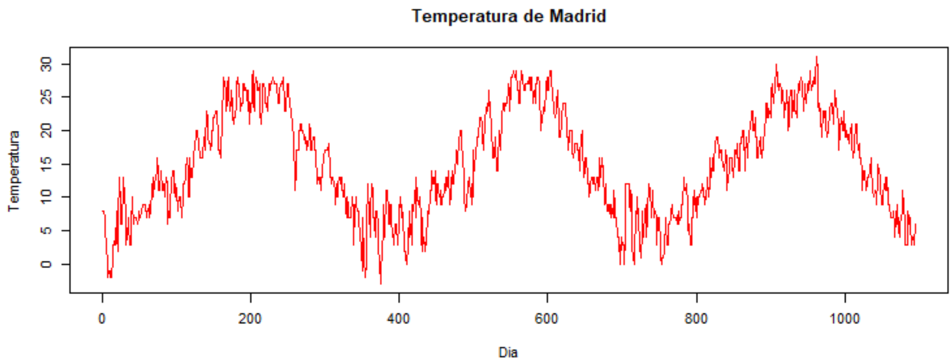
   És la component aleatòria de la sèrie. $\{Y_t\}_{t\in\mathbb{Z}}$ és una sèrie temporal estacionària. Recull alteracions sense cap pauta periòdica ni tendencial. Conforma la incapacitat del model per explicar a la perfecció el comportament de la sèrie temporal.

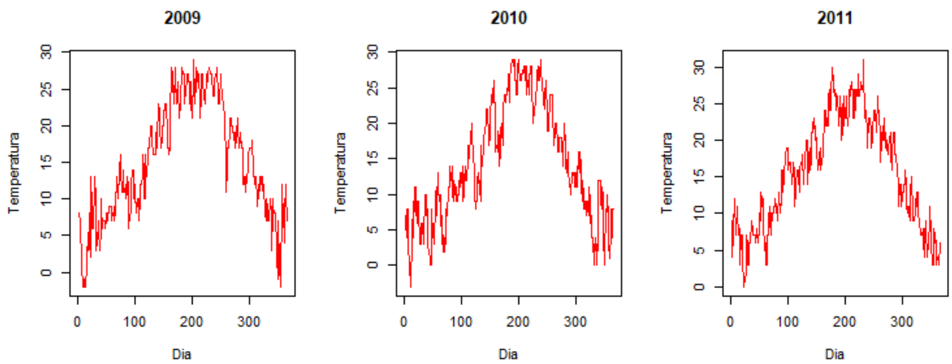
Observació 1.8. $\{Y_t\}_{t\in\mathbb{Z}}$ és una sèrie temporal estacionària. En particular, té moments de segon ordre finits, el que implica que la sèrie temporal $\{X_t\}_{t\in\mathbb{Z}}$ també. Concretament, la variància és constant:

$$\forall t \in \mathbb{Z},\ \gamma_X(t,t) = Var(X_t) = Var(Y_t) = \gamma_Y(0).$$

Exemple 1.9. Seguint els 3 exemples anteriors, suposarem que les sèries temporals $\{X_t^k\}_{t\in\mathbb{Z}}$, per a $k \in \{1,2,3\}$, admeten descomposició clàssica. Tractem cadascuna per separat per a distingir les components de cada sèrie. Per a la tendència, ens basarem amb el comportament a llarg termini de les observacions, dibuixant-hi, si hi cal, una corba que ens mostri, visualment, com evoluciona la mostra al llarg del temps. Aquesta corba, és només una estimació que realitzem "a ull", per a tenir una idea si la sèrie presenta una certa tendència. Per a la periodicitat, tractarem cada exemple per separat, amb l'objectiu de trobar una sèrie de patrons en les dades que es repeteixin anualment.

- Respecte la temperatura de Madrid:



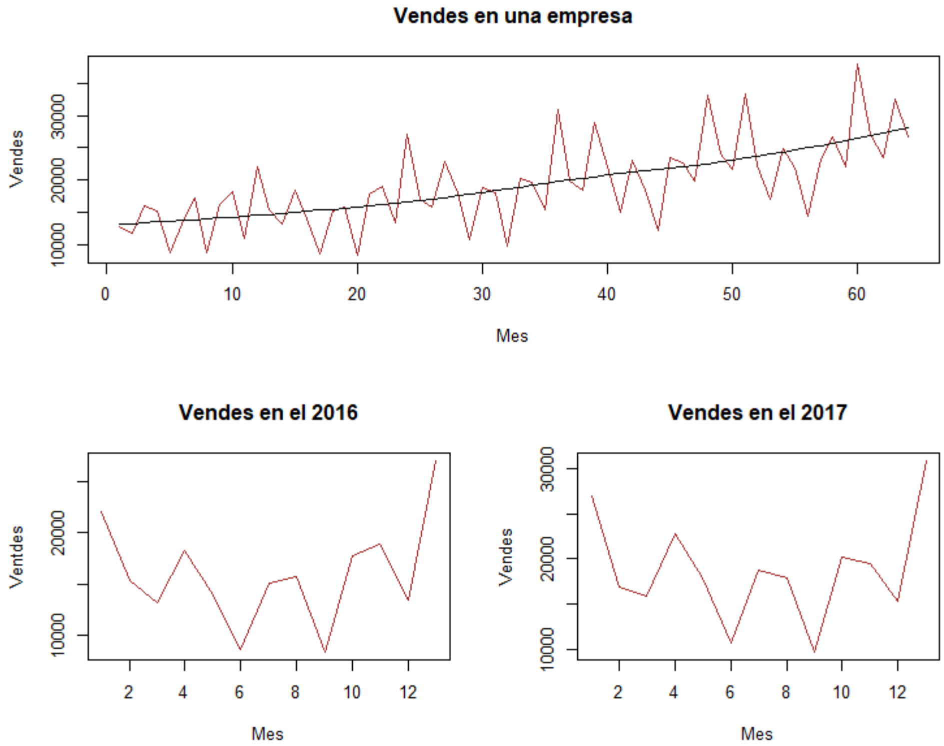


  Fixem-nos que cada 365 dies, com veiem en els gràfics de la temperatura en cada any diferent, es repeteix un patró molt semblant, degut a que cada any té les mateixes estacions. Efectivament, durant l'inici de cadascun dels tres anys la temperatura es manté més constant, llavors creix fins al voltant del dia 200 (primavera i estiu), es torna estabilitzar un temps (estiu) i finalment decreix fins al darrer dia de l'any (tardor i hivern). Llevat d'aquesta clara component periòdica, també sembla que, en general i a llarg termini, la temperatura va creixent molt lentament en cada any, de forma que en el tercer any la temperatura és una mica més alta que en els dos anteriors. Per tant, podem suposar que la sèrie $\{X_t^1\}_{t\in\mathbb{Z}}$, es descomposa de forma següent, per a tot enter t:

$$X_t^1 = m_t^1 + s_t^1 + Y_t^1,$$

on m_t^1 recull aquest petit creixament a llarg termini en funció del temps, s_t^1 és la component d'estacionalitat de període 365 (anual), i Y_t^1, és el soroll i la component aleatòria de la sèrie.

- Respecte les vendes de l'empresa:

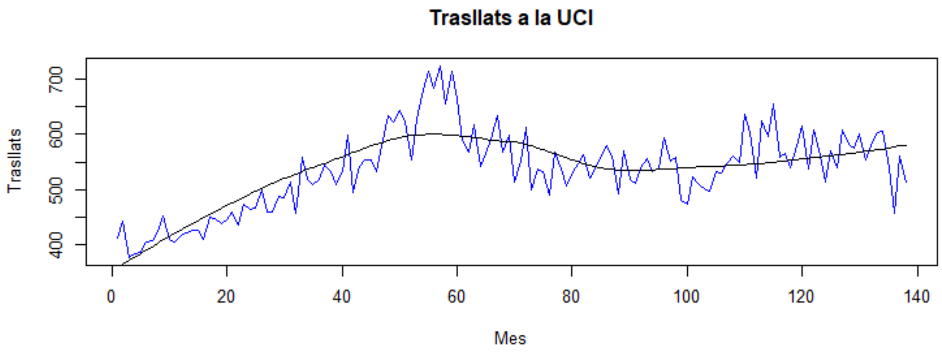


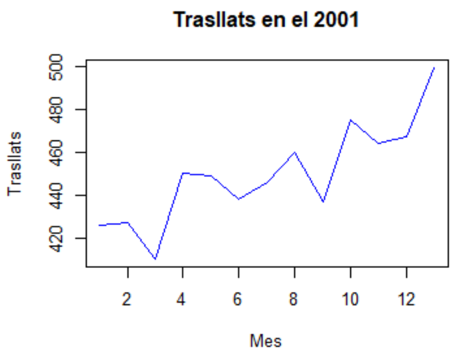
En primer lloc, al primer gràfic, a llarg termini, observem que les vendes van creixent a mesura que el temps avança (aquest fet és representat mitjançant la corba negra que hem dibuixat). De fet, fixem-nos que, en general, el creixament és lineal (és una recta). En segon lloc, observem que cada any, (cada 12 mesos), es repeteix un patró molt semblant. Per a visualtzar-ho fixem-nos amb els dos gràfics d'abaix, on el primer mostra les vendes de l'empresa des del desembre de 2015 fins el desembre del 2016, i el segon mostra les vendes de l'empresa des del desembre de 2016 fins al desembre de 2017: al llarg dels anys 2016 i 2017, el nombre de vendes, durant el gener i el febrer cau, pel març creix bastant, durant l'abril i maig baixa considerablament, per juny i juliol puja, llavors torna a baixar a l'agost, al septembre i a l'octubre torna a pujar, al novembre cau, mentre que finalment al desembre creix. Com aquest patró es repeteix durant cada any, podem suposar que la sèrie $\{X_t^2\}_{t\in\mathbb{Z}}$, es descomposa de forma següent, per a tot enter t:

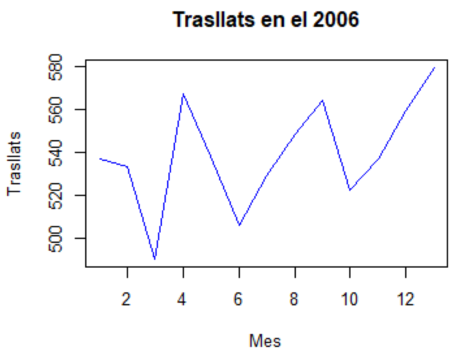
$$X_t^2 = m_t^2 + s_t^2 + Y_t^2,$$

on m_t^2 recull aquest comportament de recta a llarg termini en funció del temps, s_t^2 és la component d'estacionalitat de període 12 (anual), i Y_t^2, és el soroll i la component aleatòria de la sèrie.

- Respecte els trasllats a la UCI:







Al gràfic d'amunt, a llarg termini, observem que els trasllats segueixen la corba negra dibuixada que hem traçat. És a dir, van creixent fins el mes 60, llavors disminueixen fins al 90, i tornen a créixer fin al 140 . Per altra banda, observem que les dades es recullen mensualment, i a més cada any, (cada 12 mesos), també es poden observar algus patrons que es repeteixen, tot i que aquests no són tan evidents com en l'exemple de les vendes. Per a visualtzar-ho, fixem-nos amb els dos gràfics d'abaix. El primer mostra els trasllats mensuals a la UCI del desembre del 2000 fins al desembre del 2001, mentre que el segon mostra els trasllats del desembre del 2005 al desembre del 2006: tant al 2001 com al 2006, a l'inici de l'any (gener i febrer) els trasllats disminueixen, i al març, amb l'arribada de la primavera, el nombre de trasllats a la UCI creix significativament. Durant els següents mesos, els trasllats varien al voltant d'aquest valor, fins al septembre, on amb el final de l'estiu i l'inici de la tardor, el nombre de trasllats comença a créixer fins al desembre. Aquestes semblançes anuals se solen reproduir en tots els anys on les dades han estat observades. Per tant, podem concloure que la sèrie $\{X_t^3\}_{t\in\mathbb{Z}}$, es descomposa de forma següent, per a tot enter t:

$$X_t^3 = m_t^3 + s_t^3 + Y_t^3,$$

on m_t^3 recull aquest comportament de la corba negra a llarg termini en funció del temps, s_t^3 és la component d'estacionalitat de període 12 (anual), i Y_t^3, és el soroll i la component aleatòria de la sèrie.

En conclusió, a partir d'ara suposarem que la sèrie temporal a determinar $\{X_t\}_{t\in\mathbb{Z}}$ que ajusti les n observacions, és a dir, que per a un cert $w_0 \in \Omega$,

$$\{X_1(w_0), ..., X_n(w_0)\} = \{x_1, ..., x_n\},$$

admet una descomposició clàssica. Tot i així, és massa complex estudiar en general totes aquestes sèries, però resulta que aquest fet no és problemàtic, ja que com veurem en la següent secció, totes elles es poden transformar en estacionàries, i aquestes últimes sí que les sabrem gestionar.

1.3 L'operador B: De la descomposició clàssica a una sèrie estacionària

Per estacionar la sèrie amb descomposició clàssica $\{X_t\}_{t\in\mathbb{Z}}$, cal introduir el concepte de l'operador de retrocés B.

De resultats bàsics d'anàlisis real i funcional, sabem que $\left(L^2(\Omega)/\sim, <.\,,\,.>\right)$, on $L^2(\Omega) = \{X : \Omega \longrightarrow \mathbb{R}\ \mathit{mesurables} : E(X^2) < \infty\}$, és un espai de Hilbert (real). És a dir, és un espai vectorial complet respecte la norma induïda pel producte escalar $<.\,,\,.>$. $L^2(\Omega)/\sim$ és el conjunt quocient derivat de la relació d'equivalència $\sim$ definida a $L^2(\Omega)$ com a:

$$X \sim Y \longleftrightarrow E\left((X-Y)^2\right) = 0,$$

i $<.\,,\,.>$ és el producte escalar definit com a:

$$\begin{aligned} L^2(\Omega)/\sim \times L^2(\Omega)/\sim &\longrightarrow \mathbb{R} \\ (X,Y) &\longmapsto E(XY). \end{aligned}$$

Per simplificar la notació, s'identifica $L^2(\Omega)$ amb $L^2(\Omega)/\sim$.

Fixem-nos també que per a tot subespai vectorial $F \subseteq L^2(\Omega)$, es pot definir

$$\mathbb{L}(F) = \{g : F \longrightarrow F : \text{són lineals}\},$$

que amb la suma habitual d'operadors és un espai vectorial real.

Sigui $\{X_t\}_{t\in\mathbb{Z}}$ una sèrie temporal qualsevol que admet una descomposició clàssica. Definim

$$M = < \{X_t\}_{t\in\mathbb{Z}} >,$$

és a dir, el subespai de $L^2(\Omega)$ generat pel conjunt de variables aleatòries que formen la sèrie temporal $\{X_t\}_{t\in\mathbb{Z}}$.

Considerem l'espai vectorial real $\mathbb{L}(M)$, i definim el següent operador lineal pertanyent a $\mathbb{L}(M)$ (definit en els elements de la base de M):

$$\begin{aligned} B : M &\longrightarrow M \\ X_t &\longmapsto X_{t-1}. \end{aligned}$$

Definim també les potències j-èssimes de l'operador B, per a tot $j \in \mathbb{Z}$, com a:

$$B^j(X_t) = X_{t-j}.$$

$B^0 = I$, és a dir, l'operador identitat de $\mathbb{L}(M)$, i, clarament, totes les potències de l'operador pertanyen a $\mathbb{L}(M)$, i a més, fixem-nos que podem manipular sèries de potències finites en B multiplicant-les i sumant-les com si es tractessin de sèries de potències finites amb variables reals, identificant la suma i la composició d'operadors amb la suma i el producte de sèries respectivament. Per exemple:

$$\begin{aligned}\left((I-2B+B^2)\circ(2I+B^{-2})\right)(X_t) &= (2I+B^{-2}-4B-2B^{-1}+2B^2+I)(X_t)\\ &= (3I+B^{-2}-2B^{-1}-4B+2B^2)(X_t)\\ &= 3X_t+X_{t+2}-2X_{t+1}-4X_{t-1}+2X_{t-2}.\end{aligned}$$

Un cop explicats tots aquests conceptes, ja podem veure com tansformar la sèrie $\{X_t\}_{t\in\mathbb{Z}}$ en estacionària.

1.4 Eliminació per diferenciació de l'estacionalitat i de la tendència

Sigui $\{X_t\}_{t\in\mathbb{Z}}$ una sèrie temporal que es pot descomposar de forma clàssica, i el conjunt $\{X_1(w_0),...,X_n(w_0)\}$ les nostres observacions. Suposem que en coneixem el període de la component estacionalitat. Tenim que,

$$X_t = m_t + s_t + Y_t,\ \forall t\in\mathbb{Z}.$$

Per a transformar la sèrie en estacionària, dividim el procés en dues etapes:

1. **Eliminació de l'estacionalitat:**

   En primer lloc eliminarem l'estacionalitat mitjançant l'operador $I-B^d$, on d és el període de la component estacionalitat. En la sèrie temporal resultant $\{Z_t\}_{t\in\mathbb{Z}}$, les variables aleatòries venen definides per

   $$Z_t = (I-B^d)(X_t) = X_t - X_{t-d},\ \forall t\in\mathbb{Z},$$

   és a dir, per a tot enter t,

   $$Z_t = m_t+s_t+Y_t-(m_{t-d}+s_{t-d}+Y_{t-d}) = m_t-m_{t-d}+Y_t-Y_{t-d}.$$

   Veiem que $\{Z_t\}_{t\in\mathbb{Z}}$ es pot descompondre de forma clàssica: té com a tendència m_t-m_{t-d}, com a component irregular Y_t-Y_{t-d} (ja que $\{Y_t-Y_{t-d}\}_{t\in\mathbb{Z}}$ és una sèrie temporal estacionària) i no presenta estacionalitat. Fixem-nos que en aquest cas les n observacions $\{X_1(w_0),...,X_n(w_0)\}$, provenents del procés $\{X_t\}_{t\in\mathbb{Z}}$ queden reduïdes a $\{Z_{d+1}(w_0),...,Z_n(w_0)\}$, on

   $$Z_i(w_0) = X_i(w_0)-X_{i-d}(w_0),\ \forall i\in\{d+1,...,n\},$$

   i aquestes serien les observacions del procés $\{Z_t\}_{t\in\mathbb{Z}}$ sense estacionalitat.

   Exemple 1.10. De l'exemple 1.9, tenim que les sèries

   $$\{X_t^1\}_{t\in\mathbb{Z}}, \{X_t^2\}_{t\in\mathbb{Z}} \text{ i } \{X_t^3\}_{t\in\mathbb{Z}},$$

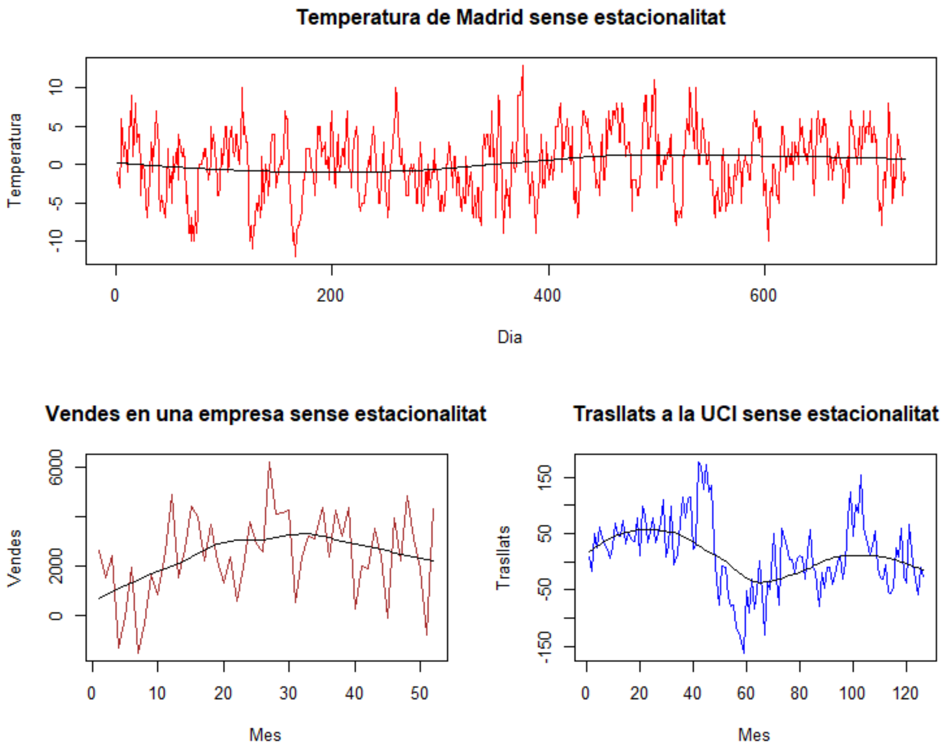
   es descomposen de forma clàssica amb components d'estacionalitat de períodes 365, 12 i 12 respectivament. Si, tenint en compte la secció 1.3, considerem els operadors de retrocés B_k i d'identitat I_k associats a les sèries $\{X_t^k\}_{t\in\mathbb{Z}}$, per a $k\in\{1,2,3\}$, podem eliminar aquesta component d'estacionalitat considerant les noves sèries temporals

   $$\{(I_1-B_1^{365})X_t^1\}_{t\in\mathbb{Z}}, \{(I_2-B_2^{12})X_t^2\}_{t\in\mathbb{Z}} \text{ i } \{(I_3-B_3^{12})X_t^3\}_{t\in\mathbb{Z}},$$

les quals es poden descomposar de forma clàssica i no presenten component d'estacionalitat. Fixem-nos que un cop eliminada aquesta component, amb la mateixa notació que a l'exemple 1.3, disposem de les mostres següents:

Temperatura:	$\{(I_1 - B_1^{365})X_{366}^1(w_0^1), ..., (I_1 - B_1^{365})X_{1095}^1(w_0^1)\}$.
Vendes:	$\{(I_2 - B_2^{12})X_{13}^2(w_0^2), ..., (I_2 - B_2^{12})X_{64}^2(w_0^2)\}$.
Trasllats:	$\{(I_3 - B_3^{12})X_{13}^3(w_0^3), ..., (I_3 - B_3^{12})X_{138}^3(w_0^3)\}$.

Aquestes són representades en els següents gràfics:



La corba negra de la temperatura, representant la tendència d'aquesta, sí sembla força invariant, però tot i així presenta lleus oscil·lacons, el que provoca que no la considerem constant del tot. A més, les altres 2 corbes negres traçades en els dos altres gràfics, representant de forma visual la tendència de cadascuna d'elles, no són constants, indicant-nos que les noves mostres sí presenten tendència. És a dir, arribem a la conclusió que les tres noves sèries temporals es poden descomposar de forma clàssica, sí tenen tendència i no presenten component d'estacionalitat.

2. **Eliminació de la tendència:**

   Un cop aplicat el primer pas, la sèrie temporal resultant no presenta estacionalitat. Vegem com eliminar la tendència.

   Sigui $\{X_t\}_{t\in\mathbb{Z}}$ una sèrie temporal que es pot descomposar de forma clàssica i que no presenta estacionalitat. És a dir,

$$X_t = m_t + Y_t, \ \forall t \in \mathbb{Z}.$$

   La idea és trobar un $j \in \mathbb{N}$ tal que la sèrie temporal $\{Z_t\}_{t\in\mathbb{Z}}$ sigui estacionària, on les respectives variables aleatòries del procés es defineixen com a

$$Z_t = ((I - B)^j)X_t, \ \forall t \in \mathbb{Z}.$$

Si $\{X_t\}_{t\in\mathbb{Z}}$ és la sèrie temporal que ajusta la nostra mostra de n observacions $\{X_1(w_0),...,X_n(w_0)\}$, llavors aquesta queda reduïda a

$$\{Z_{j+1}(w_0), Z_{j+2}(w_0), ..., Z_n(w_0)\},$$

on

$$Z_i(w_0) = (I-B)^j X_i(w_0),\ \forall i \in \{j+1,...,n\},$$

representen les observacions del procés estacionari $\{Z_t\}_{t\in\mathbb{Z}}$. Això és degut a que si $i \leq j$, llavors per a calcular $(I-B)^j X_i(w_0)$ necessitariem tenir alguna mostra de $X_k(w_0)$ per a $k \leq 0$, cosa que no posseïm.

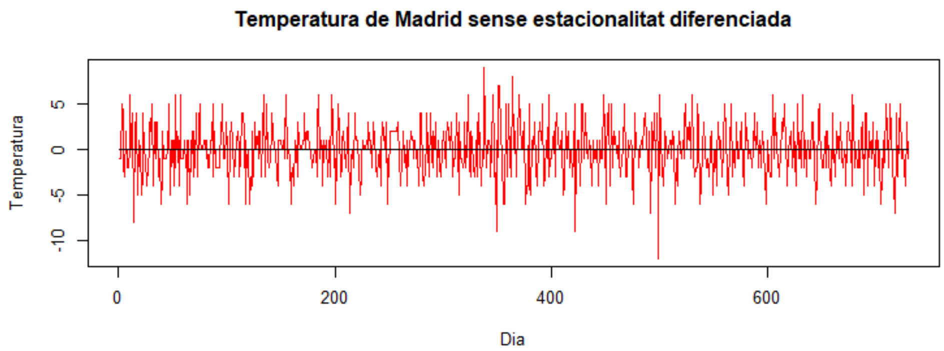
Per a determinar el j ens hem de basar amb la nostra mostra. La idea és la següent:

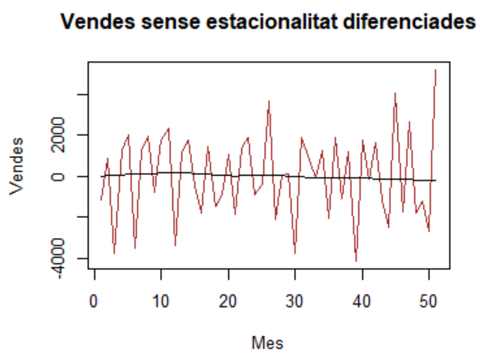
- Començant per a $j = 1$, graficar les $n-1$ observacions $\{(I-B)X_2(w_0),...,(I-B)X_n(w_0)\} = \{X_2(w_0)-X_1(w_0),...,X_n(w_0)-X_{n-1}(w_0)\}$.
- Comprovar que la nova mostra provengui d'un procés estacionari, veient que no presenta una tendència que evolucioni al llarg del temps, i que la variància és gairabé constant al llarg de la mostra. Si es compleixen aquestes condicions, llavors considerem $j = 1$, i per tant la sèrie temporal $\{(I-B)X_t\}_{t\in\mathbb{Z}}$, és estacionària. En cas de que no es compleixi aquest fet, cal considerar $j = j+1$, i tornar al pas anterior, és a dir, tornar a graficar les noves observacions per a comprovar que provenguin d'una sèrie temporal estacionària. A la pràctica, el j triat és molt petit, sent 1 en la majoria de casos.

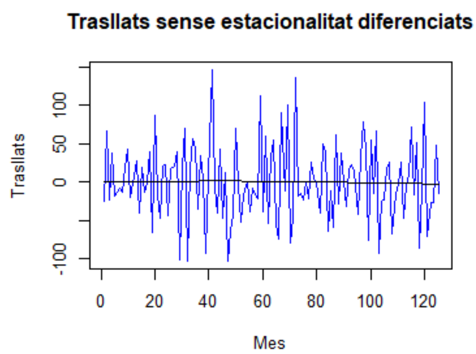
Exemple 1.11. De l'exemple anterior, tenim que les sèries

$$\{(I_1 - B_1^{365})X_t^1\}_{t\in\mathbb{Z}}, \{(I_2 - B_1^{12})X_t^2\}_{t\in\mathbb{Z}} \text{ i } \{(I_3 - B_3^{12})X_t^3\}_{t\in\mathbb{Z}},$$

es descomposen de forma clàssica i no presenten estacionalitat (però sí tendència). Si diferenciem, per a $j = 1$, cadascun del conjunt d'observacions dels 3 processos, obtenim els següents gràfics:







Observem que les 3 corbes negres traçades en els gràfics, representant de forma visual la tendència de cadascuna d'elles, són gairebé constants, indicant-nos que les noves mostres no presenten tendència. Per altra banda, en cap dels 3 casos trobem un patró on la variància de les dades creixi o decreixi al llarg del temps. En els 3 gràfics, trobem alguns casos puntuals extrems, com la dada 501 en la temperatura de Madrid, amb un valor de -12, o l'última dada (51) de les vendes, amb un valor de 5160, o les observacions 42 i 72 dels trasllats, amb valors de 146 i 136 respectivament. Llevat d'aquests pocs casos extrems, la variabilitat sembla mantenir-se constant a mesura que el temps avança, motiu pel qual concloem que les 3 mostres provenen de sèries temporals estacionàries. En particular, si per a tot enter t definim:

$$Z_t^1 = (I_1 - B_1)(I_1 - B_1^{365})X_t^1,$$

$$Z_t^2 = (I_2 - B_2)(I_2 - B_2^{12})X_t^2 \text{ i}$$

$$Z_t^3 = (I_3 - B_3)(I_3 - B_3^{12})X_t^3,$$

llavors les sèries temporals

$$\{Z_t^1\}_{t\in\mathbb{Z}}, \{Z_t^2\}_{t\in\mathbb{Z}} \text{ i } \{Z_t^3\}_{t\in\mathbb{Z}}$$

són estacionàries i s'ajusten als 3 gràfics d'adalt. És a dir, amb la mateixa notació que en l'exemple 1.4,

Temperatura sense estacionalitat diferenciada:	$\{Z_{367}^1(w_0^1), ..., Z_{1095}^1(w_0^1)\}$.
Vendes sense estacionalitat diferenciades:	$\{Z_{14}^2(w_0^2), ..., Z_{64}^2(w_0^2)\}$.
Trasllats sense estacionalitat diferenciats:	$\{Z_{14}^3(w_0^3), ..., Z_{138}^3(w_0^3)\}$.

D'aquesta forma, si $\{X_t\}_{t\in\mathbb{Z}}$ és la sèrie temporal que admet descomposicó clàssica de la que provenen les observacions $\{X_1(w_0), ..., X_n(w_0)\}$, sabent el període de la component estacionalitat podem transformar la sèrie en una altra que admet descomposició clàssica i que no presenta estacionalitat. Posteriorment, podem transformar aquesta darrera sèrie en una estacionària. Finalment, reordenant els índexos, tindrem una mostra $\{Z_1(w_0), ..., Z_m(w_0)\}$, (amb $m \leq n$), provenent d'una sèrie temporal estacionària $\{Z_t\}_{t\in\mathbb{Z}}$. En conclusió, podem reduir el problema a l'estudi de les sèries temporals estacionàries, ja que a partir de l'estudi de $\{Z_t\}_{t\in\mathbb{Z}}$ podrem extrapolar la informació a la sèrie original $\{X_t\}_{t\in\mathbb{Z}}$, com veurem en el capítol 4. D'aquesta forma, en els següents capítols ens centrarem en l'estudi dels processos estacionaris.

2 Processos estacionaris

Seguint els mètodes explicats en el capítol anterior, en tot aquest capítol podem suposar que tenim un nombre finit n d'observacions, que provenen d'un procés estacionari $\{X_t\}_{t\in\mathbb{Z}}$. És a dir, que $\exists w_0 \in \Omega$, tal que les nostres dades corresponen a

$$\{X_1(w_0), ..., X_n(w_0)\}.$$

Per tant, ens centrarem en l'estudi del procés estacionari $\{X_t\}_{t\in\mathbb{Z}}$.

2.1 Predicció en els models estacionaris

En aquesta secció, $\{X_t\}_{t\in\mathbb{Z}}$ és una sèrie temporal estacionària (amb una mostra de n dades $\{X_1(w_0), ..., X_n(w_0)\}$) amb mitjana μ_X i funcions d'autocovariància $\gamma_X(.)$ i d'autocorrelació $p_X(.)$ conegudes.

El nostre objectiu és, amb només aquesta informació, intentar predir observacions futures a partir de combinacions lineals de les dades observades. És a dir, sigui $h \in \mathbb{N}$ fixat, trobar la variable aleatòria pertanyent al subespai

$$F =< 1, X_1, ..., X_n >\subseteq L^2(\Omega),$$

(on 1 és la variable aleatoria constant igual a 1), a la que direm el predictor lineal d'F i denotarem per $P_F X_{n+h}$, que millor aproximi X_{n+h}. Equivalentment, determinar els coeficients reals $\{a_0, .., a_n\}$ per a trobar la variable aleatòria

$$P_F X_{n+h} = a_0 + \sum_{i=1}^{n} a_i X_{n+1-i},$$

que minimitzi l'aplicació no negativa

$$\begin{aligned} S : \mathbb{R}^{n+1} &\longrightarrow \mathbb{R} \\ (a_0, ..., a_n) &\longmapsto E\left((X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i})^2\right). \end{aligned}$$

Fixem-nos que $F =< 1, X_1, ..., X_n >$, és un subespai de dimensió finita (i per tant tancat) de $L^2(\Omega)$. Com $(L^2(\Omega), < . >)$ és un espai de Hilbert, utilitzant el teorema de descomposició ortogonal [2] , sabem que per a tota variable aleatòria $x \in L^2(\Omega)$, existeix un únic $y_x \in F$ i un únic $z_x \in F^\perp$, tal que

$$x = y_x + z_x, \text{ i } d(x, F) = d(x, y_x) = \sqrt{< x - y_x >} = \sqrt{E((x - y_x)^2)}.$$

Aquest resultat ens indica que $P_F X_{n+h}$ és definit de forma única g.p.t., o equivalentment, que l'aplicació S té un mínim absolut $(a_0, ..., a_n) \in \mathbb{R}^{n+1}$. Ara bé, com aquest mínim és

[2]Teorema de descomposició ortogonal: Sigui H un espai de Hilbert i sigui $Y \subset H$ un subespai tancat. Llavors, donat $x \in H$, existeix un únic $y_x \in Y$ i un únic $z_x \in Y^\perp$ tal que $x = y_x + z_x$. A més es té:

1. y_x és l'únic vector de Y que compleix $x - y_x \perp Y$.
2. $d(x, Y) = |z_x|$ i $d(x, Y^\perp) = |y_x|$.
3. $Y = Y^\perp$.

únic g.p.t, podria haver-hi més d'una sol·lució, pero totes elles derivarien en predictors lineals que serien iguals llevat d'un conjunt amb probabilitat 0.

Degut al teorema de descomposició ortogonal, $P_F X_{n+h}$ és l'única variable de F tal que

$$X_{n+h} - P_F X_{n+h} \in F^{\perp}.$$

Per tant, per a trobar els coeficients $(a_0, ..., a_n) \in \mathbb{R}^{n+1}$ que minimitzin S, és suficient imposar

$$< X_{n+h} - P_F X_{n+h}, 1 >= 0, \text{ i que per a } 1 \leq i \leq n, \ < X_{n+h} - P_F X_{n+h}, X_i >= 0.$$

Respecte a la variable aleatòria 1, fent ús de les propietats elementals del producte escalar:

$$\begin{aligned} <X_{n+h} - P_F X_{n+h}, 1 > &= < X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}, 1 > \\ &= E(X_{n+h}) - a_0 - \sum_{j=1}^{n} a_j E(X_{n+1-i}) = \mu_X - \mu_X \sum_{j=1}^{n} a_j - a_0. \end{aligned}$$

Per tant, cal imposar que

$$a_0 = \mu_X(1 - \sum_{j=1}^{n} a_j).$$

Respecte a X_i, per a $1 \leq i \leq n$, i utilitzant que $a_0 = \mu_X(1 - \sum_{j=1}^{n} a_j)$:

$$\begin{aligned} <X_{n+h} - P_F X_{n+h}, X_i > &= < X_{n+h} - a_0 - \sum_{i=1}^{n} a_i X_{n+1-i}, X_i > \\ &= E(X_{n+h} X_i) - a_0 E(X_i) - \sum_{j=1}^{n} a_j E(X_{n+1-j} X_i) \\ &= Cov(X_{n+h}, X_i) + \mu_X^2 - \left(\sum_{j=1}^{n} a_j \left(Cov(X_{n+1-j}, X_i) + \mu_X^2 \right) \right) - a_0 \mu_X \\ &= \gamma_X(n+h-i) - \sum_{j=1}^{n} a_j \gamma_X(n+1-j-i) - \mu_X^2 \sum_{j=1}^{n} a_j + \mu_X^2 - a_0 \mu_X \\ &= \gamma_X(n+h-i) - \sum_{j=1}^{n} a_j \gamma_X(n+1-j-i) + \mu_X a_0 - a_0 \mu_X. \end{aligned}$$

És a dir, cal imposar, per a $1 \leq i \leq n$, que

$$\gamma_X(n+h-i) = \sum_{j=1}^{n} a_j \gamma_X(n+1-j-i).$$

Fixem-nos que els coeficients $(a_1, ..., a_n)$ són sol·lució del sistema d'equacions lineal

$$\Gamma_n a_n^* = \gamma_n(h).$$

Γ_n és la matriu real definida per $(\gamma_X(i-j))_{1 \leq i,j \leq n}$, $a_n^* = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$,

$\gamma_n(h) = \begin{pmatrix} \gamma_X(h) \\ \gamma_X(h+1) \\ \vdots \\ \gamma_X(h+n-1) \end{pmatrix}$, i finalment, $a_0 = \mu_X(1 - \sum_{j=1}^{n} a_j)$.

El teorema de descomposició ortogonal ens assegura que el sistema $\Gamma_n a_n^* = \gamma_n(h)$, és compatible, i que en el cas en que tingui més d'una sol·lució, tots els predictors lineals derivats d'elles seran únics g.p.t.

Observació 2.1. Per assegurar la unicitat dels coeficients $\{a_0, ..., a_n\}$ cal imposar que la matriu Γ_n sigui invertible.

Observació 2.2. Degut a les hipòtesis del Teorema de la projecció ortogonal, considerant qualsevol subespai G tancat de $L^2(\Omega)$, tenim que $P_G X_{n+h}$ està ben definit i és únic g.p.t.

Un cop calculat aquests coeficients, sabem com és explícitament $P_F X_{n+h}$, i per a estimar X_{n+h} simplement cal aplicar el predictor lineal de F a la nostra mostra, és a dir, calcular $P_F X_{n+h}(w_0)$ a partir de les nostres dades $\{X_1(w_0), ..., X_n(w_0)\}$, essent aquest el valor numèric que estimi $X_{n+h}(w_0)$.

A continuació recollirem algunes de les propietats dels predictors lineals. La 2 i la 4 són imprescindibles per a validar el model ARMA i calcular els seus residus en el següent capítol.

Propietats de $\mathbf{P_n X_{n+h}}$

1) $P_F X_{n+h} = \mu_X + \sum_{i=1}^{n} a_i (X_{n+1-i} - \mu_X)$ on $a_n^* = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ satisfà l'equació $\Gamma_n a_n^* = \gamma_n(h)$,

2) $E\left((X_{n+h} - P_F X_{n+h})^2\right) = \gamma_X(0) - a_n^{*T} \gamma_n(h)$,

3) $E(X_{n+h} - P_F X_{n+h}) = 0$,

4) $E[(X_{n+h} - P_F X_{n+h}) X_j] = 0,\ j = 1, .., n$.

Demostració. 1) Del raonament anterior sabem que $P_F X_{n+h} = a_0 + \sum_{i=1}^{n} a_i X_{n+1-i}$ on $a_n^* = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, satisfà l'equació $\Gamma_n a_n^* = \gamma_n(h)$, i $a_0 = \mu_X (1 - \sum_{j=1}^{n} a_j)$, del que deduïm la primera propietat.

2)

$$E\left((X_{n+h} - P_F X_{n+h})^2\right) = E\left[X_{n+h}^2 - 2X_{n+h}\left(\mu_X + \sum_{i=1}^{n} a_i(X_{n+1-i} - \mu_X)\right)\right.$$

$$\left. + \left(\mu_X + \sum_{i=1}^{n} a_i(X_{n+1-i} - \mu_X)\right)^2\right]$$

$$= E\left[X_{n+h}^2 - 2X_{n+h}\mu_X - 2\sum_{i=1}^{n} a_i(X_{n+h}X_{n+1-i} - X_{n+h}\mu_X) + \mu_X^2\right.$$

$$\left. + 2\mu_X \sum_{i=1}^{n} a_i(X_{n+1-i} - \mu_X) + \left(\sum_{i=1}^{n} a_i(X_{n+1-i} - \mu_X)\right)^2\right]$$

$$= \gamma_X(0) + \mu_X^2 - 2\mu_X^2 - 2\sum_{i=1}^{n} a_i\left(\gamma_X(h+i-1) + \mu_X^2 - \mu_X^2\right) + \mu_X^2$$

$$+ E\left(\left(\sum_{i=1}^{n} a_i(X_{n+1-i} - \mu_X)\right)\left(\sum_{j=1}^{n} a_j(X_{n+1-j} - \mu_X)\right)\right)$$

$$= \gamma_X(0) - 2\sum_{i=1}^{n} a_i\gamma_X(h+i-1) + \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\gamma_X(i-j)a_j$$

$$= \gamma_X(0) - \left(2\sum_{i=1}^{n} a_i\gamma_X(h+i-1) - \sum_{i=1}^{n} a_i \sum_{j=1}^{n} a_j\gamma_X(i-j)\right)$$

$$= \gamma_X(0) - \left(2\sum_{i=1}^{n} a_i\gamma_X(h+i-1) - \sum_{i=1}^{n} a_i\gamma_X(h+i-1)\right).$$

I per les definicions de a_n^* i Γ_n, l'expressió anterior és igual a

$$\gamma_X(0) - a_n^{*T}\gamma_n(h).$$

3)

$$E(X_{n+h} - P_F X_{n+h}) = E\left(X_{n+h} - \mu_X - \sum_{i=1}^{n} a_i(X_{n+1-i} - \mu_X)\right)$$

$$= \mu_X - \mu_X - 0 = 0.$$

4) És conseqüència de $X_{n+h} - P_F X_{n+h} \in F^{\perp}$.

□

Observació 2.3. En realitat la metodologia i les propietats descrites les podem aplicar a qualsevol sèrie temporal amb moments de segon ordre finit. Efectivament, suposem que $\{X_t\}_{t\in\mathbb{Z}}$ és una sèrie temporal qualsevol de la qual coneixem les seves propietats de segon ordre, és a dir les funcions $\mu_X(h)$ i $\gamma_X(i,j)$ definides a la secció 1. Sigui $h \in \mathbb{N}$. Si, considerant $F =< 1, X_1, ..., X_n >$, volem calcular $P_F X_{n+h}$, prosseguint de la mateixa forma que hem definit abans, aribem a la conclusió que

$P_F X_{n+h} = \mu_X(n+h) + \sum_{i=1}^{n} a_i(X_{n+1-i} - \mu_X(n+1-i))$, on $a_n^* = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, satisfà l'equació

$$\Gamma_n a_n^* = \gamma_n(h).$$

En aquest cas Γ_n és la matriu real definida per $(\gamma_X(n+1-i, n+1-j))_{1\leq i,j\leq n}$, i

$$\gamma_n(h) = \begin{pmatrix} \gamma_X(n+h,n) \\ \gamma_X(n+h,n-1) \\ \vdots \\ \gamma_X(n+h,1) \end{pmatrix}.$$

2.1.1 Els predictors pas a pas

Com acabem de veure, els predictors lineals es poden definir per a qualsevol sèrie temporal amb moments de segon ordre finits. Tot i que nosaltres ens centram amb les sèries estacionàries, anem a definir els predictors pas a pas per a qualsevol sèrie en general tal que per a tot enter t, $\mu_X(t) = 0$. L'objectiu d'aquests és justificar que les dades provenen de la sèrie temporal $\{X_t\}_{t\in\mathbb{Z}}$. Per ara ens centrarem en definir-los, i en la següent secció, quan estudiem els models ARMA, veurem la seva utilitat.

Fixem-nos que amb el mateix mètode de l'apartat anterior, fixat $h \geq 0$, per a tot $1 \leq m \leq n$, podem considerar el subespai tancat de dimensió finita

$$F_m =< 1, X_1, ..., X_m >$$

i calcular $P_{F_m}X_{m+h}$. És a dir, calcular el predictor lineal de X_{m+h} a partir del subespai F_m. Aquesta idea és fonamental per a calcular els predictors pas a pas.

Sigui $\{X_t\}_{t\in\mathbb{Z}}$, una sèrie temporal qualsevol amb moments de segon ordre finits, de la que tenim un nombre n de dades observades, i de la que coneixem les funcions $\mu_X(h)$ i $\gamma_X(i,j)$. Definim els predictors pas a pas com a:

$$\hat{X_m} = \begin{cases} 0, & \text{si } m = 1, \\ P_{F_{m-1}}X_m, & \text{si } m \in \{2, .., n\} \subseteq \mathbb{Z}, \\ P_{F_n}X_m, & \text{si } m \geq n+1. \end{cases}$$

Fixem-nos que si $m \in \{2, .., n\}$, $\hat{X_m}$ representa la predicció de la variable aleatòria X_m com a combinació lineal del subespai $F_{m-1} =< 1, X_1, ..., X_{m-1} >$. Si $m > n$, llavors $\hat{X_m}$, és la predicció de X_m com a combinació lineal de $F =< 1, X_1, ..., X_n >$.

Introduïm també, per a tot i $\in \{1, ..., n\}$, els errors quadràtics $v_{i-1} = E\left((X_i - \hat{X}_i)^2\right)$, els quals sabem calcular per la propietat 2, i els errors de predicció $U_i = X_i - \hat{X}_i$, .

Degut a 2.1 i a que $\mu_X = 0$, tenim que $\hat{X}_i$ és combinació lineal de $< X_1, ..., X_{i-1} >$. És a dir, si $U_n^* = (U_1, ..., U_n)^T$ i $X_n^* = (X_1, ..., X_n)^T$, llavors existeix una matriu real

$$A_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{11} & 1 & 0 & \cdots & 0 \\ a_{22} & a_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ a_{n-1,n-1} & a_{n-1,n-2} & a_{n-1,n-3} & \cdots & 1 \end{pmatrix},$$

que compleix $U_n^* = A_n X_n^*$. A més, com A_n és triangular, $|A_n| = 1$, amb el que deduïm que A_n és invertible i per tant $\exists$ una matriu real

$$C_n = A_n^{-1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ \theta_{11} & 1 & 0 & \cdots & 0 \\ \theta_{22} & \theta_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \theta_{n-1,n-1} & \theta_{n-1,n-2} & \theta_{n-1,n-3} & \cdots & 1 \end{pmatrix},$$

que compleix $X_n^* = C_n U_n^*$.

Definint $\Theta_n = C_n - Id$, el vector de predictors pas a pas $\hat{X}_n^* = (\hat{X}_1, \hat{X}_2, \hat{X}_3, ..., \hat{X}_n)^T$, es pot expressar matricialment com

$$\hat{X}_n^* = X_n^* - U_n^* = C_n U_n^* - U_n^* = (C_n - Id)U_n^* = \Theta_n U_n^*.$$

Equivalentment,

$$\hat{X_{m+1}} = \begin{cases} 0, & \text{si } m = 0, \\ \displaystyle\sum_{j=1}^{m} \theta_{mj}(X_{m-j+1} - \hat{X}_{m-j+1}), & \text{si } m \in \{1, 2, .., n-1\}. \end{cases}$$

Fixem-nos que els coeficients $\{a_{ij}\}$,(i per tant els $\{\theta_{ij}\}$), únicament depenen del nombre de dades observades (n), de $\mu_X(h)$ i de $\gamma_X(i, j)$.

Observació 2.4. Per a $i, j \in \{1, ..., n\}, i \neq j$,

$$Cov(X_i - \hat{X}_i, X_j - \hat{X}_j) = 0.$$

Vegem-ho: Suposem, sense pèrdua de generalitat, que $i > j$. És suficient veure que $E((X_i - \hat{X}_i)(X_j - \hat{X}_j)) = 0$, ja que a causa de la propietat 3, $\forall k \in \{1, ..., n\}$,

$$E(X_k - \hat{X}_k) = 0.$$

Fixem-nos que per la propietat 1,

$$X_j - \hat{X}_j = X_j - \mu_X - \sum_{k=1}^{j-1} a_k(X_{j-k} - \mu_X),$$

per tant, fent ús de les propietats 3 i 4

$$\begin{aligned} E\Big((X_i - \hat{X}_i)(X_j - \hat{X}_j)\Big) &= E\Big((X_i - \hat{X}_i)X_j\Big) - \mu_X E\Big(X_i - \hat{X}_i\Big) \\ &- \sum_{k=1}^{j-1} \Big(a_k E((X_i - \hat{X}_i)X_{j-k}) - \mu_X E(X_i - \hat{X}_i)\Big) \\ &= 0. \end{aligned}$$

2.2 Processos lineals

Els processos lineals són els processos estacionaris principals que tractarem. Anem a veure en què consisteixen:

Notació 2. *Diem que la sèrie temporal* $\{Z_t\}_{t\in\mathbb{Z}} \sim WN(0, \sigma^2)$, *si les variables aleatòries* Z_t *compleixen les següents propietas, per a tots enters* t, t_1 *i* t_2*:*

- $E(Z_t^2) = \sigma^2 < \infty$.
- $E(Z_t) = 0$.
- $Cov(Z_{t_1}, Z_{t_2}) = 0$.

Observació 2.5. Seguint amb la notació anterior, en general, la distribució que segueixen les variables aleatòries Z_t és desconeguda. En el cas en què, per a tot enter t, $Z_t \sim N(0,\sigma^2)$, fixem-nos que les variables aleatòries Z_t serien independents.

Definició 2.6. *La sèrie temporal* $\{X_t\}_{t\in\mathbb{Z}}$ *és un procés lineal si*

$$X_t = \sum_{j=-\infty}^{\infty} \delta_j Z_{t-j},$$

per a tot $t \in \mathbb{Z}$, *on* $\{Z_t\}_{t\in\mathbb{Z}} \sim WN(0,\sigma^2)$, *amb* $\delta_j \in \mathbb{R}, \forall j \in \mathbb{Z}$ *i* $\sum_{j=-\infty}^{\infty} |\delta_j| < \infty$.

La següent proposició justifica que els processos lineals estiguin ben definits i que siguin estacionaris.

Proposició 2.7. *Sigui* $\{Y_t\}_{t\in\mathbb{Z}}$ *una sèrie temporal estacionària amb mitjana* $\mu_Y = 0$ *i funció de covariància* $\gamma_Y(.)$. *Si* $\{\chi_i\}_{i\in\mathbb{Z}}$ *és una successió de reals tals que* $\sum_{j=-\infty}^{\infty} |\chi_j| < \infty$, *llavors la sèrie temporal* $\{X_t\}_{t\in\mathbb{Z}}$, *definida per a tot enter* t, *com a*

$$X_t = \sum_{j=-\infty}^{\infty} \chi_j Y_{t-j},$$

és estacionària, amb mitjana $\mu_X = 0$, *i funció d'autocovariància*

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \chi_j \chi_k \gamma_Y(h+k-j) < \infty, \text{ per a tot enter } h.$$

En particular, si $\{X_t\}_{t\in\mathbb{Z}}$ *és un procés lineal, és a dir, si* $\{Y_t\}_{t\in\mathbb{Z}} \sim WN(0,\sigma^2)$, *llavors*

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \chi_j \chi_{j+h} \sigma^2,$$

i

$$\sum_{h=-\infty}^{\infty} |\gamma_X(h)| < \infty.$$

Demostració. En primer lloc vegem que la sèrie és convergent quasi segurament.

La condició $\sum_{j=-\infty}^{\infty} |\chi_j| < \infty$, implica que les variables aleatòries $\{X_t\}$ prenguin valors reals amb probabilitat 1, és a dir, que per a tot enter t, la successió de variables aleatòries

$$\{X_{t_n} = \sum_{j=-n}^{n} \chi_j Y_{t-j}\}_{n\geq 1},$$

convergeixi quasi segurament a X_t. Vegem-ho:

Sigui $t \in \mathbb{Z}$. Considerem la variable aleatòria no negativa

$$Z_t : \Omega \longmapsto \bar{\mathbb{R}},$$

expressada com a $Z_t = \sum_{j=-\infty}^{\infty} |\chi_j||Y_{t-j}|$, la qual està ben definida i és mesurable com a conseqüència del Teorema de Beppo Levi [3]. A més, fent ús d'aquest mateix teorema, tenim que

$$E(Z_t) = \sum_{j=-\infty}^{\infty} |\chi_j| E(|Y_{t-j}|).$$

[3] Teorema de Beppo Levi: Si $\{X_j\}_{j\geq 1}$ és una successió de variables aletòries no negatives, llavors $X = \sum_{j=1}^{\infty} X_j$ és una variable aleatòria no negativa i $E(X) = \sum_{j=1}^{\infty} E(X_j)$.

Per altra banda, fixem-nos que per a tot enter s, (Desigualtat de Schwarz):

$$E(|Y_s|) \leq \sqrt{E(Y_s^2)} = \sqrt{\gamma_Y(0)},$$

amb el que deduïm que

$$E(Z_t) \leq \sum_{j=-\infty}^{\infty} |\chi_j| \sqrt{\gamma_Y(0)} = \sqrt{\gamma_Y(0)} \sum_{j=-\infty}^{\infty} |\chi_j| < \infty.$$

Que Z_t tingui esperança finita, implica que $\exists \mathbb{A}_t \subseteq \Omega$, amb $\mathbb{P}(\mathbb{A}_t) = 1$, tal que $Z_t(\mathbb{A}_t) < \infty$. És a dir, $\forall w \in \mathbb{A}_t$,

$$Z_t(w) = \sum_{j=-\infty}^{\infty} |\chi_j||Y_{t-j}|(w) < \infty$$

Com que, per a tot $n \geq 1$,

$$|X_{t_n}(w)| \leq \sum_{j=-n}^{n} |\chi_j||Y_{t-j}|(w) \leq Z_t(w) < \infty,$$

deduim que, per a tot $w \in \mathbb{A}_t$, si $n \to \infty$, la sèrie de nombres reals $\{|X_{t_n}(w)|\}_{n\geq 1}$ covergeix, i com a conseqüència, la sèrie $\{X_{t_n}(w)\}_{n\geq 1}$ també (ja que si una sèrie real és asolutament convergent llavors és convergent). Concloem que, per a tot $w \in \mathbb{A}_t$,

$$X_t(w) = \lim_{n\to+\infty} X_{t_n}(w) = \sum_{j-\infty}^{\infty} \chi_j Y_{t-j}(w) \in \mathbb{R}.$$

Per tant, $\forall t \in \mathbb{Z}$, $\exists \mathbb{A}_t \subseteq \Omega$, amb $P(\mathbb{A}_t) = 1$, tal que $\forall w \in \Omega$,

$$\lim_{n\to+\infty} X_{t_n}(w) = X_t(w) \in \mathbb{R}.$$

Aquest fet implica que, per a tot enter t, $X_t : \Omega \longrightarrow \mathbb{R}$ és una variable aleatòria ben definida, i que $\{X_{t_n} = \sum_{j=-n}^{n} \chi_j Y_{t-j}\}_{n\geq 1}$ convergeix quasi segurament a X_t.

Calculem $\mu_X(t)$. Sigui $t \in \mathbb{Z}$. Amb la mateixa notació utilitzada fins ara, fixem-nos que $\{X_{t_n}\}_{n\geq 1}$ és una successió de variables aleatòries tals que g.p.t $w \in \Omega$, es compleix que

$$\lim_{n\to+\infty} X_{t_n} = X_t.$$

A més, per a tot $n \geq 1$

$$|X_{t_n}| \leq Z_t,$$

i

$$E(Z_t) \leq \sqrt{\gamma_Y(0)} \sum_{j=-\infty}^{\infty} |\chi_j| < \infty.$$

Fent ús del teorema de convergència dominada de Lebesgue [4], i degut a que per a tot enter s, $E(Y_s) = 0$, veiem que

$$E(X_t) = \sum_{j=-\infty}^{\infty} \chi_j E(Y_{t-j}) = 0.$$

[4]Teorema de la convergència dominada de Lebesgue: Sigui $\{X_j\}_{j\geq 1}$ una successió de variables aleatòries que convergeix puntualment g.p.t. cap a una variable aleatòria X, i que està dominada per una variable aleatòria no negativa integrable Y, aleshores les variables aleatòries $|X_j|$ i $|X|$ tenen esperança finita, i compleixen $\lim_{j\to+\infty} E(|X_j - X|) = 0$ i $E(X) = \lim_{j\to+\infty} E(X_j)$

Per tant, $\mu_X(t)$ està ben definida i és 0 per a tot enter t.

Calculem la funció d'autocovariància $\gamma_X(t, t+h) = E\left(X_{t+h}X_t\right)$. Siguin t i h dos enters qualsevols.

En aquest cas, tenim que g.p.t $w \in \Omega$,

$$\lim_{n \to +\infty} X_{(t+h)_n} X_{t_n} = X_{t+h} X_t,$$

on per a tot $n \geq 1$,

$$X_{(t+h)_n} X_{t_n} = \sum_{j=-n}^{n} \sum_{k=-n}^{n} \chi_j \chi_k Y_{t+h-j} Y_{t-k}.$$

Considerem també la variable aleatòria no negativa

$$R_{th} : \Omega \longmapsto \bar{\mathbb{R}},$$

definida per

$$R_{th} = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |\chi_j||\chi_k||Y_{t+h-j}||Y_{t-k}|,$$

la qual està ben definida i és mesurable com a conseqüència, de nou, del Teorema de Beppo Levi explicat abans. A més, fent ús d'aquest mateix teorema,

$$E(R_{th}) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |\chi_j||\chi_k| E\left(|Y_{t+h-j}||Y_{t-k}|\right).$$

Per altra banda, fixem-nos que per a tots enters j i k, (Desigualtat de Schwarz):

$$E\left(|Y_{t+h-j}||Y_{t-k}|\right) \leq \sqrt{E(Y_{t+h-j}^2)E(Y_{t-k}^2)} = \gamma_Y(0),$$

amb el que deduïm que

$$E(R_{th}) \leq \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} |\chi_j||\chi_k|\gamma_Y(0) = \gamma_Y(0) \sum_{j=-\infty}^{\infty} |\chi_j| \sum_{k=-\infty}^{\infty} |\chi_k| < \infty.$$

Llavors, per a tot $n \geq 1$

$$|X_{(t+h)_n} X_{t_n}| \leq R_{th},$$

i

$$E(R_{th}) < \infty.$$

Fent ús de nou del teorema de convergència dominada de Lebesgue , veiem que

$$\begin{aligned}
\gamma_X(t+h, t) = E(X_{t+h}X_t) &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \chi_j \chi_k E\left(Y_{t+h-j} Y_{t-k}\right) \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \chi_j \chi_k \gamma_Y(h+k-j) < \infty.
\end{aligned}$$

Per tant, $\gamma_X(t+h, t)$ està ben definida i és independent de t, amb el que deduïm que $\{X_t\}_{t \in \mathbb{Z}}$ és una sèrie temporal estacionària.

En particular, si $\{X_t\}_{t\in\mathbb{Z}}$ és un procés lineal, és a dir, si $\{Y_t\}_{t\in\mathbb{Z}} \sim WN(0,\sigma^2)$, tenim que

$$\gamma_X(h) = \begin{cases} 0, & \text{si } h \neq 0, \\ \sigma^2, & \text{si } h = 0. \end{cases}$$

Per tant, per a tot enter h,

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \chi_j \chi_{j+h} \sigma^2.$$

Per altra banda, com $\sum_{j=-\infty}^{\infty} |\chi_j| < \infty$, llavors

$$\left(\sum_{h=-\infty}^{\infty} |\chi_h|\right)\left(\sum_{j=-\infty}^{\infty} |\chi_j|\right) = \sum_{h=-\infty}^{\infty}\left(|\chi_h| \sum_{j=-\infty}^{\infty} |\chi_j|\right) < \infty.$$

Per tant,

$$\begin{aligned}\sum_{h=-\infty}^{\infty} |\gamma_X(h)| &\leq \sigma^2 \sum_{h=-\infty}^{\infty}\left(\sum_{j=-\infty}^{\infty} |\chi_j||\chi_{j+h}|\right) \\ &= \sigma^2 \sum_{h=-\infty}^{\infty}\left(|\chi_h| \sum_{j=-\infty}^{\infty} |\chi_j|\right) < \infty.\end{aligned}$$

□

Observació 2.8. Fixem-nos que, amb la mateixa notació dels subespais i de l'operador B de 1.3, si $\{X_t\}_{t\in\mathbb{Z}}$ és una sèrie temporal estacionària amb mitjana 0, i M el subespai de $L^2(\Omega)$ generat per les variables aleatòries del procés, llavors qualsevol sèrie de potències a coeficients reals

$$\phi(z) = \sum_{j=-\infty}^{\infty} \phi_j z^j,$$

tal que $\sum_{j=-\infty}^{\infty} |\phi_j| < \infty$, la podem identificar amb l'operador:

$$\phi(B) = \sum_{j=-\infty}^{\infty} \phi_j B^j,$$

el qual, per la proposició anterior, està ben definit , i, per com està definit, és lineal sobre M. És a dir, $\phi(B) \in \mathbb{L}(M)$. A més, exactament igual que amb les sèries finites, podem relacionar la suma i producte de sèries amb la suma i composició d'operadors.

Efectivament, si

$$\phi(B) = \sum_{j=-\infty}^{\infty} \phi_j B^j, \text{ i } \theta(B) = \sum_{j=-\infty}^{\infty} \phi_j B^j,$$

són dos operadors de $\mathbb{L}(M)$, amb $\sum_{j=-\infty}^{\infty} |\phi_j| < \infty$ i $\sum_{j=-\infty}^{\infty} |\theta_j| < \infty$, llavors l'operador $\phi(B) \circ \theta(B)$ es pot expressar com a:

$$\psi(B) = \sum_{j=-\infty}^{\infty} \phi_j B^j \circ \sum_{j=-\infty}^{\infty} \theta_j B^j = \sum_{j=-\infty}^{\infty}\left(\sum_{k=-\infty}^{\infty} \phi_k \theta_{j-k}\right) B^j = \sum_{j=-\infty}^{\infty} \psi_j B^j,$$

on

$$\sum_{j=-\infty}^{\infty} |\psi_j| \le \left(\sum_{j=-\infty}^{\infty} |\phi_j|\right) \left(\sum_{k=-\infty}^{\infty} |\theta_k|\right) < \infty,$$

el que implica, per la proposició anterior, que està ben definit i que pertany a $\mathbb{L}(M)$.

Aquest fet, juntament amb que $\mathbb{L}(M)$ és un espai vectorial, implica que poguem manipular (amb combinacions lineals i composant) aquests tipus d'operadors com si de sèries de potències es tractessin, de forma que l'operador resultant continui pertanyent a $\mathbb{L}(M)$.

Per altra banda, si tenim un nombre $k \geq 0$ de processos estacionaris:

$$\{X_t^1\}_{t\in\mathbb{Z}},\ \ldots\ ,\ \{X_t^k\}_{t\in\mathbb{Z}},$$

podem ampliar tots els conceptes explicats al subespai generat per totes les variables aleatòris dels k processos,

$$M_2 =< \{X_t^1\}_{t\in\mathbb{Z}} \cup\ \ldots\ \cup\ \{X_t^k\}_{t\in\mathbb{Z}} >\subseteq L^2(\Omega),$$

i a l'operador $B \in \mathbb{L}(M_2)$, definit per a tot $i \in \{1, ..., k\}$, i per a tot enter t, com a:

$$\begin{aligned} B : M_2 &\longrightarrow M_2 \\ X_t^i &\longmapsto X_{t-1}^i. \end{aligned}$$

Recordem que nosaltres partim d'una sèrie estacionària $\{X_t\}_{t\in\mathbb{Z}}$, que ajusta les nostres dades $\{X_1(w_0), ..., X_n(w_0)\}$. Fixem-nos, que definint la sèrie temporal $\{Y_t\}_{t\in\mathbb{Z}}$, on, per a tot enter t,

$$Y_t = X_t - \mu_X,$$

resulta que aquesta és una sèrie estacionària amb mitjana 0 i, per a tot enter h,

$$\gamma_Y(h) = \gamma_X(h).$$

Com els processos lineals són les sèries estacionàries (amb mitjana 0) més usuals, nosaltres suposarem que la sèrie $\{Y_t\}_{t\in\mathbb{Z}}$, és un procés lineal. Fixem-nos, que, d'aquesta forma, per a tot enter t,

$$X_t = \mu_X + Y_t,$$

on $\{Y_t\}_{t\in\mathbb{Z}}$ és un procés lineal.

Exemple 2.9. Aplicada aquesta idea als nostres tres exemples, i amb la mateixa notació utilitzada en el capítol 1, considerem que, per a $k \in \{1, 2, 3\}$,

$$\{Z_t^k\}_{t\in\mathbb{Z}} = \{W_t^k + \mu_{Z^k}\}_{t\in\mathbb{Z}},$$

on μ_{Z^k} és la mitjana de la sèrie estacionària $\{Z_t^k\}_{t\in\mathbb{Z}}$, i $\{W_t^k\}_{t\in\mathbb{Z}}$ és un procés lineal.

2.3 Estimació de la mitjana d'una sèrie temporal estacionària

Disposem de n dades observades que provenen d'una sèrie estacionària $\{X_t\}_{t\in\mathbb{Z}}$, és a dir, que $\exists w_0 \in \Omega$, tal que les dades corresponen a $\{X_1(w_0), ..., X_n(w_0)\}$. Aquesta sèrie, normalment, té mitjana μ_X desconeguda. El nostre objeciu en aquesta secció és estimar-la, per a restar-li a la sèrie temporal, i així poder suposar que aquesta té mitjana 0, (en particular que és un procés lineal).

Definició 2.10. *Sigui n la grandària de la mostra, i $\{X_t\}_{t\in\mathbb{Z}}$ la sèrie temporal estacionària, de la qual provenen les n dades observades. Definim el següent estadístic:*

$$\bar{X_n} = \frac{1}{n}\sum_{t=1}^{n} X_t.$$

$\bar{X_n}$ es diu la mitjana mostral (SM).

Proposició 2.11. *Sigui n la grandària de la mostra i $\{X_t\}_{t\in\mathbb{Z}}$ una sèrie temporal estacionària amb mitjana μ_X, funció d'autocovariància γ_X, i $\bar{X_n} = \frac{1}{n}\sum_{k=1}^{n} X_k$. Llavors:*

1. *$\bar{X_n}$ és un estimador de μ_X sense biaix. És a dir, $E(\bar{X_n}) = \mu_X$.*

2. *$Var(\bar{X_n}) = E\left((\bar{X_n} - \mu_X)^2\right) = n^{-1}\sum_{h=-n}^{n}(1 - \frac{|h|}{n})\gamma_X(h)$.*

Demostració. 1. És inmediat, ja que per a tot enter t, $E(X_t^2) < \infty$, (l'esperança està ben definida) i per la linealitat d'aquesta:

$$E(\bar{X_n}) = \frac{n\mu_X}{n} = \mu_X.$$

2. Aplicant de nou la linealitat de l'esperança:

$$\begin{aligned}
Var(\bar{X_n}) &= E\left((\bar{X_n} - \mu_X)^2\right) = E\left(\bar{X_n}^2 - 2\mu_X\bar{X_n} + \mu_X^2\right) \\
&= \frac{1}{n}E\left(\frac{1}{n}\sum_{k=1}^{n} X_k^2 + \frac{1}{n}\sum_{i\neq j} X_iX_j\right) - \mu_X^2 \\
&= \frac{1}{n}\left(\gamma_X(0) + \mu_X^2 + \frac{1}{n}\sum_{i\neq j}\left(Cov(X_i, X_j) + \mu_X^2\right)\right) - \mu_X^2 \\
&= \frac{1}{n}\left(\gamma_X(0) + \mu_X^2 + \left(\frac{1}{n}\sum_{h=-n,h\neq 0}^{n}(n - |h|)\gamma_X(h)\right) + \frac{n(n-1)}{n}\mu_X^2\right) - \mu_X^2 \\
&= \frac{1}{n}\left(\left(\sum_{h=-n}^{n}(1 - \frac{|h|}{n})\gamma_X(h)\right) + n\mu_X^2\right) - \mu_X^2 = \frac{1}{n}\left(\sum_{h=-n}^{n}(1 - \frac{|h|}{n})\gamma_X(h)\right)
\end{aligned}$$

.

□

De la proposició anterior deduïm el següent corol·lari:

Corol·lari 2.12. *Amb les condicions de la proposició anterior, si $\sum_{h=-\infty}^{\infty}|\gamma_X(h)| < \infty$:*

$$\lim_{n\to+\infty} nVar(\bar{X_n}) = \sum_{|h|<\infty}\gamma_X(h).$$

En particular,

$$\lim_{n\to+\infty} Var(\bar{X_n}) = 0.$$

Sabem que, de l'anterior secció, el procés estacionari $\{X_t\}_{t\in\mathbb{Z}}$, que ajusta les nostres dades, compleix que, per a tot enter t,

$$X_t = \mu_X + Y_t,$$

on $\{Y_t\}_{t\in\mathbb{Z}}$ és un procés lineal. En concret,

$$\gamma_X = \gamma_Y,$$

i com, per la proposició 2.7, $\sum_{h=-\infty}^{\infty}|\gamma_Y(h)| < \infty$, llavors $\sum_{h=-\infty}^{\infty}|\gamma_X(h)| < \infty$. Fixemnos que, el Corolari 2.12, ens indica que $\bar{X_n}$, assimptòticament, estima amb exactitud la funció μ_X. És a dir, que per a n prou gran, g.p.t $w \in \Omega$,

$$\bar{X_n}(w) \simeq \mu_X.$$

Amb aquest raonament, aplicant aquest estadístic a w_0, és a dir, a la nostra mostra, calculant $\bar{X_n}(w_0) = n^{-1}\sum_{j=1}^{n} X_j(w_0)$, obtenim una bona aproximació de μ_X. Considerant que es tracta d'una igualtat, és a dir, que $\bar{X_n}(w_0) = \mu_X$, podem treballar amb el procés lineal $\{Y_t\}_{t\in\mathbb{Z}}$ definit per

$$Y_t = X_t - \mu_X = X_t - \bar{X_n}(w_0),$$

del qual disposem de les dades

$$\{Y_1(w_0), ..., Y_n(w_0)\} = \{X_1(w_0) - \bar{X_n}(w_0), ..., X_n(w_0) - \bar{X_n}(w_0)\}.$$

Exemple 2.13. Seguint amb els exemples de l'anterior secció, recordem que disposem de 3 conjunt de dades provenents de 3 sèries temporals estacionàries amb mitjanes desconegudes.

Nosaltres coneixem els valors

$$\begin{aligned}
&\{Z^1_{367}(w^1_0), ..., Z^1_{1095}(w^1_0)\}, \text{ (Temperatura de Madrid)},\\
&\{Z^2_{14}(w^2_0), ..., Z^2_{64}(w^2_0)\}, \text{ (Vendes), i}\\
&\{Z^3_{14}(w^2_0), ..., Z^3_{138}(w^2_0)\}, \text{ (Trasllats a la UCI)},
\end{aligned}$$

corresponents a les sèries temporals

$$\begin{aligned}
&\{Z^1_t\}_{t\in\mathbb{Z}}, \text{ amb mitjana } \mu_{Z^1},\\
&\{Z^2_t\}_{t\in\mathbb{Z}}, \text{ amb mitjana } \mu_{Z^2}, \text{ i}\\
&\{Z^3_t\}_{t\in\mathbb{Z}}, \text{ amb mitjana } \mu_{Z^3}.
\end{aligned}$$

Considerant els estadístics

$$\begin{aligned}
\bar{Z^1_{729}} &= \frac{1}{729}\sum_{t=367}^{1095} Z^1_t,\\
\bar{Z^2_{51}} &= \frac{1}{51}\sum_{t=14}^{64} Z^2_t, \text{ i}\\
\bar{Z^3_{125}} &= \frac{1}{125}\sum_{t=14}^{138} Z^3_t,
\end{aligned}$$

i aplicant-los a les nostres dades, és a dir, a w_0^1, a w_0^2, i a w_0^3, respectivament, obtenim unes estimacions de les tres mitjanes. Admetent que

$$\begin{aligned}
\mu_{Z^1} &= \bar{Z}^1_{729}(w_0^1) = -0.001371742,\\
\mu_{Z^2} &= \bar{Z}^2_{51}(w_0^2) = 33.60784, \text{ i}\\
\mu_{Z^3} &= \bar{Z}^3_{125}(w_0^3) = -0.28,
\end{aligned}$$

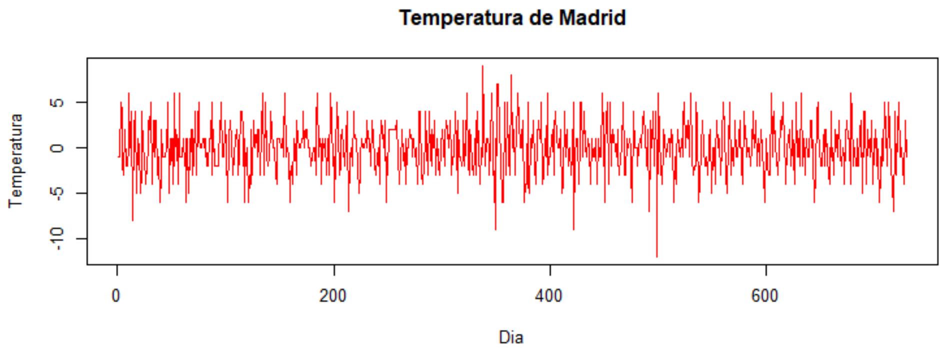
i definint, per a tot enter t,

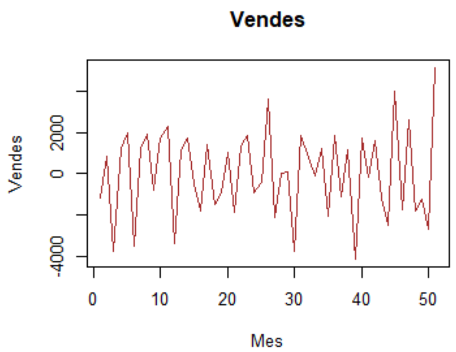
$$\begin{aligned}
W_t^1 &= Z_t^1 + 0.001371742,\\
W_t^2 &= Z_t^2 - 33.60784, \text{ i}\\
W_t^3 &= Z_t^3 + 0.28,
\end{aligned}$$

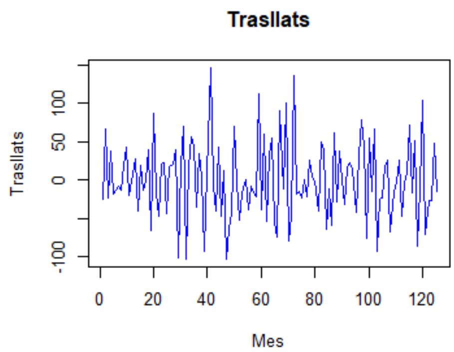
llavors, per l'exemple 2.9, $\{W_t^k\}_{t\in\mathbb{Z}}$, per a $k \in \{1,2,3\}$, és un procés lineal. Fixem-nos, que disposem de les dades

$$\begin{aligned}
\{W^1_{367}(w_0^1), ..., W^1_{1095}(w_0^1)\} &= \{Z^1_{367}(w_0^1) + 0.001371742, ..., Z^1_{1095}(w_0^1) + 0.001371742\},\\
\{W^2_{14}(w_0^2), ..., W^2_{64}(w_0^2)\} &= \{Z^2_{14}(w_0^2) - 33.670784, ..., Z^2_{64}(w_0^2) - 33.60784\}, \text{ i}\\
\{W^3_{14}(w_0^3), ..., W^3_{138}(w_0^3)\} &= \{Z^3_{14}(w_0^3) + 0.28, ..., Z^3_{138}(w_0^3) + 0.28\},
\end{aligned}$$

les quals són representades en els següents gràfics:







D'aquesta forma, arribem a un conjunt d'observacions, que provenen d'un procés lineal $\{X_t\}_{t\in\mathbb{Z}}$. És a dir, que $\exists w_0 \in \Omega$, tal que les dades corresponen a $\{X_1(w_0), ..., X_n(w_0)\}$.

3 Processos ARMA

Dins de les sèries lineals, els processos ARMA són els més utilitzats i per tant, en la nostra recerca d'una sèrie lineal $\{X_t\}_{t\in\mathbb{Z}}$, que ajusti les dades observades $\{X_1(w_0), ..., X_n(w_0)\}$, suposarem que es tracta d'un procés ARMA a determinar. En la primera part d'aquest capítol, explicarem les principals propietats dels processos ARMA, i en la segona part, veurem com determinar el procés ARMA que millor ajusti la nostra mostra i a validar-lo, per a poder utilitzar les tècniques de predicció explicades a la primera part.

3.1 Els models ARMA(p, q)

Definició 3.1. *Una sèrie temporal $\{X_t\}_{t\in\mathbb{Z}}$, és un procés ARMA(p, q) si és un procés estacionari i si existeixen enters p i q tals que per a tot enter t,*

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q},$$

on $\{Z_t\}_{t\in\mathbb{Z}} \sim WN(0, \sigma^2)$, i els polinomis a coeficients reals $\phi(z) = (1 - \phi_1 z - ... - \phi_p z^p)$ i $\theta(z) = (1 + \theta_1 z + ... + \theta_q z^q)$ no tenen arrels comunes.

Observació 3.2. $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA(p, q) de mitjana μ, si la sèrie temporal $\{Y_t\}_{t\in\mathbb{Z}}$, on $Y_t = X_t - \mu$, $\forall t \in \mathbb{Z}$, és un procés ARMA(p, q).

En termes de l'operador de retrocés B explicat a la secció 1.3, definint el subespai

$$F =< \{X_t\}_{t\in\mathbb{Z}} \cup \{Z_t\}_{t\in\mathbb{Z}} >\subseteq L^2(\Omega),$$

i els operadors de $\mathbb{L}(F) = \{g : F \longrightarrow F : \text{ lineals}\}$,

$$\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p \text{ i } \theta(B) = (1 + \theta_1 B + ... + \theta_q B^q),$$

on

$$\begin{aligned} B : F &\longrightarrow F \\ X_t &\longmapsto X_{t-1} \\ Z_t &\longmapsto Z_{t-1}, \end{aligned}$$

tenim que el procés ARMA $\{X_t\}_{t\in\mathbb{Z}}$, es pot expressar com a

$$\phi(B)(X_t) = \theta(B)(Z_t), \quad \forall t \in \mathbb{Z}.$$

Anem a caracteritzar com han de ser els polinomis $\phi(z)$ i $\theta(z)$ perquè estigui ben definit el procés ARMA(p, q).

Proposició 3.3. *Siguin $\phi(z) = (1 - \phi_1 z - ... - \phi_p z^p)$ i $\theta(z) = (1 + \theta_1 z + ... + \theta_q z^q)$ dos polinomis a coeficients reals que no tenen arrels comunes, i $\{Z_t\}_{t\in\mathbb{Z}}$ un procés estacionari tal que $\{Z_t\}_{t\in\mathbb{Z}} \sim WN(0, \sigma^2)$. Llavors, existeix un procés estacionari ARMA(p, q), $\{X_t\}_{t\in\mathbb{Z}}$, definit, per a tot enter t, com a*

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q},$$

si i només si el polinomi $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p \neq 0$, $\forall z \in \mathbb{C}$, tal que $|z| = 1$. A més, en aquest cas, la sèrie ARMA(p, q) és única g.p.t.

Demostració. Suposem primer que el polinomi $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p \neq 0$, $\forall z \in \mathbb{C}$, tal que $|z| = 1$. En aquest cas, per continuitat, existeix un $\delta > 0$, i un entorn de la frontera del disc unitat $U = \{z \in \mathbb{C} \text{ tals que } 1 - \delta < |z| < 1 + \delta\} \subseteq \mathbb{C}$, tal que $\phi(z) \neq 0, \forall z \in U$. Per tant,

$$\begin{aligned} \frac{1}{\phi} : \ U &\longrightarrow \mathbb{C} \\ z &\longmapsto \frac{1}{\phi(z)}, \end{aligned}$$

és una funció holomorfa a U (és el quocient de dues funcions holomorfes i el denominador no s'anul·la). Per tant, com a conseqüència de la descomposició de sèries de Laurent [5], existeix una successió de reals $\{\chi_j\}_{j \in \mathbb{Z}}$, tal que

$$\frac{1}{\phi(z)} = \sum_{j=-\infty}^{\infty} \chi_j z^j,$$

per a $1 - \delta < |z| < 1 + \delta$, i a més la sèrie és uniformament i absolutament convergent en compactes de U. En particular, $(z = 1)$,

$$\sum_{j=-\infty}^{\infty} |\chi_j| < \infty.$$

Aquesta darrera condició suposa que, considerant el subespai $G =< Z_t, t \in \mathbb{Z} > \subseteq L^2(\Omega)$, l'operador de retrocés $B \in \mathbb{L}(G)$ i els operadors lineals

$$\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p \text{ i } \theta(B) = (1 + \theta_1 B + ... + \theta_q B^q),$$

podem definir l'operador $\chi(B) : F \longrightarrow F$, definit com a

$$\chi(B) = \frac{1}{\phi(B)} = \sum_{j=-\infty}^{\infty} \chi_j B^j,$$

que està ben definit i és lineal (Proposició 2.7). Basant-nos amb l'observació 2.8, podem definir també l'operador de $\mathbb{L}(G)$:

$$\psi(B) = \chi(B) \circ \theta(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j,$$

el qual compleix (observació 2.8) que,

$$\sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

Per tant, fent ús un altre cop de la proposició 2.7, podem definir, per a tot enter t, el procés estacionari (que a més és lineal):

$$X_t = \psi(B)(Z_t) = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}.$$

[5] Teorema de Descomposició en sèries de Laurent: Sigui $a \in \mathbb{C}$; $0 \leq r \leq R < \infty$, i sigui $U = \{z \in \mathbb{C} : r \leq |z - a| \leq R\}$. Si $f : U \longrightarrow \mathbb{C}$ és una funció holomorfa a U, existeix una única sèrie de Laurent tal que $f(z) = \sum_{n=-\infty}^{\infty} c_n (z - a)^n$, $r < |z - a| < R$, la qual és absolutament convergent en compactes de U.

Observem que es tracta d'un procés ARMA(p,q), ja que composant a ambdós costats per $\phi(B)$, identificant-ho amb el producte de sèries, (ho podem fer sense problema degut a l'observació 2.8), obtenim

$$\phi(B)(X_t) = \theta(B)(Z_t), \ \forall t \in \mathbb{Z}.$$

A més, el procés estacionari $\{X_t\}_{t\in\mathbb{Z}}$, és l'única sol·lució possible g.p.t, ja que si $\{Y_t\}_{t\in\mathbb{Z}}$ és un altre procés ARMA(p,q), utilitzant la notació dels operadors de B al subespai $F =< \{Y_t\}_{t\in\mathbb{Z}} \cup \{Z_t\}_{t\in\mathbb{Z}} >\subseteq L^2(\Omega)$, llavors

$$\phi(B)(Y_t) = \theta(B)(Z_t), \quad \forall t \in \mathbb{Z},$$

i per tant composant ambdós costats per $\chi(B)$, i identificant la composició amb el producte de sèries, (utilitzam de nou l'observació 2.8, ja que $\{Y_t\}_{t\in\mathbb{Z}}$ també és estacionari),

$$Y_t = \psi(B)(Z_t) = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} = X_t.$$

És a dir, per a tot enter t, $Y_t = X_t$ a $L^2(\Omega)$, amb el que cocloem que són iguals g.p.t.

Recíprocament, suposem que $\exists z^* \in \mathbb{C}$ de mòdul 1 tal que $\phi(z^*) = 0$. Sigui $\{X_t\}_{t\in\mathbb{Z}}$ un procés ARMA(p,q), és a dir, es compleix que

$$X_t - \phi_1 X_{t-1} - ... - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + ... + \theta_q Z_{t-q}.$$

Fent ús d'alguns resultats de densitats espectrals, els quals no detallarem per llargària i complexitat, es pot veure que $\{X_t\}_{t\in\mathbb{Z}}$ no és estacionari, el que implica que no és un procés ARMA(p,q) i per tant, arribem a una contradicció i deduïm que en aquest cas no existeix cap procés ARMA(p,q). La idea de la demsotració d'aquesta implicació la podem veure al problema 3.24 del final del capítol 3 de del llibre "Time Series: Theory and Methods", el qual apareix a la bibliografia final del treball.

□

De la demostració anterior deduïm que si $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA(p,q), llavors existeix una successió de reals$\{\psi_j\}_{j\in\mathbb{Z}}$, amb $\sum_{j\in\mathbb{Z}}|\psi_j| < \infty$, tal que, per a tot enter t,

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j}.$$

És a dir, en deduïm el següent corol·lari:

Corołari 3.4. *Si $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA (p,q), llavors és un procés lineal.*

No estudiarem tots els processos ARMA en detall. Demanarem que tinguin propietats "bones", el que traduirem en que siguin causals i invertibles.

Definició 3.5. *Un procés ARMA(p,q) $\{X_t\}_{t\in\mathbb{Z}}$ és causal, si existeix una successió de nombres reals $\{\psi_j\}_{j\geq 0}$ tals que $\sum_{j=0}^{\infty}|\psi_j| < \infty$, i que per a tot enter t*

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

Definició 3.6. *Un procés ARMA(p, q) $\{X_t\}_{t\in\mathbb{Z}}$ és invertible, si existeix una successió de nombres reals $\{\pi_j\}_{j\geq 0}$ tals que $\sum_{j=0}^{\infty}|\pi_j| < \infty$, i que per a tot enter t*

$$Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}.$$

Amb les següents dues proposicions caracteritzarem els processos ARMA causals i invertibles.

Proposició 3.7. *Un procés ARMA(p, q) $\{X_t\}_{t\in\mathbb{Z}}$ és causal, si i només si $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p \neq 0,\ \forall z \in \mathbb{C},$ tals que $|z| \leq 1$.*

Demostració. Suposem primer que $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p \neq 0,\ \forall z \in \mathbb{C},$ tals que $|z| \leq 1$. Utilitzant la mateixa notació que en la demostració de la proposició anterior, en aquest cas $\frac{1}{\phi}$ és una funció holomorfa a un obert U que conté el disc unitat, és a dir,

$$D = \{z \in \mathbb{C} : |z| \leq 1\} \subset U.$$

Per la descomposició en sèries de potències de funcions holomorfes [6], existeix una successió de reals $\{\chi_j\}_{j\geq 0}$, tal que

$$\frac{1}{\phi(z)} = \sum_{j=0}^{\infty} \chi_j z^j,$$

per a $z \in U$. En particular, es absolutament convergent en compactes de D, i per tant $(z = 1)$,

$$\sum_{j=0}^{\infty} |\chi_j| < \infty.$$

Aquesta darrera condició suposa que, considerant el subespai $G =< \{Z_t\}_{t\in\mathbb{Z}} \cup \{X_t\}_{t\in\mathbb{Z}} > \subseteq L^2(\Omega)$, l'operador de retrocés $B \in \mathbb{L}(G)$ i els operadors lineals

$$\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p \text{ i } \theta(B) = (1 + \theta_1 B + ... + \theta_q B^q),$$

podem definir l'operador

$$\chi(B) = \frac{1}{\phi(B)} = \sum_{j=0}^{\infty} \chi_j B^j,$$

que està ben definit i és lineal (Proposició 2.7). Com $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA, per a tot enter t

$$\phi(B)X_t = \theta(B)Z_t,$$

i composant a amdós costats per $\chi(B)$, basant-nos amb l'observació 2.8, obtenim que per a tot enter t

$$X_t = \chi(B) \ o \ \theta(B)(Z_t) = \sum_{j=0}^{\infty} \psi_j Z_{t-j},$$

el qual compleix (observació 2.8) que, $\sum_{j=0}^{\infty}|\psi_j| < \infty$, és a dir, $\{X_t\}_{t\in\mathbb{Z}}$ és un procés causal.

[6] Teorema de Descomposició en sèries de potències: Sigui $U \subseteq \mathbb{C}$, un obert i f una funció complexa holomorfa a U. Si $D(a, R)$ és un disc (obert) centrat en $a \in U$ amb radi $R \geq 0$, tal que l'adherència del disc es troba a U, llavors f és expressada com una sèrie de potències al voltant d'a: $f(z) = \sum_{j=0}^{\infty} c_n(z - a)^n$, per a $z \in D(a, R)$. A més, f és absolutament convergent en compactes de $D(a, R)$.

Recíprocament, suposem que $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA(p,q) causal. En aquest cas, existeix una successió de nombres reals $\{\psi_j\}_{j\geq 0}$ tals que $\sum_{j=0}^{\infty}|\psi_j| < \infty$, i que per a tot enter t

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

Considerant, novament, el subespai $G =< \{Z_t\}_{t\in\mathbb{Z}} \cup \{X_t\}_{t\in\mathbb{Z}} >\subseteq L^2(\Omega)$, l'operador de retrocés $B \in \mathbb{L}(G)$ i els operadors lineals

$$\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p \text{ i } \theta(B) = (1 + \theta_1 B + ... + \theta_q B^q),$$

podem definir l'operador

$$\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j,$$

que està ben definit i és lineal (Proposició 2.7). Per tant, tenim que per a tot enter t,

$$X_t = \psi(B)(Z_t).$$

Composant a ambdós costats per $\phi(B)$, i recordant que $\phi(B)X_t = \theta(B)Z_t$, deduim que per a tot enter t,

$$\theta(B)(Z_t) = \phi(B) \circ \psi(B)(Z_t).$$

Basant-nos de nou amb l'observació 2.8, podem definir l'operador lineal, pertanyent a $\mathbb{L}(G)$,

$$\delta(B) = \phi(B) \circ \psi(B) = \sum_{j=0}^{\infty} \delta_j B^j,$$

on $\sum_{j=0}^{\infty} \delta_j < \infty$. Fixem-nos,que, per a tot enter t,

$$\theta(B)(Z_t) = \delta(B)(Z_t).$$

Per tant, si considerem el suespai $G_2 =< \{Z_t\}_{t\in\mathbb{Z}} >\subseteq G$, resulta que per a tot $W \in G_2$,

$$\theta(B)(W) = \delta(B)(W).$$

Sigui $t \in \mathbb{Z}$. Fixem $k \geq 0$. Llavors, com $\delta(B)(Z_t) = \theta(B)(Z_t)$,

$$< \delta(B)(Z_t), Z_{t-k} >=< \theta(B)(Z_t), Z_{t-k} >,$$

el que implica, ja que $\{Z_t\}_{t\in\mathbb{Z}} \sim WN(0,\sigma^2)$, que

$$\delta_k \sigma^2 = \theta_k \sigma^2.$$

És a dir, per a tot $k \geq 0$, tenim que

$$\delta_k = \theta_k.$$

En termes de sèries de potències (observació 2.8),

$$\theta(z) = \delta(z) = \psi(z)\phi(z).$$

Fixem-nos que el radi de convergència de la sèrie de potències $\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j$ és major o igual que 1, ja que $\sum_{j=0}^{\infty}|\psi_j| < \infty$, el que implica que per a tot $z \in \mathbb{C}$ tal que $|z| \leq 1$,

$$\theta(z) = \delta(z) = \psi(z)\phi(z).$$

Si existís un $z_0 \in \mathbb{C}$, tal que $|z_0| \leq 1$, i $\phi(z_0) = 0$, llavors $\theta(z_0) = \psi(z_0)\phi(z_0) = 0$, el que implicaria que $\phi(z)$ i $\theta(z)$ tenen una arrel comuna, fet que és una contradició ja que $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA(p,q). Per tant, arribem a una contradicció i demostrem, per reducció a l'absurd, que en aquest cas $\phi(z) = 1 - \phi_1 z - ... - \phi_p z^p \neq 0, \ \forall z \in \mathbb{C}, \$ tals que $|z| \leq 1$. □

Proposició 3.8. *Un procés ARMA(p, q)* $\{X_t\}_{t\in\mathbb{Z}}$ *és invertible, si i només si* $\theta(z) = 1 + \theta_1 z + ... + \theta_q z^q \neq 0, \ \forall z \in \mathbb{C}, \$ *tals que* $|z| \leq 1$.

Demostració. La demostració de la primera implicació és anàlega a la primera implicació de la proposició anterior. És a dir, intercanviant els papers de $\phi(z)$ per $\theta(z)$ i amb la mateixa metodologia que abans, suposant que $\theta(z) \neq 0, \ \forall z \in \mathbb{C}$, tals que $|z| \leq 1$, arribariem de forma inmediata a que $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA invertible.

Recíprocament, suposem que $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA(p,q) invertible. En aquest cas, considerant el subespai $G =< \{Z_t\}_{t\in\mathbb{Z}} \cup \{X_t\}_{t\in\mathbb{Z}} >\subseteq L^2(\Omega)$, l'operador de retrocés $B \in \mathbb{L}(G)$ i els operadors lineals

$$\phi(B) = 1 - \phi_1 B - ... - \phi_p B^p \text{ i } \theta(B) = (1 + \theta_1 B + ... + \theta_q B^q),$$

per definició de procés invertible existeix una successió de nombres reals $\{\pi_j\}_{j\geq 0}$ tals que $\sum_{j=0}^{\infty} |\pi_j| < \infty$, i que per a tot enter t

$$\sum_{j=0}^{\infty} \pi_j X_{t-j} = Z_t.$$

Podem definir l'operador de $\mathbb{L}(G)$,

$$\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j,$$

que està ben definit i és lineal (Proposició 2.7). Per tant, tenim que per a tot enter t,

$$\pi(B)(X_t) = Z_t.$$

Composant a ambdós costats per $\phi(B)$, i basant-nos en l'observació 2.8, identificant la composicó d'operadors amb el producte de sèries, per a tot enter t:

$$\phi(B) \circ \pi(B)(X_t) = \phi(B)(Z_t) \longleftrightarrow \pi(B) \circ (\phi(B)(X_t)) = \phi(B)(Z_t).$$

Com $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA (p,q), per a tot enter t:

$$\pi(B) \circ \theta(B)(Z_t) = \phi(B)(Z_t).$$

Basant-nos de nou amb l'observació 2.8, podem definir l'operador lineal, pertanyent a $\mathbb{L}(G)$,

$$\epsilon(B) = \pi(B) \circ \theta(B) = \sum_{j=0}^{\infty} \epsilon_j B^j,$$

on $\sum_{j=0}^{\infty} \epsilon_j < \infty$. Fixem-nos,que, per a tot enter t,

$$\epsilon(B)(Z_t) = \phi(B)(Z_t).$$

Per tant, si considerem el suespai $G_2 =< \{Z_t\}_{t\in\mathbb{Z}} >\subseteq G$, resulta que per a tot $W \in G_2$,

$$\epsilon(B)(W) = \phi(B)(W).$$

Sigui $t \in \mathbb{Z}$. Considerem $k \geq 0$. Llavors, com $\epsilon(B)(Z_t) = \phi(B)(Z_t)$,

$$< \epsilon(B)(Z_t), Z_{t-k} >=< \phi(B)(Z_t), Z_{t-k} >,$$

el que implica, ja que $\{Z_t\}_{t\in\mathbb{Z}} \sim WN(0, \sigma^2)$,

$$\epsilon_k\sigma^2 = \phi_k\sigma^2.$$

És a dir, per a tot $k \geq 0$,

$$\epsilon_k = \phi_k.$$

En termes de sèries de potències (observació 2.8),

$$\phi(z) = \epsilon(z) = \pi(z)\theta(z).$$

Fixem-nos que el radi de convergència de la sèrie de potències $\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j$ és major o igual que 1, ja que $\sum_{j=0}^{\infty}|\pi_j| < \infty$, el que implica que per a tot $z \in \mathbb{C}$, tal que $|z| \leq 1$,

$$\phi(z) = \epsilon(z) = \pi(z)\theta(z).$$

Si existís un $z_0 \in \mathbb{C}$, tal que $|z_0| \leq 1$, i $\theta(z_0) = 0$, llavors $\phi(z_0) = \pi(z_0)\theta(z_0) = 0$, el que implicaria que $\phi(z)$ i $\theta(z)$ tenen una arrel comuna, fet que és una contradició ja que $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA(p, q). Per tant, arribem a una contradicció i demostrem, per reducció a l'absurd, que en aquest cas $\theta(z) = 1+\phi_1 z+...+\theta_p z^p \neq 0, \ \forall z \in \mathbb{C},$ tals que $|z| \leq 1$. □

A partir d'ara ens centrarem amb l'estudi dels processos ARMA(p, q) causals i invertibles. Fixem-nos que, a partir de la proposició 2.7, per a calcular la seva funció d'autocovariància, necessitarem saber els coeficients $\{\psi_j\}_{j\geq 0}$ tals que per a tot enter t,

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

Vegem com calcular-los a partir de $(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q)$.

3.1.1 Càlcul de $\{\psi_j\}_{j\geq 0}$ d'un procés ARMA causal i invertible

Utilitzant les mateixes notacions que en la demostració de la primera implicació de la proposició 3.3, fixem-nos que si $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA(p, q) causal i invertible llavors els coeficients $\{\psi_j\}_{j\geq 0}$, en termes de sèries de potències, venen determinats per la relació $\psi(z) = \chi(z)\theta(z) = \frac{\theta(z)}{\phi(z)}$, o equivalentment

$$\phi(z)\psi(z) = \theta(z),$$

és a dir,

$$(1 - \phi_1 z - ... - \phi_p z^p)(\sum_{j=0}^{\infty} \psi_j z^j) = 1 + \theta_1 z + ... + \theta_q z^q.$$

Amb el que igualant els coeficients de z^j, per a tot $j \geq 0$:

$$\psi_j - \sum_{k=1}^{p} \phi_k \psi_{j-k} = \theta_j \text{ (Suposant que si } k < 0,\ \psi_k = 0, \text{ que } \theta_0 = 1, \text{ i que si } k > q,\ \theta_k = 0).$$

D'aquesta forma, calculant en primer lloc ψ_0, podrem calcular ψ_1, i així, recursivament, tots els coeficients de $\{\psi_j\}_{j\geq 0}$, a partir de $(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q)$, per a expressar la nostra sèrie com a

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}.$$

Un cop calculats $\{\psi_i\}_{i\geq 0}$, podem fer ús de la proposició 2.7 i calcular la funció d'autocovariància d'un procés ARMA(p, q) causal i invertible.

3.1.2 ACVF dels models ARMA(p, q) causals i invertibles

Sigui $\{X_t\}_{t\in\mathbb{Z}}$ un procés ARMA(p, q) causal i invertible. Si volem calcular $\gamma_X(.)$ i $p_X(.)$ fixem-nos que a causa de la proposició 2.7, (ja que els processos ARMA són lineals), per a tot enter h

$$\gamma_X(h) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}.$$

Per tant, coneixent els paràmetres del model $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$ *i* σ^2 podem calcular explícitament γ_X i en conseqüència p_X.

Observació 3.9. Com $\sum_{j=0}^{\infty} |\psi_j| < \infty$, en particular $\lim_{j\to+\infty} \psi_j = 0$, i en conseqüència, $\gamma_X(h) \to 0$ si $h \to \infty$. A la pràctica, la funció d'autocovariància d'un model ARMA qualsevol sol tendir cap a 0 ràpidament.

Els predictors pas a pas, definits a l'anterior secció, en els processos ARMA es poden calcular únicament a partir dels coeficients $(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q)$, fet essencial per al següent apartat determinar quin model ARMA seleccionar. Vegem-ho:

3.1.3 Els predictors pas a pas en els processos ARMA

Sigui $\{X_t\}_{t\in\mathbb{Z}}$ un procés ARMA(p, q) causal i invertible, que ajusta les nostres observacions $\{X_1(w_0), ..., X_n(w_0)\}$, del qual coneixem els valors $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$ i σ^2.

Lògicamenent, podem realitzar prediccions exactament de la mateixa forma que en 2.1 i fer ús de les seves propietats, ja que un procés ARMA és un procés estacionari amb mitjana 0. La característica essencial que tenen aquests models és que els predictors pas a pas no depenen de σ^2, fet que ens serà de molta utilitat en la següent secció a l'hora de triar el model ARMA que millor ajusti la nostra mostra. Definim els subespais, per a $1 \leq m \leq n$,

$$F_m =< 1, X_1, ..., X_m > \subseteq L^2(\Omega),$$
$$F_0 =< 1 > \subseteq L^2(\Omega).$$

Podem veure que cap de les prediccions $X_{F_m} P_{m+h},\ \forall m \in \{0, 1, ..., n\},\ \forall h \geq 0$, depenen de σ^2. Anem a comprovar-ho:

Per tot el que hem vist fins ara,

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j},$$

on els coeficients $\{\psi_j\}_{j\geq 0}$ els podem calcular a partir de $(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q)$, i a més, per a tot enter h

$$\gamma_X(h) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h},$$

amb el que deduïm que $\frac{\gamma_X}{\sigma^2}$ es pot expressar en funció de $(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q)$. Sigui $m \in \{0, 1, ..., n\}$ i $h \geq 0$, fent ús d'allò explicat a la secció 2.1:

$$P_{F_m} X_{m+h} = \sum_{i=1}^{m} a_i X_{m+1-i},$$

on $a_m^* = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$ satisfà l'equació $\Gamma_m a_m^* = \gamma_m(h)$. Fixem-nos que

$$\Gamma_m, \text{ és la matriu real definida per } (\gamma_X(i-j))_{1\leq i,j\leq m}, \text{ i } \gamma_m(h) = \begin{pmatrix} \gamma_X(h) \\ \gamma_X(h+1) \\ \vdots \\ \gamma_X(h+m-1) \end{pmatrix}.$$

A causa del raonament del paràgraf anterior, $\frac{\Gamma_m}{\sigma^2}$ i $\frac{\gamma_m(h)}{\sigma^2}$ depenen únicament de $(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q)$. Ara bé, a_m^* satisfà

$$\Gamma_m a_m^* = \gamma_m(h) \longleftrightarrow \frac{\Gamma_m}{\sigma^2} a_m^* = \frac{\gamma_m(h)}{\sigma^2},$$

amb el que concloem que, $a_m^* = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix}$, es pot calcular en funció de $(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q)$, i per tant $P_{F_m} X_{m+h}$ també, i és independent de σ^2. Recordem que $a_0 = 0$, ja que $\mu_X = 0$. Conseqüentment, el valor numèric $P_{F_m} X_{m+h}(w_0)$ també és independent de σ^2.

En particular, per a $m \in \{0, 1, ..., n-1\}$, els predictors lineals $\hat{X_{m+1}} = P_{F_m} X_{m+1}$, no depenen de σ^2. A més, fixem-nos que de 2.1:

$$v_m = E((X_{m+1} - P_m X_{m+1})^2) = \gamma_X(0) - a_m^T \gamma_m(1) = \sigma^2 (\frac{\gamma_X(0)}{\sigma^2} - a_m^T \frac{\gamma_m(1)}{\sigma^2}) = \sigma^2 r_m,$$

on r_m és independent de σ^2 i es pot expressar úniament a partir de $(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q)$.

En aquest capítol ens hem dedicat a veure les propietats pincipals dels processos ARMA causals i invertibles. A partir d'ara, el nostre objectiu serà determinar el procés ARMA causal i invertible que millor ajusti la nostra mostra $\{X_1(w_0), ..., X_n(w_0)\}$. És a dir, volem determinar els enters (p, q) i el valor dels paràmetres $(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q, \sigma^2)$ d'un procés ARMA causal i invertible $\{X_t\}_{t\in\mathbb{Z}}$, que, per a $w_0 \in \Omega$, sigui justificable que $\{X_1(w_0), ..., X_n(w_0)\}$ sigui la nostra mostra.

3.2 La distribució Gaussiana en els processos ARMA

Recordem, de l'inici del capítol 3, que nosaltres partim d'una mostra $\{X_1(w_0), ..., X_n(w_0)\}$, la qual suposem que prové d'un procés ARMA causal i invertible, amb enters p, q i paràmetres $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$ i σ^2 a priori desconeguts. En aquesta secció suposarem que sabem quins són els valors de p i q, i el nostre objectiu serà estimar els $p+q+1$ paràmetres desconeguts, $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$ *i* σ^2, que determinin el procés ARMA que millor ajusti les nostres n dades.

La idea és simple, suposem que el procés ARMA(p, q) causal i invertible $\{X_t\}_{t\in\mathbb{Z}}$, amb els $p+q+1$ paràmetres desconeguts, que representa la nostra mostra, és un procés Gaussià, és a dir, que per a tot enter t

$$X_t \sim N(0, \gamma_X(0)).$$

Fixem-nos que en aquest cas, la funció γ_X determina el model complet de la sèrie temporal, és a dir totes les distribucions conjuntes en dimensió finita, ja que $\mu_X = 0$, i γ_X determina $Cov(X_r, X_s)$, per a tots enters r i s. A més, pel que hem vist en la secció anterior, γ_X, i per tant totes les propietats de segon ordre de la sèrie, es poden expressar en funció dels paràmetres $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$ i σ^2.

Cal limitar l'espai de paràmetres de $(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q)$ a la regió tal que els polinomis, $\phi(z)$ i $\theta(z)$, ens defineixin un procés ARMA causal i invertible. Definim

$$\begin{aligned} U = \Big\{ &(\phi_1, ..., \phi_p, \theta_1, ..., \theta_q, \sigma^2) \in \mathbb{R}^{p+q+1} : \sigma^2 > 0,\ \phi(z) = 1 - ... - \phi_p z^p \text{ i} \\ &\theta(z) = 1 + ... + \theta_q z^q \text{ no tenen arrels comunes, i per a tot } z \in \mathbb{C},\ \text{tal que } |z| \leq 1, \\ &\phi(z) \neq 0 \text{ i } \theta(z) \neq 0 \Big\}. \end{aligned}$$

Llavors, per a cada $\lambda = (\phi_1, ..., \phi_p, \theta_1, ..., \theta_q, \sigma^2) \in U$, queda definit un procés ARMA (p, q) causal i invertible $\{X_t\}_{t\in\mathbb{Z}}$, del qual podem calcular γ_X, i en conseqüència totes les distribucions conjuntes en dimensió finita.

Facem inferència estadística sobre el vector aleatori $(X_1, ..., X_n)$ (de la que prové la nostra mostra). En particular, per a cada $\lambda \in U$, la terna

$$(\mathbb{R}^n, \mathbb{B}(\mathbb{R}^n), \mathbb{P}_\lambda),$$

és un espai de probabilitat (que segueix la distribució gaussiana del vector $(X_1, ..., X_n)$ derivada del procés ARMA(p, q) definit a partir del valor de λ).

És a dir, $(\mathbb{R}^n, \mathbb{B}(\mathbb{R}^n), \mathbb{P})$, on $\mathbb{P} = \{\mathbb{P}_\lambda : \lambda \in U\}$, és un model estadístic paramètric, amb U com a espai de paràmetres (dimensió $p+q+1$). Aleshores, com per a cada valor de λ el vector $(X_1, ..., X_n)$ és gaussià, podem definir la seva funció de versemblança.

$$\begin{aligned} L : \mathbb{R}^n \times U &\longrightarrow \mathbb{R} \\ (y_1, ..., y_n, \phi_1, ..., \phi_p, \theta_1, ..., \theta_q, \sigma^2) &\longmapsto (2\pi)^{-\frac{n}{2}} (det\ \Gamma_n)^{-\frac{1}{2}} exp(-\frac{1}{2} {y_n^*}^T \Gamma_n^{-1} y_n^*) \end{aligned}$$

on

$$\Gamma_n \text{ és la matriu real definida per } (\gamma_X(i-j))_{1\leq i,j\leq n},$$

la qual podem expressar en funció de $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$ *i* σ^2, i suposem que és invertible. Per altra banda,

$$y_n^* = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n \text{ és una mostra qualsevol de } (X_1, ..., X_n).$$

Per a cada $\lambda \in U$, considerem $\gamma_X = f(\lambda)$, la funció d'autocovariància provenent del corresponent procés ARMA (p, q). Amb les mateixes notacions que en 2.1.1, considerem el vector aleatori $X_n^* = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$, els predictors pas a pas $\hat{X}_n^* = \begin{pmatrix} \hat{X_1} \\ \vdots \\ \hat{X_n} \end{pmatrix}$, els errors quadràtics $v_i = E((X_{i+1} - \hat{X_{i+1}})^2)$ per a $i \in \{0, 1, ..., n-1\}$, els errors de predicció $U_n^* = \begin{pmatrix} X_1 - \hat{X_1} \\ \vdots \\ X_n - \hat{X_n} \end{pmatrix}$, i les matrius reals A_n, C_n i Θ_n. A causa de 3.1.3, els coeficients de A_n són independents de σ^2, i com a conseqüència, C_n i Θ_n també, i tots ells es poden calcular en funció de $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$. A més, si coneixem $(X_1(w_1), ..., X_n(w_1)) = (y_1, ..., y_n)$ (on $w_1 \in \Omega$), el vector de predictors, (avaluats a $w_1 \in \Omega$), també depenen únicament de $\phi_1, ..., \phi_p, \theta_1, ..., \theta_q$.

Per l'observació 2.4, les $n + 1$ compoents del vector aleatori U_n^* estan intercorrelacionades, i per tant la seva matriu de covariàncies (D_n) és diagonal i:

$$D_n = diag(v_0, ..., v_{n-1}).$$

Com U_n^* és una transformació afí del vetor aleatori Gaussià X_n^*, ja que $U_n^* = A_n X_n^*$, llavors U_n^* també és Gaussià. Per tant, tenim que:

$$X_n^* \sim N(0, \Gamma_n) \text{ i } U_n^* \sim N(0, D_n).$$

Per altra banda, X_n^* també és una transformació afí de U_n^*, ja que $X_n^* = C_n U_n^*$, amb el que deduïm que:

$$\Gamma_n = C_n D_n C_n^T.$$

Per tant, recordant de 2.1.1 que $det(C_n) = det(A_n) = 1$,

$$det(\Gamma_n) = det(C_n)^2 v_0 ... v_{n-1} = v_0 ... v_{n-1},$$

i fixem-nos que,

$$\begin{aligned} {X_n^*}^T \Gamma_n^{-1} X_n^* &= {X_n^*}^T (C_n^T)^{-1} D_n^{-1} C_n^{-1} X_n^* = {X_n^*}^T A_n^T D_n^{-1} A_n X_n^* = {U_n^*}^T D_n^{-1} U_n^* \\ &= \sum_{j=1}^{n} \frac{(X_j - \hat{X}_j)^2}{v_{j-1}}. \end{aligned}$$

Concretament, si $(y_1, ..., y_n) = (X_1(w_1), ..., X_n(w_1)) \in \mathbb{R}^n$, és una mostra qualsevol del vector aleatori X_n^*, llavors

$$({X_n^*}^T \Gamma_n^{-1} X_n^*)(w_1) = {y_n^*}^T \Gamma_n^{-1} y_n^* = \sum_{j=1}^{n} \frac{(X_j - \hat{X}_j)^2}{v_{j-1}}(w_1).$$

Definint $\hat{y_j} = \hat{X_j}(w_1) = \hat{X_j}(y_1, ..., y_n)$, independent de σ^2 per 3.1.3, aquesta darrera avaluació es pot expressar com a

$$\sum_{j=1}^{n} \frac{(y_j - \hat{y_j})^2}{v_{j-1}}.$$

Com de 3.1.3 sabem que per a $i \in \{0,1,...,n-1\}$, $v_i = \sigma^2 r_i$, amb r_i independent de σ^2, llavors

$$(2\pi)^{-\frac{n}{2}}(v_0...v_{n-1})^{-\frac{1}{2}}exp(-\frac{1}{2}\sum_{j=1}^{n}\frac{(y_j-\hat{y_j})^2}{v_{j-1}})$$
$$=(2\pi\sigma^2)^{-\frac{n}{2}}(r_0...r_{n-1})^{-\frac{1}{2}}exp(-\frac{1}{2\sigma^2}\sum_{j=1}^{n}\frac{(y_j-\hat{y_j})^2}{r_{j-1}}).$$

I per tant,

$$L(y_1,...,y_n,\phi_1,...,\phi_p,\theta_1,...,\theta_q,\sigma^2)=(2\pi\sigma^2)^{-\frac{n}{2}}(r_0...r_{n-1})^{-\frac{1}{2}}exp\left(-\frac{1}{2\sigma^2}\sum_{j=1}^{n}\frac{(y_j-\hat{y_j})^2}{r_{j-1}}\right),$$

on per a $j \in \{1,...,n\}$, r_{j-1} depen únicament de $(\phi_1,...,\phi_p,\theta_1,...,\theta_q)$, i $\hat{y_j}$ depen de $(\phi_1,...,\phi_p,\theta_1,...,\theta_q)$ i $(y_1,...,y_n)$. És a dir, són independents de σ^2.

Un cop tenim definida la funció de versemblança del nostre model estadístic, l'objectiu és estimar els paràmetres $\phi_1,...,\phi_p,\theta_1,...,\theta_q$ i σ^2. El mètode que utilitzarem per durho a terme serà el de màxima versemblança, ja que els estimadors trobats per aquesta metodologia són assimptòticament normals i no esbiaixats.

3.2.1 Estimadors per màxima versemblança

Recordem que nosaltres partim d'una mostra que prové d'un procés ARMA(p,q) $\{X_t\}_{t\in\mathbb{Z}}$, amb enters p,q coneguts i paràmetres $(\phi_1,...,\phi_p,\theta_1,...,\theta_q,\sigma^2)$ desconguts. És a dir, que les observacions corresponen a $\{X_1(w_0),...,X_n(w_0)\}$ per a un cert $w_0 \in \Omega$. La idea és trobar un estadístic $\hat{\lambda}(X_1,...,X_n)=(\hat{\phi_1},...,\hat{\phi_p},\hat{\theta_1},...,\hat{\theta_q},\hat{\sigma^2})(X_1,...,X_n)$ que maximitzi la funció de versemblança

$$L(y_1,...,y_n,\hat{\lambda}(y_1,...,y_n)).$$

per a tota mostra $(y_1,...,y_n)\in\mathbb{R}^n$. A la pràctica, trobar els estimadors explícitament en funció de les variables aleatòries $(X_1,...,X_n)$ és inviable, motiu pel qual el que farem és estimar-los numèricament, és a dir, trobar el valor $\hat{\lambda}$ que maximitzi

$$L(X_1(w_0),...,X_n(w_0),\hat{\lambda}(X_1(w_0),...,X_n(w_0))),$$

on $\{X_1(w_0),...,X_n(w_0)\}$ són les dades observades, és a dir, la única mostra que posseïm. Amb aquesta metodologia, es troba $\hat{\lambda}(X_1(w_0),...,X_n(w_0))$, és a dir, l'estimador de màxima versemblança avaluat únicament a la nostra mostra.

Estimació de $\hat{\sigma^2}$:

Suposem que $(y_1,...,y_n)\in\mathbb{R}^n$ i que $(\hat{\phi_1},...,\hat{\phi_p},\hat{\theta_1},...,\hat{\theta_q})$ són coneguts, és a dir, són constants. En aquest cas, per a $j\in\{1,...,n\}$, tant $\hat{y_j}$ com r_{j-1} són també coneguts i definint

$$S(\hat{\phi},\hat{\theta})=\sum_{j=1}^{n}\frac{(y_j-\hat{y_j})^2}{r_{j-1}},$$

la funció de versemblança és unidimensional, i queda reduïda a

$$l:(0,\infty)\longrightarrow\mathbb{R}$$
$$\sigma^2\longmapsto(2\pi\sigma^2)^{-\frac{n}{2}}(r_0...r_{n-1})^{-\frac{1}{2}}exp(-\frac{1}{2\sigma^2}S(\hat{\phi},\hat{\theta}))$$

Fixem-nos que $l(\sigma^2)$, és resultat d'operacions elementals de funcions de classe $C^\infty(0,\infty)$, i per tant també és de classe $C^\infty(0,\infty)$. Si derivem un cop:

$$\frac{dl}{d\sigma^2}(\sigma^2) = -\frac{1}{2\sqrt{2\pi r_0..r_{n-1}}}\left((\sigma^2)^{\frac{-n-4}{4}} exp(-\frac{1}{2\sigma^2}S(\hat{\phi},\hat{\theta}))\right)\left(n\sigma^2 - S(\hat{\phi},\hat{\theta})\right),$$

tenim que $\frac{dl}{d\sigma^2}$ només s'anul·la a $\sigma^2 = n^{-1}S(\hat{\phi},\hat{\theta})$. En particular, tenim que:

$$\frac{dl}{d\sigma^2}(\sigma^2) = \begin{cases} > 0, & \text{si } \sigma^2 \in (0, n^{-1}S(\hat{\phi},\hat{\theta})), \\ = 0, & \text{si } \sigma^2 = n^{-1}S(\hat{\phi},\hat{\theta}), \\ < 0, & \text{si } \sigma^2 > n^{-1}S(\hat{\phi},\hat{\theta}). \end{cases}$$

Amb el que deduïm que $\sigma^2 = n^{-1}S(\hat{\phi},\hat{\theta})$ és un màxim global a $(0,\infty)$. En conclusió, fixats $(y_1,...,y_n)$ i $(\hat{\phi_1},...,\hat{\phi_p},\hat{\theta_1},...,\hat{\theta_q})$, l'estimador que maximitza la funció de versemblança és

$$\hat{\sigma^2} = n^{-1}S(\hat{\phi},\hat{\theta}).$$

Estimació de $(\hat{\phi_1},...,\hat{\phi_p},\hat{\theta_1},...,\hat{\theta_q})$:

Si per a $(y_1,...,y_n) \in \mathbb{R}^n$ fixat, considerem

$$S(\hat{\phi},\hat{\theta}) = \sum_{j=1}^{n} \frac{(y_j - \hat{y_j})^2}{r_{j-1}} \quad i \quad \hat{\sigma^2} = n^{-1}S(\hat{\phi},\hat{\theta}),$$

llavors, gràcies a l'estimació de $\hat{\sigma^2}$ acabada de justificar, el nostre objectiu és trobar $(\hat{\phi},\hat{\theta}) = (\hat{\phi_1},...,\hat{\phi_p},\hat{\theta_1},...,\hat{\theta_q}) \in U$ per a maximitzar la funció

$$L(\hat{\phi},\hat{\theta}) = exp(-\frac{n}{2})\frac{1}{\sqrt{(2\pi)^n}}\frac{1}{\sqrt{(n^{-1}S(\hat{\phi},\hat{\theta}))^n r_0...r_{n-1}}},$$

el que equival a trobar $(\hat{\phi},\hat{\theta})$ per a minimitzar la funció

$$(n^{-1}S(\hat{\phi},\hat{\theta}))^n r_0...r_{n-1},$$

que coincideix en minimitzar

$$(n^{-1}S(\hat{\phi},\hat{\theta}))(r_0...r_{n-1})^{\frac{1}{n}},$$

i prenent logaritmes, en determinar $(\hat{\phi},\hat{\theta})$ per a trobar el mínim de

$$l(\phi,\theta) = log(n^{-1}S(\phi,\theta)) + n^{-1}\sum_{j=1}^{n} log(r_{j-1}).$$

En conclusió, per a calcular els estimadors de màxima versemblança aplicats a la nostra mostra $\{X_1(w_0),...,X_n(w_0)\}$:

1. Trobar els valors $(\hat{\phi},\hat{\theta}) \in U$ que minimitzin la funció

$$l(\hat{\phi},\hat{\theta}) = log(n^{-1}S(\hat{\phi},\hat{\theta})) + n^{-1}\sum_{j=1}^{n} log(r_{j-1}).$$

Aquest procés es du a terme de forma numèrica i actualment existeixen una gran quantitat de programes que a partir d'uns valors inicials (ϕ_0, θ_0) busquen de forma sistemàtica i troben els valors $(\hat{\phi}, \hat{\theta})$ de U que minimitzen la funció l. Fixem-nos que per a cada valor de (ϕ, θ) que el programa provi, juntament amb la mostra $\{X_1(w_0), ..., X_n(w_0)\}$, es calculen els predictors lineals $\{\hat{X}_1(w_0), ..., \hat{X}_n(w_0)\}$ i els coeficients $\{r_0, ..., r_{n-1}\}$, i es pot avaluar sense problema el valor de $l(\phi, \theta)$.

2. Un cop trobats $(\hat{\phi}, \hat{\theta})$, i juntament amb la nostra mostra $\{X_1(w_0), ..., X_n(w_0)\}$ s'estima $\hat{\sigma^2}$:
$$\hat{\sigma^2} = n^{-1} S(\hat{\phi}, \hat{\theta})$$
I d'aquesta forma es tenen els estimadors de màxima versemlança $(\hat{\phi}, \hat{\theta}, \hat{\sigma^2})$ aplicats a les nostres dades observades.

Exemple 3.10.

- Respecte la temperatura de Madrid, recordem que diposem de la mostra $\{W^1_{367}(w^1_0), ..., W^1_{1095}(w^1_0)\}$, provenent del procés ARMA $\{W^1_t\}_{t\in\mathbb{Z}}$, amb paràmetres desconeguts. Considerant que $p = 5$ i $q = 4$, (ho argumentarem en la següent secció), és a dir, que per a tot enter t,
$$W^1_t - \phi^1_1 W^1_{t-1} - ... - \phi^1_5 W^1_{t-5} = V^1_t + \theta^1_1 V^1_{t-1} + ... + \theta^1_3 V^1_{t-3} + \theta^1_4 V^1_{t-4},$$
on $\{V^1_t : \Omega_1 \longrightarrow \mathbb{R}\}_{t\in\mathbb{Z}} \sim WN(0, \sigma^2_1)$, el nostre objectiu és estimar els paràmetres
$$(\hat{\phi^1_1}, ..., \hat{\phi^1_5}, \hat{\theta^1_1}, ..., \hat{\theta^1_4}, \hat{\sigma^2_1}).$$
Aplicant la teoria anterior, suposant que el model és gaussià, calculem els estimadors de màxima versemblança mitjançant el programa R, obtenint els següents resultats:
$$\begin{aligned}(\hat{\phi^1_1},\ \hat{\phi^1_2},\ \hat{\phi^1_3},\ \hat{\phi^1_4},\ \hat{\phi^1_5}) &= (1.11067,\ -0.17912,\ -0.976662,\ 0.770369, -0.067951),\\ (\hat{\theta^1_1},\ \hat{\theta^1_2},\ \hat{\theta^1_3},\ \hat{\theta^1_4}) &= (-1.29311,\ 0.152339,\ 1.03649,\ -0.895701),\\ \hat{\sigma^2_1} &= 7.13248.\end{aligned}$$

- Respecte les vendes de l'empresa, recordem que diposem de la mostra $\{W^2_{14}(w^2_0), ..., W^2_{64}(w^2_0)\}$, provenent del procés ARMA $\{W^2_t\}_{t\in\mathbb{Z}}$, amb paràmetres desconeguts. Considerant que $p = 0$ i $q = 1$, (ho argumentarem en la següent secció), és a dir, que per a tot enter t,
$$W^2_t = V^2_t + \theta^2_1 V^2_{t-1},$$
on $\{V^2_t : \Omega_2 \longrightarrow \mathbb{R}\}_{t\in\mathbb{Z}} \sim WN(0, \sigma^2_2)$, el nostre objectiu és estimar els dos paràmetres
$$(\hat{\theta^2_1}, \hat{\sigma^2_2}).$$
Aplicant de nou la teoria anterior, suposant que el model és gaussià, mitjançant l'R obtenim els estimadors de màxima versemblança següents:
$$\begin{aligned}\hat{\theta^2_1} &= -0.853365,\\ \hat{\sigma^2_2} &= 278353.\end{aligned}$$

- Respecte els trasllats a la UCI, recordem que diposem de la mostra $\{W^3_{14}(w^3_0), ...,$ $W^3_{138}(w^3_0)\}$, provenent del procés ARMA $\{W^3_t\}_{t\in\mathbb{Z}}$, amb paràmetres desconeguts. Considerant que $p = 4$ i $q = 4$, (ho argumentarem en la següent secció), és a dir, que per a tot enter t,

$$W^3_t + \phi^3_1 W^3_{t-1} + ... + \phi^3_4 W^3_{t-4} = V^3_t + \theta^3_1 V^3_{t-1} + ... + \theta^3_4 V^3_{t-4},$$

  on $\{V^3_t : \Omega_3 \longrightarrow \mathbb{R}\}_{t\in\mathbb{Z}} \sim WN(0, \sigma^2_3)$, el nostre objectiu és estimar els paràmetres

$$(\hat{\phi^3_1}, ..., \hat{\phi^3_4}, \hat{\theta^3_1}, ..., \hat{\theta^3_4}, \hat{\sigma^2_3}).$$

  Aplicant la teoria anterior, suposant que el model és gaussià, els estimadors de màxima versemblança calculats per l'R són els següents:

$$\begin{aligned}(\hat{\phi^3_1},\ \hat{\phi^3_2},\ \hat{\phi^3_3},\ \hat{\phi^3_4}) &= (0.491451,\ 0.0771776\ -0.542833,\ 0.536081),\\ (\hat{\theta^3_1},\ \hat{\theta^3_2},\ \hat{\theta^3_3}, \hat{\theta^3_4}) &= (-0.961459, -0.00000221896, 0.961454, -0.999992),\\ \hat{\sigma^2_3} &= 1601.33.\end{aligned}$$

3.2.2 Elecció de l'ordre p i q

En la tria del model ARMA que representi la nostres dades hem donat com a sabuts els ordres p i q. Ara bé, com triar aquests nombres? Ens basarem amb el criteri de l'AIC com a bondat de l'ajust:

Recordatori de l'AIC d'un model

El criteri d'informació d'Akaike (AIC) és una mesura de qualitat relativa d'un model estadístic paramètric.

Si considerem el model estadístic paramètric sobre el vector aleatori $(X_1, ..., X_n)$, és a dir,

$$(\mathbb{R}^n, \mathbb{B}(\mathbb{R}^n), \{\mathbb{P}_\lambda : \lambda \in U \subseteq \mathbb{R}^m\}),$$

tenim una mostra $\{X_1(w_0), ..., X_n(w_0)\}$ per a un cert $w_0 \in \Omega$, i considerem la funció de versemblança

$$\begin{aligned}L : \mathbb{R}^n \times U &\longrightarrow \mathbb{R}\\ (y, \lambda) &\longmapsto L(y, \lambda),\end{aligned}$$

llavors l'AIC es definex com a

$$AIC = 2m - 2log(L^*)$$

on m és el nombre de paràmetres del model estadístic (dimensió de U) i L^* representa la funció de versemblança màxima del model avaluada a $(X_1(w_0), ..., X_n(w_0))$. És a dir,

$$L^* = L(X_1(w_0), ..., X_n(w_0), \hat{\lambda}),$$

on $\hat{\lambda}$ és l'estimador de λ de màxima versemblança avaluat a les nostres dades observades. Fixem-nos que L^* ens indica la "versemblança", que la mostra $(X_1(w_0), ..., X_n(w_0))$ provengui realment del model estadístic considerat, de forma que quan més alt és aquest valor, millor és la bondat de l'ajust del model considerat. Ara bé, quan en un model esadístic afegim més paràmetres, la bondat d'aquest (L^*) augmenta. La idea de definir l'AIC així, és la de, amb el menor nombre de paràmetres, trobar el millor model possible. S'acostuma a seleccionar models estadístics amb menor AIC (equilibri entre la bondat i el nombre de paràmetres).

En el nostre cas, $\hat{\lambda} = (\hat{\phi}, \hat{\theta}, \hat{\sigma^2})$ i $m = p + q + 1$.

El procés que realitzarem nosaltres és el següent:

1. Determinar uns rangs per a p i q (no massa amplis, ja que no volem sobrecarregar de paràmetres el nostre model).
2. Plantejar-nos tots els models estadístics paramètrics possibles amb les parelles (p, q).
3. Trobar els estimadors de màxima versemblança per a cadascun dels models (amb la mostra $\{X_1(w_0), ..., X_n(w_0)\}$) i calcular l'AIC de cadascun d'ells.
4. Seleccionar el que tengui un menor AIC, i per tant el que ajusta millor les nostres dades segons el criteri d'Akaike.

Exemple 3.11. Aquest ha estat el criteri seleccionat per a la tria dels ordres p i q en l'exemple 3.10. El procediment ha estat el següent:

1. Determinar els rang per a p i q. Hem considerat $0 \leq p, q \leq 5$.
2. Plantejar-nos tots els models estadístics paramètrics possibles amb les parelles (p, q).
3. Trobar els estimadors de màxima versemblança per a cadascun dels models (amb la mostra de cadascun dels 3 exemples) i calcular l'AIC de cadascun d'ells.
4. Seleccionar el que tengui un menor AIC, i per tant el que ajusta millor les nostres dades segons el criteri d'Akaike.

Amb aquesta metodologia, hem arribat a les següents conclusions:

- Respecte la temperatura, el model ARMA amb $p = 5$ i $q = 4$ ha estat el que menor AIC ha obtingut, amb un valor de 3526.46.
- Respecte les vendes, el model ARMA amb $p = 0$ i $q = 1$ ha estat el que menor AIC ha obtingut, amb un valor de 906.9641.
- Respecte els trasllats, el model ARMA amb $p = 4$ i $q = 4$ ha estat el que menor AIC ha obtingut, amb un valor de 1305.257.

3.2.3 Diagnòstics

Arribats a aquest punt, ja sabem com seleccionar el model ARMA(p, q) causal i invertible, que millor representi les nostres observacions. Nosaltres suposarem que els paràmetres estimats per màxima versemblança $(\hat{\phi_1}, ..., \hat{\phi_p}, \hat{\theta_1}, ..., \hat{\theta_q}, \hat{\sigma^2})$ són els paràmetres reals que defineixen el procés ARMA del qual provenen les observacions. (Tot i que en teoria són només una aproximació dels paràmetres reals, aquestes estimacions per màxima versemblança són molt precissos i no hi ha inconvenients greus en realitzar aquesta suposició). És a dir, després d'aquesta suposició, arribem a la conclusió que $\exists w_0 \in \Omega$, tal que

$\{X_1(w_0), ..., X_n(w_0)\}$ són les nostres observacions, on $\{X_t\}_{t\in\mathbb{Z}}$ és un procés ARMA(p,q) (amb p,q i tots els paràmetres coneguts). Concretament, per a tot enter t,

$$X_t - \hat{\phi}_1 X_{t-1} - ... - \hat{\phi}_p X_{t-p} = Z_t + \hat{\theta}_1 Z_{t-1} + ... + \hat{\theta}_q Z_{t-q}$$

on $\{Z_t\}_{t\in\mathbb{Z}} \sim WN(0, \hat{\sigma^2})$.

A més, recordem que, per la suposició de l'anterior secció, $\{X_t\}_{t\in\mathbb{Z}}$ és un procés Gaussià.

Exemple 3.12. Lligant amb l'exemple 3.10, 3.11 i amb les consideracions que acabem d'explicar, arribem a la conclusió que:

- Respecte la temperatura de Madrid, la nostra mostra $\{W^1_{367}(w^1_0), ..., W^1_{1095}(w^1_0)\}$, prové del procés ARMA $\{W^1_t\}_{t\in\mathbb{Z}}$, definit per a tot enter t com a,

  $$W^1_t - \hat{\phi}^1_1 W^1_{t-1} - \hat{\phi}^1_2 W^1_{t-2} - \hat{\phi}^1_3 W^1_{t-3} - \hat{\phi}^1_4 W^1_{t-4} - \hat{\phi}^1_5 W^1_{t-5} = V^1_t + \hat{\theta}^1_1 V^1_{t-1} + \hat{\theta}^1_2 V^1_{t-2} + \hat{\theta}^1_3 V^1_{t-3} + \theta^1_4 \hat{V}^1_{t-4},$$

  on $\{V^1_t\}_{t\in\mathbb{Z}} \sim WN(0, \hat{\sigma^2_1})$.

- Respecte les vendes de l'empresa, la nostra mostra $\{W^2_{14}(w^2_0), ..., W^2_{64}(w^2_0)\}$, prové del procés ARMA $\{W^2_t\}_{t\in\mathbb{Z}}$, definit per a tot enter t com a,

  $$W^2_t = V^2_t + \hat{\theta}^2_1 V^2_{t-1},$$

  on $\{V^2_t\}_{t\in\mathbb{Z}} \sim WN(0, \hat{\sigma^2_2})$.

- Respecte els trasllats a la UCI, la nostra mostra $\{W^3_{14}(w^3_0), ..., W^3_{138}(w^3_0)\}$, prové del procés ARMA $\{W^3_t\}_{t\in\mathbb{Z}}$, definit per a tot enter t com a,

  $$W^3_t - \hat{\phi}^3_1 W^3_{t-1} - \hat{\phi}^3_2 W^3_{t-2} - \hat{\phi}^3_3 W^3_{t-3} - \hat{\phi}^3_4 W^3_{t-4} = V^3_t + \hat{\theta}^3_1 V^3_{t-1} + \hat{\theta}^3_2 V^3_{t-2} + \hat{\theta}^3_3 V^3_{t-3} + \hat{\theta}^3_4 V^3_{t-4},$$

  on $\{V^3_t\}_{t\in\mathbb{Z}} \sim WN(0, \hat{\sigma^2_3})$.

Ara bé, és certa aquesta conclusió a la qual hem arribat? Podem assegurar que aquest model ARMA, amb aquests paràmetres estimats, ajusta les nostres dades? Fixem-nos, que al llarg de tot el treball hem hagut de suposar com a certs molts de fets, com que la sèrie estacionària a determinar és un procés lineal, o que aquest darrer és, en particular, un procés ARMA, o que els paràmetres estimats per màxima versemblança són els paràmetres reals. Podria donar-se el fet que les dades no poguéssin representar-se mitjançant un procés ARMA.

L'eina que tenim per a validar i assegurar que efectivament les nostres observacions provenen del procés ARMA(p,q) $\{X_t\}_{t\in\mathbb{Z}}$, que hem trobat a partir dels paràmetres estimats, és l'anàlisi dels residus.

Anàlisi dels residus

Sota la hipòtesis de que el procés ARMA $\{X_t\}_{t\in\mathbb{Z}}$ (amb els paràmetres estimats $(\hat{\phi}, \hat{\theta}, \hat{\sigma^2})$) ajusta les nostres n dades, amb la mateixa notació que a 2.1.1, tenim que :

$$U_n^* = A_n X_n^*,$$

i com X_n^* és Gaussià, llavors U_n^* també ho és (és una transformació afí). A més, sabem que té mitjana 0, i que la seva matriu de covariànces és la matriu diagonal:

$$D_n = diag(v_0, ..., v_{n-1}) = \hat{\sigma^2} diag(r_0, ..., r_{n-1}).$$

Per tant, per a tot $1 \leq i \leq n$, les variables aleatòries:

$$\frac{1}{\sqrt{r_{i-1}}} U_i = \frac{X_i - \hat{X}_i}{\sqrt{r_{i-1}}} \sim N(0, \hat{\sigma^2}),$$

i a més són independents (ja que són intercorrelacionades i són normals).

Per tant, si el model ajusta les nostres dades, és a dir, si $\exists w_0 \in \Omega$, tal que $\{X_1(w_0), ..., X_n(w_0)\}$ són les nostres n observacions, llavors els residus

$$\hat{R}_i = \frac{X_i - \hat{X}_i}{\sqrt{r_{i-1}}}(w_0),$$

haurien de provenir de n variables aleatòries independents normals amb variància constant ($\hat{\sigma^2}$). Per a comprovar-ho existeixen molts de mètodes. Nosaltres utilitzarem les següents eines:

- **Graficar els residus per a comprovar homocedasticitat i independència:** la idea és veure en un gràfic els n residus $\hat{R}_i$, per a cada $1 \leq i \leq n$. Cal observar que aquests presentin variància constant (homocedasticitat), i que estiguin distribuïts completament a l'atzar al voltant del 0 (independència).

- **Comprovar normalitat dels residus:** hi ha diferents tests i mètodes per a comprovar que els residus $\hat{R}_i$ provenen d'una distribució normal. Nosaltres farem ús del test de Shapiro-Wilk, el qual amb un estadístic W calculat a partir dels residus $\hat{R}_i$, intenta resoldre el següent contrast d'hipòtesis:

  $$\begin{cases} H_0: & \text{els residus provenen d'una distribució normal }, \\ H_1: & \text{els residus no provenen d'una distribució normal }. \end{cases}$$

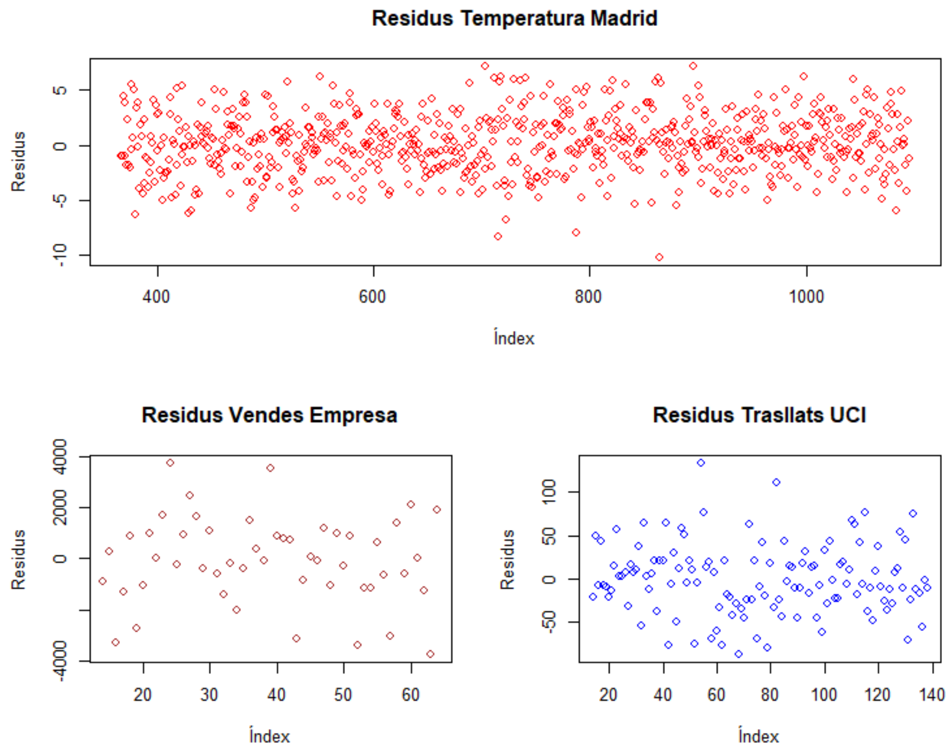
  Es rebutja l'hipòtesis nul·la de normalitat si l'estadístic W és menor que el valor crític (p-valor) proporcionat per la taula, elaborada pels autors del test, per a una grandària de la mostra (n), i un nivell de significació donats. Nosaltres treballarem sempre amb un nivell de significació del 5 per cent. És a dir, rebutjarem H_0 si el p-valor del test és menor que 0.05.

Si després d'aquests dos mètodes no tenim evidències significatives per a rebutjar que el model ajusti les nostres dades, haurem validat el model, és a dir, suposarem que la sèrie temporal ARMA(p, q) $\{X_t\}_{t \in \mathbb{Z}}$, amb els paràmetres estimats a l'anterior secció, és la que ajusta les nostres observacions, la qual compleix, per a tot enter t:

$$X_t - \hat{\phi}_1 X_{t-1} - ... - \hat{\phi}_p X_{t-p} = Z_t + \hat{\theta}_1 Z_{t-1} + ... + \hat{\theta}_q Z_{t-q}, \text{ amb } \{Z_t\}_{t \in \mathbb{Z}} \sim WN(0, \hat{\sigma^2}).$$

Per tant, existeix un $w_0 \in \Omega$, tal que les nostres observacions són $\{X_1(w_0), ..., X_n(w_0)\}$.

Exemple 3.13. Els models ARMA considerats en els nostres 3 exemples ajusten les nostres dades? En el següent gràfic vegem els residus acabats d'explicar per a cadascun dels exemples.



En els 3 gràfics podem veure que els residus es troben distribuïts aleatòriament al voltant de l'origen. A més, llevat d'alguns pocs valors específics, la variància es manté constant al llarg de cada gràfic. Efectivament, en el de la temperatura, gairabé tots els residus varien entre $(-7, 7)$, en el de les vendes, gairabé tots ells varien entre $(-4000, 4000)$, i, finalment, els residus dels trasllats semblen variar de forma constant entre $(-100, 100)$. És a dir, visualment, podem comprovar la homocedasticitat i independència dels residus.

Per altra banda, cal comprovar si els tres residus corresponents als tres exemples provenen d'una distribució normal. Realitzant el test de Shapiro-Wilk, obtenim p-valors de 0.1108 (Temperatura) 0.3651 (Vendes) i 0.2248 (Trasllats). Els 3, treballant amb un nivell de significació del 5 per cent, no són sinificatius (són majors que 0.05). És a dir, no tenim evidències significatives per a concloure que els residus no provenen de distribucions normals.

En conclusió, d'aquesta forma els 3 models ARMA que ajusten les nostres observacions queden validats i, per tant, donam com a certs els models considerats a l'exemple 3.12.

Al llarg d'aquest capítol hem vist com estimar els paràmetres del model ARMA causal i invertible del qual provenien les n dades. Ara bé, recordem del capítol 1, que originalment, la mostra prové d'una sèrie que admetia descomposició clàssica, la qual segueix un model que desconeixem. Un cop validat el model ARMA , com podem traslladar les prediccions, al model orginal? Aquesta és la qüestió que resoldrem en la següent secció.

4 Processos SARIMA

Recordem, del capítol 1, que nosaltres partim d'una sèrie temporal $\{X_t\}_{t\in\mathbb{Z}}$, que es podia descomposar de forma clàssica, i que a través d'una sèrie de procediments, fent ús de l'operador B, transformem en una sèrie estacionària, la qual hem modelitzat mitjançant un procés ARMA, del qual n'hem estimat els paràmetres. D'aquest en coneixem totes les propietats de segon ordre, i podem realitzar prediccions sense problemes a partir de les tècniques explicades en 2.1. Ara bé, de la sèrie original $\{X_t\}_{t\in\mathbb{Z}}$, no tenim aquesta informació, motiu pel qual no podem realitzar prediccions com a 2.1. L'objectiu en aquesta secció es veure com traslladar les prediccions del procés ARMA, a la de la sèrie original $\{X_t\}_{t\in\mathbb{Z}}$.

4.1 Els models SARIMA

Definició 4.1. *Si d i s són dos enters no negatius, diem que la sèrie temporal $\{X_t\}_{t\in\mathbb{Z}}$ (amb la mateixa notació de l'operador B de 1.3), és un procés SARIMA$(p,d,q)_s$ amb mitjana μ, si la sèrie temporal diferenciada $\{Y_t\}_{t\in\mathbb{Z}}$, definida per a tot enter t com a*

$$Y_t = \Big((I-B)^d(I-B^s)X_t\Big) - \mu,$$

és un procés ARMA(p,q) causal i invertible. És a dir,

$$\phi(B)Y_t = \theta(B)Z_t,$$

on $\{Z_t\}_{t\in\mathbb{Z}} \sim WN(0,\sigma^2)$, $\phi(z) = (1-\phi_1 z - ... - \phi_p z^p)$ i $\theta(z) = (1+\theta_1 z + ... + \theta_q z^q)$.

Exemple 4.2. Seguint amb els tres exemples tractats al llarg del treball,

- Respecte la temperatura de Madrid, recordem que les observacions inicials provenen de la sèrie temporal $\{X_t^1\}_{t\in\mathbb{Z}}$, i que la sèrie diferenciada $\{W_t^1\}_{t\in\mathbb{Z}}$, definida per a tot enter t com a (exemples 1.11 i 2.13)

  $$W_t^1 = \big((I_1-B_1)(I_1-B_1^{365})X_t^1\big) - (-0.001371742),$$

  és un procés ARMA$(5,3)$ amb els paràmetres coneguts (exemples 3.12 i 3.13). És a dir, concloem que $\{X_t^1\}_{t\in\mathbb{Z}}$, és un procés SARIMA $(5,1,3)_{365}$ amb mitjana (-0.001371742).

- Respecte les vendes de l'empresa , recordem que les observacions inicials provenen de la sèrie temporal $\{X_t^2\}_{t\in\mathbb{Z}}$, i que la sèrie diferenciada $\{W_t^2\}_{t\in\mathbb{Z}}$, definida per a tot enter t com a (exemples 1.11 i 2.13)

  $$W_t^2 = \big((I_2-B_2)(I_2-B_2^{12})X_t^2\big) - 33.60784,$$

  és un procés ARMA$(0,1)$ amb els paràmetres coneguts (exemples 3.12 i 3.13). És a dir, concloem que $\{X_t^2\}_{t\in\mathbb{Z}}$, és un procés SARIMA $(0,1,1)_{12}$ amb mitjana 33.60784.

- Respecte els traslats a la UCI, recordem que les observacions inicials provenen de la sèrie temporal $\{X_t^3\}_{t\in\mathbb{Z}}$, i que la sèrie diferenciada $\{W_t^3\}_{t\in\mathbb{Z}}$, definida per a tot enter t com a (exemples 1.11 i 2.13)

  $$W_t^3 = \big((I_3-B_3)(I_3-B_3^{12})X_t^3\big) - (-0.28),$$

  és un procés ARMA$(4,4)$ amb els paràmetres coneguts (exemples 3.12 i 3.13). És a dir, concloem que $\{X_t^3\}_{t\in\mathbb{Z}}$, és un procés SARIMA $(4,1,4)_{12}$ amb mitjana (-0.28).

4.2 Predicció en els models SARIMA

Sigui $\{X_t\}_{t\in\mathbb{Z}}$ un procés SARIMA $(p,d,q)_s$ amb mitjana μ , (amb tots els paràmetres coneguts) i $\{Y_t\}_{t\in\mathbb{Z}}$ el corresponent procés ARMA (p,q). És a dir, per a tot enter t,

$$Y_t + \mu = (I-B)^d(I-B^s)X_t.$$

Suposem que disposem d'una mostra de $n > d+s$ observacions del procés $\{X_t\}_{t\in\mathbb{Z}}$. És a dir, que $\exists w_0 \in \Omega$, tal que disposem de $\{X_{-d-s+1}(w_0),...,X_0(w_0),X_1(w_0)...,X_{n-d-s}(w_0)\}$. Per tant, per les tècniques de diferenciació explicades al primer capítol, disposem de la mostra $\{Y_1(w_0),...,Y_{n-d-s}(w_0)\}$ del procés ARMA $\{Y_t\}_{t\in\mathbb{Z}}$, (utilitzem aquesta enumeració perque la mostra del procés ARMA quedi de 1 fins a $n-d-s$, però aquesta pot ser de distint ordre, com en els exemples que estem tractant).

Considerem els subespais de $L^2(\Omega)$,

$$F_1 =< 1, X_{-d-s+1},...,X_0,X_1,...,X_{n-d-s} >,$$
$$F_2 =< 1, X_{-d-s+1},...,X_0,Y_1,...,Y_{n-d-s} > .$$

Amb la mateixa notació que en 2.1, el nostre objectiu és, per a $h > 0$, trobar el predictor lineal $P_{F_1}X_{n-d-s+h}$, per així calcular $P_{F_1}X_{n-d-s+h}(w_0)$ i estimar numèricament l'observació en el futur.

Per la definició de $\{Y_t\}_{t\in\mathbb{Z}}$, tenim que existeix $(a_1,...,a_{d+s}) \in \mathbb{R}^{d+s}$, tal que per a tot enter t,

$$Y_t + \mu = X_t + \sum_{j=1}^{d+s} a_j X_{t-j}.$$

Fixem-nos que per a $1 \le t \le n-d-s$, podem observar que $Y_t \in F_1$. És a dir, $F_2 \subseteq F_1$.

Per altra banda,

$$X_1 = \mu + Y_1 - \sum_{j=1}^{d+s} a_j X_{1-j} \in F_2.$$

De la mateixa forma, recursivament, es veu que per a $2 \le t \le n-d-s$, tenim que $X_t \in F_2$. És a dir, $F_1 \subseteq F_2$. Per tant, cocloem que $F_1 = F_2$.

Calculem, per a $h > 0$, $P_{F_1}X_{n-d-s+h}$. Com

$$X_{n-d-s+h} = \mu + Y_{n-d-s+h} - \sum_{j=1}^{d+s} a_j X_{n-d-s+h-j},$$

deduïm que (és conseqüència del teorema de la projecció ortogonal)

$$P_{F_1}X_{n-d-s+h} = \mu + P_{F_1}Y_{n-d-s+h} - \sum_{j=1}^{d+s} a_j P_{F_1}X_{n-d-s+h-j}.$$

Ara bé, com $F_1 = F_2$, tenim que $P_{F_1}Y_{n-d-s+h} = P_{F_2}Y_{n-d-s+h}$.

Si suposem que $(X_{-d-s+1},...,X_0)$ està incorrelacionat amb $\{Y_t\}_{t\ge 1}$, és a dir, que $Cov(X_{s_1}, Y_{s_2}) = 0$, si $s_1 \in \{-d-s+1,...,0\}$ i $s_2 \ge 1$, llavors per l'observació 2.3 arribem a la conclusió que, considerant el subespai $G =< 1, Y_1,...,Y_{n-d-s} >\subseteq L^2(\Omega)$,

$$P_{F_2}Y_{n-d-s+h} = P_G Y_{n-d-s+h}.$$

Aquest fet es deu a que sota aqueixa hipòtesis les equacions dels coeficients d'ambdos predictors lineals coincideixen. Com $\{Y_t\}_{t\in\mathbb{Z}}$ és un procés ARMA(p,q) amb els paràmetres coneguts (en coneixem les propietats de segon ordre), podem calcular $P_G Y_{n-d-s+h}$ com a 2.1, i en conseqüència també $P_G Y_{n-d-s+h}(w_0)$ a partir de la mostra $\{Y_1(w_0), ..., Y_{n-d-s}(w_0)\}$. Per tant,

$$P_{F_1} X_{n-d-s+h} = \mu + P_G Y_{n-d-s+h} - \sum_{j=1}^{d+s} a_j P_{F_1} X_{n-d-s+h-j}.$$

Notem que, per a $h = 1$,

$$P_{F_1} X_{n-d-s+1-j}(w_0) = X_{n-d-s+1-j}(w_0) \in F_1, \text{ si } j \in \{1, .., d+s\}.$$

És a dir, podem calcular $P_{F_1} X_{n-d-s+1}(w_0)$ a partir de les dades

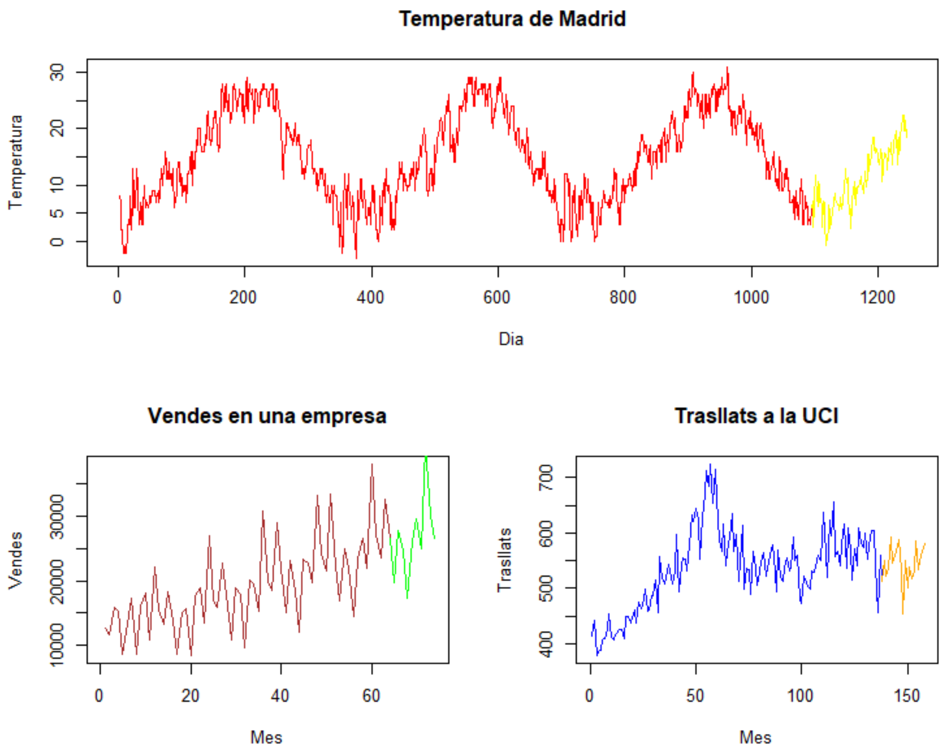
$$\{Y_1(w_0)..., Y_{n-d-s}(w_0), X_{-d-s+1}(w_0), ..., X_0(w_0), ..., X_{n-d-s}(w_0)\}.$$

Així, recursivament ($h = 2$, llavors $h = 3$, ...) , podem deteminar $P_{F_1} X_{n-d-s+h}(w_0)$ per a $h \geq 1$, a partir d'aquest mateix conjunt de dades.

Exemple 4.3. Amb aquesta metodologia, podem calcular prediccions per a les sèries temporals SARIMA $\{X_t^1\}_{t\in\mathbb{Z}}$, $\{X_t^2\}_{t\in\mathbb{Z}}$, i $\{X_t^2\}_{t\in\mathbb{Z}}$. És a dir, per a tot $h \in \mathbb{N}$, podem estimar numèricament

$$X^1_{1095+h}(w^1_0),\ X^2_{64+h}(w^2_0) \text{ i } X^3_{138+h}(w^3_0).$$

El següent gràfic mostra les primeres 150 prediccions per a $\{X_t^1\}_{t\in\mathbb{Z}}$ (la temperatura dels 150 primers dies del 2012, de color groc), les primeres 10 per $\{X_t^2\}_{t\in\mathbb{Z}}$ (les vendes del maig de 2020 fins al febrer del 2021, de color verd) i les primeres 20 per a $\{X_t^3\}_{t\in\mathbb{Z}}$ (els trasllats del maig de 2011 fins el desembre de 2012, de color taronja).



5 Conclusions

Com havíem explicat al resum inicial, hem partit de 3 conjunts d'observacions que evolucionen al llarg del temps, els quals, des del punt de vista probabilístic, hem modelitzat mitjançant un procés SARIMA. Ens hem encarregat de justificar que els models són adients, és a dir, que ajusten els tres conjunts de dades, mitjançant els diagnòstics de la secció 3.2.3. Finalment, hem estat capaços de realitzar futures prediccions per a cadascun dels models, com es pot visualitzar a l'últim gràfic de l'anterior capítol, el qual reflexa l'últim objectiu d'aquest treball.

La idea principal de tota la tasca realitzada, és que el lector, seguint els passos utiltzats al llarg del projecte, pugui intentar modelitzar qualsevol conjunt de n dades $\{x_1, ..., x_n\}$ recollides cada cert període temporal mitjanant un procés SARIMA. Ara bé, cal dir que no sempre serà possible ajustar un d'aquests models, és a dir, existeix la possibilitat que els diagnòstics de la secció 3.2.3 fallin, fet que ens indicaria que caldria buscar altres models per intentar ajustar les nostres observacions.

Referències

[1] Peter J. Brockwell and Richard A. Davis. *Introduction to Time Series and Forecasting.* Springer, Third Edition, 2016.

[2] Peter J. Brockwell and Richard A. Davis. *Time Series: Theory and Methods.* Colorado State University, 1987.

[3] Daniel Peña *Análisis de series temporales.* Alianza editorial, Madrid, 2010.

[4] Olga Julià i David Márquez-Carreras. *Un primer curs d'estadísica.* Publicacions i Edicions de la Universitat de Barcelona, 2011.

[5] Ahlfors, L.V. *Complex analysis: an introduction to the theory of analytic functions of one complex variable.* 3th ed.New York:McGraw-Hill, 1979.

[6] Bruna, J. *Anàlisi Real.* Barcelona: Universitat Autònoma de Barcelona, Servei de Publicacions, 1996.

[7] Sanz-Solé, M. *Probabilitats*, Vol 28, Publicacions i Edicions de la UB, Barcelona, 1999.

[8] Shumay, R.H and Stoffer D.S. *Time series analysis and its applications with R examples.* Second Edition. Springer, 2006.

[9] George Weigt. *ITSM-R Reference Manual.* [Recurs online].
`https://georgeweigt.github.io/itsmr-refman.pdf`

[10] `https://www.kaggle.com/juliansimon/weather_madrid_lemd_1997_2015.csv`

[11] `https://archive.ics.uci.edu/ml/datasets/Horton+General+Hospital`

[12] `https://www.kaggle.com/podsyp/time-series-starter-dataset?select=Month_Value_1.csv`