

Llicenciatura en matemàtiques

Títol: Anàlisi multivariant de sèries temporals

Autor: Mazaira Font, Ferran-Antonio

Director: Muñoz Gracia, Pilar

Departament: Estadística i Investigació Operativa

Convocatòria: 2012



Facultat de Matemàtiques
i Estadística

UNIVERSITAT POLITÈCNICA DE CATALUNYA



FACULTAT DE MATEMÀTIQUES I ESTADÍSTICA

Anàlisi multivariant de sèries temporals

Autor:

Ferran A. Mazaira

Supervisora:

Dra. Pilar Muñoz Gràcia

Pròleg

Els mètodes per a l'anàlisi de sèries temporals han viscut una autèntica revolució en els darrers anys, atès l'interés que ha despertat la seva capacitat predictiva. Per al tractament de sèries univariants, s'han desenvolupat mètodes i algorismes que permeten el seu modelatge en una gran diversitat de situacions: des de sèries amb volatilitat com les dels mercats de valors fins a sèries amb estacionalitat com l'atur. No obstant, per a l'anàlisi multivariant de sèries, és a dir, per a modelar diverses sèries conjuntament, la investigació encara es troba en una fase més inicial, i fenòmens freqüents com l'estacionalitat o les tendències de creixement o decreixement a llarg termini no estan del tot resolts. L'objectiu d'aquest treball és presentar els conceptes i mètodes fonamentals de l'anàlisi multivariant de sèries temporals, així com veure'n algunes aplicacions.

El primer capítol del projecte es centra en introduir el marc teòric dels processos VARMA, autoregressius i de mitjanes mòbils vectorials, que són l'extensió dels processos ARMA. Es presentaran les propietats estadístiques que permeten detectar els processos VARMA, això és, la constància dels seus moments de primer i segon ordre (estacionarietat) i els diferents paradigmes que permeten modelar-los: els models autoregressius, els models de mitjanes mòbils i els que combinen tots dos. En aquest sentit, el treball conté un resultat original que relaciona l'estacionarietat d'un procés amb les propietats d'una corba tancada en el pla complex associada al mateix. Finalment, posarem especial ènfasi en la identificació i diagnosi del model, i introduïrem els darrers resultats en la recerca relacionada amb processos multivariants estacionals multiplicatius, presentant un criteri nou basat en l'ajustament del soroll no observable de les sèries.

El segon capítol està dedicat a l'explicació de les tècniques de mínims quadrats i maximo semblança per a l'estimació dels paràmetres dels models VARMA. Encara que no és freqüent en el camp de l'estimació, atès que els procediments sempre inclouen una estimació de l'error estàndard dels paràmetres, l'estudi de l'error i la convergència dels algorismes ocupa una part central del capítol, per a situar al lector sobre les debilitats i virtuts dels mateixos. Els resultats més importants que veurem són, d'una banda,

l'impacte desastrós que pot tenir en l'estimació dels paràmetres de processos que contenen operadors de mitjanes mòbils, el fet de necessitar estimar el soroll, que no és observable. D'altra banda, veurem que algunes propietats dels models VARMA, relacionades les seves components principals o amb els vectors propis de les matrius que els defineixen, poden provocar que no siguin estadísticament estimables els seus paràmetres sense cometre errors de gran magnitud.

Finalment, el tercer capítol presenta l'estudi de les sèries econòmiques de la Renta Disponible per càpita, el Consum Personal i la Renta del Capital dels Estats Units d'Amèrica. Les tres sèries es modelitzaran conjuntament a partir d'un procés VARMA i s'estudiaran amb deteniment les implicacions econòmiques de les relacions entre elles que dit model posarà de manifest.

Voldria expressar el meu agraïment a la supervisora d'aquest treball, la Dra. Pilar Muñoz Gràcia, per la llibertat que m'ha donat en el plantejament del mateix i per haver-me orientat i ofert sempre la seva ajuda. Al Dr. Joan Solà Morales i a Xavier Ros-Otón per les aportacions que han fet en la demostració d'alguns dels resultats d'aquest treball i a l'Alberto López Moreno per la seva ajuda en l'adaptació del codi dels algorismes programats en `MATLAB` a `R`.

Ferran A. Mazaira

Juny de 2012

Índex

1	Models multivariants	5
1.1	Sèries multivariants estacionàries	5
1.1.1	Anàlisi exploratòria de sèries temporals	6
1.1.2	Matrius de covariància, correlació simple i parcial	11
1.2	Models ARMA vectorials	13
1.2.1	Els processos VAR(p)	16
1.2.2	Els processos VMA(q)	18
1.2.3	Els processos VARMA(p,q)	19
1.3	Construcció de models VARMA per a sèries estacionàries	23
1.3.1	Identificació de l'estructura del model	23
1.3.2	Estimació i selecció de models	25
1.3.3	Diagnosi	25
1.4	Previsió	26
1.4.1	Predictor amb Error Quadràtic Mitjà Mínim	26
1.4.2	Previsió puntual i per intervals i regions de confiança per als models VARMA	27
1.5	Models VARMA estacionals multiplicatius	29
1.5.1	Selecció del millor model estacional multiplicatiu	30
2	Estimació i previsió	34
2.1	Estimació dels paràmetres d'un procés VAR(p)	34
2.1.1	Estimació sotmesa a constriccions lineals	35
2.2	Estimació dels paràmetres d'un procés VMA(q)	37
2.3	Estimació dels paràmetres d'un model VARMA(p,q)	39
2.3.1	Equacions normals de la maximització de la versemblança	40

2.3.2	Algorisme de màxima versemblança o dels “scores ”	41
2.3.3	Estimació sotmesa a constriccions lineals	44
2.4	Anàlisi de la convergència i l'error de l'algorisme	46
2.4.1	Metodologia	47
2.4.2	Algorisme de mínims quadrats per a processos VAR	47
2.4.3	Algorisme de màxima versemblança per a processos VMA	49
2.4.4	Algorisme de màxima versemblança per a processos VARMA	55
2.5	Estimació de previsions	60
3	Cas pràctic	65
3.1	Anàlisi exploratòria	65
3.2	Identificació del model	68
3.3	Estimació del model	70
3.4	Diagnosi	71
3.5	Efectes multiplicadors	76
3.6	Conclusions	79
A	Simulació de variables	81

Capítol 1

Models multivariants

En els darrers anys, la possibilitat d'emmagatzemar i operar amb grans quantitats de dades ha portat al desenvolupament d'una gran varietat de models i tècniques estadístiques per a la seva anàlisi. En concret, un dels camps que més s'ha desenvolupat ha estat el de la previsió, especialment per la seva importància en el camp de l'economia, la política i els mercats financers. Per exemple, preveure l'evolució del creixement econòmic, de la taxa d'atur o del nivell general de preus és de vital importància per poder desenvolupar les polítiques econòmiques adequades.

L'objectiu d'aquest capítol és presentar de forma breu el marc teòric dels models ARMA vectorials, que permeten modelar conjuntament les relacions entre les components d'un vector de sèries temporals, sota determinades hipòtesis, i fer previsions de la seva evolució futura. Es detallarà la manera de construir aquests models a partir de sèries empíriques i com seleccionar el més adequat.

1.1 Sèries multivariants estacionàries

Sigui $Y_t = (Y_t^1, \dots, Y_t^K)$, $t = 0, 1, 2, \dots$ un vector K -dimensional de sèries temporals, les quals, en el cas més general, estan interrelacionades tant en el mateix instant de temps com respecte alguns dels anteriors.

Definició 1.1. Diem que $\{Y_t\}$ és un *procés estacionari en sentit estricte* si les distribucions de probabilitat dels vectors aleatoris $\{Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}\}$ i $\{Y_{t_1+l}, Y_{t_2+l}, \dots, Y_{t_n+l}\}$ són les mateixes per a qualssevol t_1, t_2, \dots, t_n i $l = 0, \pm 1, \pm 2, \dots$

Per tant, la distribució de probabilitat de les observacions d'un procés estacionari vectorial és invariant respecte els desplaçaments en el temps. Assumint que els moments de primer i segon ordre existeixen i són finits, es compleix que $E(Y_t) = \mu$ és constant per a tot t , amb $\mu = (\mu_1, \mu_2, \dots, \mu_K)'$ el vector mitjana

del procés. A més a més, els vectors Y_t tenen matriu de covariàncies constant per a tot t , que notarem per $\Sigma_y \equiv \Gamma(0) = E[(Y_t - \mu)(Y_t - \mu)']$.

Les condicions d'estacionaritat en sentit estricte són molt fortes i difícils de verificar empíricament, de forma que utilitzarem una definició menys restrictiva, que es refereix només als moments de primer i segon ordre.

Definició 1.2. Direm que $\{Y_t\}$ és un *procés estacionari de segon ordre o en sentit ampli* si satisfà que $E(Y_t) = \mu$ no depèn de t i $E[(Y_t - \mu)(Y_{t+l} - \mu)']$ només depèn de l .

1.1.1 Anàlisi exploratòria de sèries temporals

Les sèries que ens trobem en la realitat sovint no són estacionàries i acostumen a presentar components de tendència (mitjana no constant), estacionalitat (oscil·lacions al voltant de la tendència, amb cicles d'una periodicitat d'ordre s constant) o variància no constant. Per aquest motiu, es realitzen transformacions sobre les sèries que eliminin aquests fenòmens i permetin considerar-les estacionàries. En primer lloc es faran les transformacions per a fer constant la variància i després les que eliminin la tendència, l'estacionalitat o estabilitzen la mitjana.

Transformacions per estabilitzar la variància

Quan ens trobem amb processos temporals que presenten variància no constant realitzem les anomenades transformacions de Box-Cox. Donada una sèrie X_t , la transformació de Box-Cox T_t de paràmetre λ es defineix com:

$$T_t = \begin{cases} \frac{X_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log(X_t) & \lambda = 0 \end{cases}$$

on λ s'estima per màxima versemblança [3].

Transformacions per eliminar la tendència

Suposem que tenim un procés estacionari X_t d'esperança μ i que observem el procés definit per

$$Y_t = at + X_t$$

És clar que Y_t no és estacionari, ja que la mitjana no és constant i presenta una tendència lineal determinista. Observem que si realitzem la transformació

$$Z_t = (1 - B)Y_t$$

on B és l'operador retard que aplicat sobre una sèrie actua com: $BY_t = Y_{t-1}$, aleshores tenim que

$$E[Z_t] = E[Y_t - Y_{t-1}] = \alpha t + \mu - \alpha(t-1) - \mu = \alpha$$

de manera que hem eliminat la tendència i tenim un procés estacionari.

En general, aplicarem la transformació $(1 - B)^d$ per eliminar tendències polinòmiques deterministes o estocàstiques de grau d , i la transformació $(1 - B)$ per estabilitzar la mitjana quan, sense presentar una tendència clara, es mostra no constant.

Transformacions per eliminar l'estacionalitat

Suposem que tenim un procés estacionari X_t d'esperança μ i que observem el procés definit per

$$Y_t = X_t + S_t$$

on S_t és una component estacional d'ordre s que recull les variacions degudes a l'evolució de la variable en instants del temps que es repeteixen de forma regular cada cert temps. Notem que això vol dir que

$$\sum_{i=1}^s S_{t+i} = 0$$

per a qualsevol t . És clar que Y_t no és estacionari, ja que la mitjana no és constant per aquesta estacionalitat. Observem que si realitzem la transformació

$$Z_t = (1 - B^s)Y_t$$

on B^s és l'operador retard d'ordre s que aplicat sobre una sèrie actua com: $B^s Y_t = Y_{t-s}$, aleshores tenim que

$$E[Z_t] = E[Y_t - Y_{t-s}] = \mu + E[S_t] - \mu - E[S_{t-s}] = 0$$

per ser S_t d'ordre s . Obtenim amb aquesta transformació un procés estacionari.

Exemple 1.1. Estudiem la sèrie *AirPassengers* continguda en la llibreria MASS del software R [4]. Aquesta sèrie, que notarem per X_t , conté el nombre de passatgers mensuals d'una aerolínea internacional,

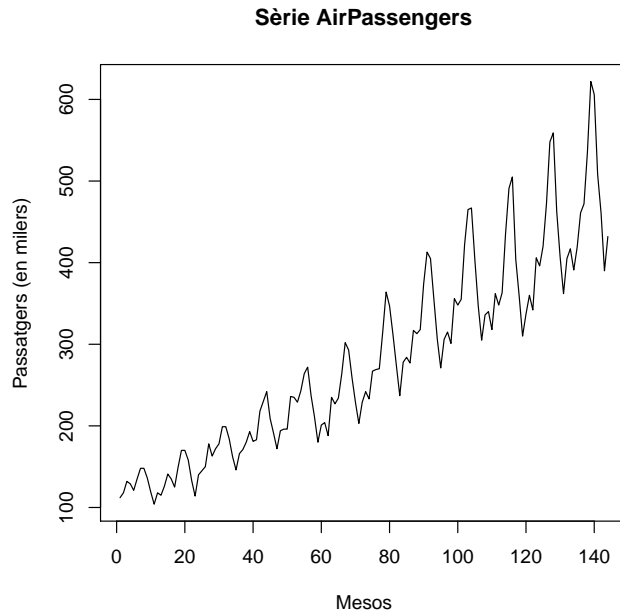


Figura 1.1: Representació de la sèrie *AirPassengers*.

durant els anys 1949-1960. En la figura 1.1 es mostra la sèrie X_t i podem veure com la variància no és constant, sinó que augmenta amb el temps, la sèrie presenta una tendència creixent i estacionalitat d'ordre $s = 12$, atès que les dades són mensuals.

Per a aconseguir que la variància sigui constant, realitzem una transformació de Box-Cox. En la figura 1.2 es mostra que el paràmetre λ estimat per màxima versemblança és $\hat{\lambda} = 0.051$, però com l'interval de confiança al 95% conté el 0, farem la transformació

$$Y_t = \log(X_t)$$

La sèrie Y_t es mostra en la figura 1.3, i s'observa clarament com la variància s'ha fet més homogènia. Per eliminar l'estacionalitat, ara de la sèrie Y_t , farem una diferenciació d'ordre 12, és a dir, obtindrem la sèrie

$$Z_t = (1 - B^{12})Y_t = (1 - B^{12})\log(X_t)$$

En la gràfica 1.4 apareix la sèrie Z_t . Com veiem, ara ja no hi ha estacionalitat però cal estabilitzar la mitjana. Per aquest motiu, fem una última diferenciació, aquest cop d'ordre 1, i obtenim la sèrie

$$W_t = (1 - B)Z_t = (1 - B)(1 - B^{12})\log(X_t)$$

representada en la figura 1.5.

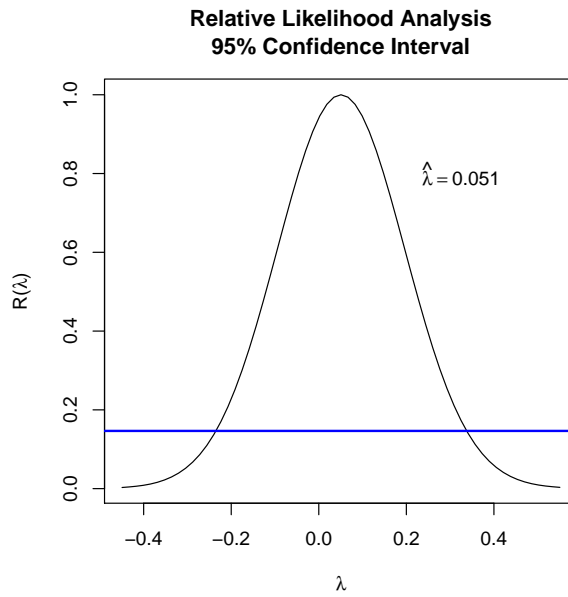


Figura 1.2: Estimació del paràmetre λ de la transformació de Box-Cox per a la sèrie X_t . Els punts de tall entre la recta blava i la distribució de l'estimador determinen l'interval de confiança al 95%.

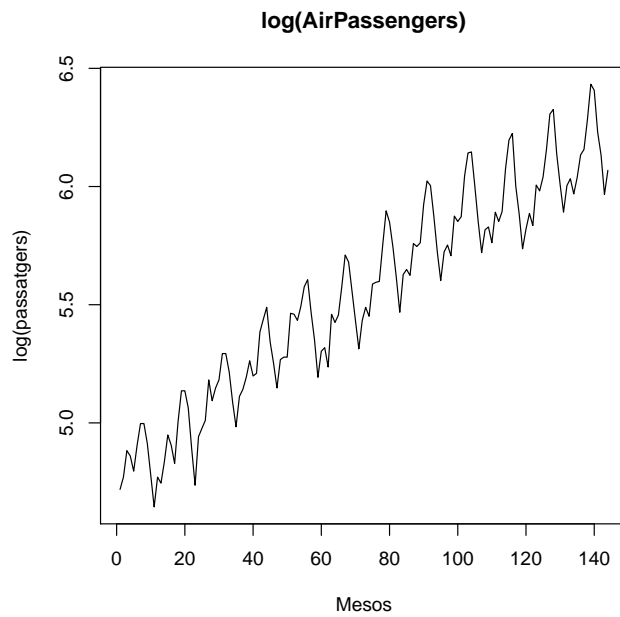


Figura 1.3: Representació de la sèrie $Y_t = \log(X_t)$.

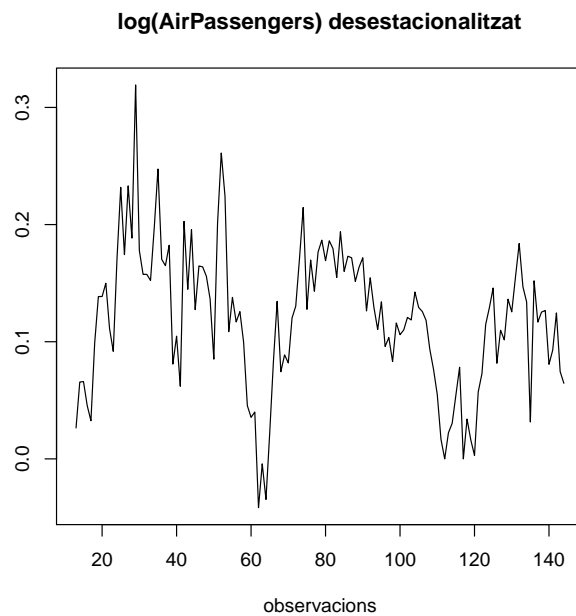


Figura 1.4: Representació de la sèrie $Z_t = (1 - B^{12})\log(X_t)$.

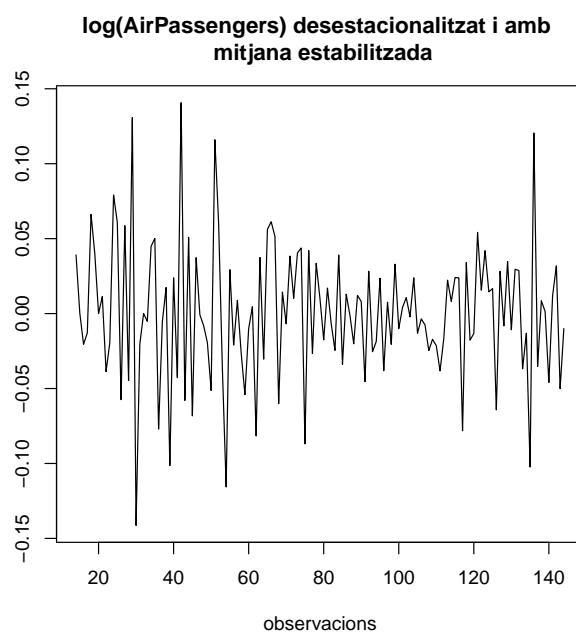


Figura 1.5: Representació de la sèrie $W_t = (1 - B)(1 - B^{12})\log(X_t)$.

1.1.2 Matrius de covariància, correlació simple i parcial

Donat un procés estocàstic $\{Y_t\}$, notarem per $\{Y_t^i\}$ el procés corresponen a la seva variable i -èssima. Per exemple, si el procés $\{Y_t\} = \{(preus_t, atur_t)'\}$ conté les dades del nivell de preus i l'atur d'un país, el procés $\{Y_t^1\}$ correspondria a l'evolució del nivell de preus $\{(preus_t)\}$, i $\{Y_t^2\}$ al procés $\{(atur_t)\}$.

En un procés estacionari $\{Y_t\}$ la covariància entre $\{Y_t^i\}$ i $\{Y_{t+l}^j\}$ depèn exclusivament del retard l , no del temps t , per a tot $i, j = 1, 2, \dots, K$ i $l = 0, \pm 1, \pm 2, \dots$. Per això, notem

$$\gamma_{ij}(l) = Cov(\{Y_t^i\}, \{Y_{t+l}^j\}) = E[(\{Y_t^i - \mu_i\})(\{Y_{t+l}^j - \mu_j\})]$$

i la matriu $K \times K$ de covariàncies creuades en el retard l com

$$\Gamma(l) = E[(\{Y_t - \mu\})(\{Y_{t+l} - \mu\})'] = \begin{pmatrix} \gamma_{11}(l) & \gamma_{12}(l) & \dots & \gamma_{1k}(l) \\ \gamma_{21}(l) & \gamma_{22}(l) & \dots & \gamma_{2k}(l) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \gamma_{k1}(l) & \gamma_{k2}(l) & \dots & \gamma_{kk}(l) \end{pmatrix}$$

Definició 1.3. Definim la *matriu de correlacions creuades* (correlació simple) en el retard l -èssim del procés $\{Y_t\}$ com

$$\rho(l) = V^{-1}\Gamma(l)V^{-1} = \begin{pmatrix} \rho_{11}(l) & \rho_{12}(l) & \dots & \rho_{1k}(l) \\ \rho_{21}(l) & \rho_{22}(l) & \dots & \rho_{2k}(l) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \rho_{k1}(l) & \rho_{k2}(l) & \dots & \rho_{kk}(l) \end{pmatrix}$$

on $V^{-1} = Diag\{\gamma_{11}(0)^{-1/2}, \dots, \gamma_{KK}(0)^{-1/2}\}$, atès que

$$\rho_{ij}(l) = Corr(\{Y_t^i\}, \{Y_{t+l}^j\}) = \gamma_{ij}(l)/[\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2}$$

amb $\gamma_{ii}(0) = Var(\{Y_t^i\})$.

Per al cas $i = j$, $\rho_{ii}(l) = \rho_{ii}(-l)$ correspon a la funció d'autocorrelació de la i -èssima sèrie Y_t^i , i per a $i \neq j$, $\rho_{ij}(l) = \rho_{ji}(-l)$ denota la correlació creuada entre les sèries $\{Y_t^i\}$ i $\{Y_t^j\}$. Noti's que $\Gamma(l)' = \Gamma(-l)$ i $\rho(l)' = \rho(-l)$, ja que $\gamma_{ij}(l) = \gamma_{ji}(-l)$. Endemés, les matrius de covariàncies $\Gamma(l)$ i les de correlacions creuades $\rho(l)$ són definides no negatives, en el sentit que $\sum_{i=1}^n \sum_{j=1}^n b_i' \Gamma(i-j) b_j \geq 0$ donat n , i qualssevols vectors K -dimensionals b_1, \dots, b_n , donat que $Var(\sum_{i=1}^n b_i Y_{t-i}) \geq 0$.

Els elements de les matrius de covariàncies i autocorrelació simple s'estimen mitjançant:

$$\hat{\gamma}_{ij}(m) = \sum_{t=1}^{T-m} (Y_t^i - \bar{Y}^i)(Y_t^j - \bar{Y}^j)$$

$$\hat{\rho}_{ij}(m) = \frac{\hat{\gamma}_{ij}(m)}{\sqrt{\hat{\gamma}_{ii}(0)\hat{\gamma}_{jj}(0)}}$$

on $\bar{y}_i = T^{-1} \sum_{t=1}^n y_{it}$. Com en el cas univariant, es demostra que si tenim un procés de soroll blanc vectorial, els coeficients estimats $\hat{\rho}_{ij}(m)$, per a $m > 0$, tenen una distribució asimptòticament normal, amb mitjana zero i variància T^{-1} . Per tant, considerarem no significativament diferents de zero amb un nivell de significació α , aquells coeficients menors, en valor absolut, a $z_{1-\alpha/2}/\sqrt{T}$, on $z_{1-\alpha/2}$ és el percentil $1 - \alpha/2$ de la distribució normal estàndard.

De la mateixa manera que hem generalitzat el concepte de matriu d'autocorrelació simple, podem introduir les matrius d'autocorrelació parcial del procés.

Definició 1.4. Definim la *matriu d'autocorrelació parcial* d'ordre m d'un procés estacionari vectorial $\{\tilde{Y}_t\} = \{Y_t - \mu\}$ d'esperança zero, com la matriu $P(m)$ obtinguda ajustant per mínims quadrats el model lineal multivariant:

$$\tilde{Y}_t = \Pi_1 \tilde{Y}_{t-1} + \dots + \Pi_{m-1} \tilde{Y}_{t-m-1} + P(m) \tilde{Y}_{t-m} + u_t$$

Notem que en aquest model cada sèrie s'explica per regressió sobre els seus m valors passats i els m valors passats de les altres sèries. Com en el cas escalar, $P(m)$ explica la relació lineal existent entre el vector \tilde{Y}_t i el vector \tilde{Y}_{t-m} un cop tinguts en compte els efectes de les variables explicatives $\tilde{Y}_{t-1}, \dots, \tilde{Y}_{t-m-1}$.

En la pràctica, aquestes matrius s'estimen a partir d'estimar $\tilde{Y}_t = Y_t - \bar{Y}$, és a dir, substituint l'esperança μ per les mitjanes mostrals.

Per a la identificació de models ARMA vectorials és cabdal conèixer com són aquestes matrius, i molt especialment si es poden considerar nul·les o no. En aquest cas, a diferència de les matrius de correlació simple, no coneixem la distribució dels coeficients de les matrius, i per això no podem plantejar contrastos d'hipòtesis sobre els mateixos. Notant per $\hat{P}(m)$ les estimacions de les matrius d'autocorrelació parcial, es pot construir un contrast [14] per decidir si les matrius es poden considerar zero o no, és a dir, un test per a resoldre:

$$H_0 : P(m) = 0$$

$$H_1 : P(m) \neq 0$$

El contrast comença amb $P(1)$. Si no és zero, es contrasta si ho es $P(2)$, condicionat a què $P(1)$ és diferent de zero, i així successivament. A continuació detallem amb profunditat aquest contrast. Sigui

$$e_{t,m} = \tilde{Y}_t - \Pi_1 \tilde{Y}_{t-1} - \dots - \Pi_{m-1} \tilde{Y}_{t-m-1} - \hat{P}(m) \tilde{Y}_{t-m}$$

els residus de l'ajustament en estimar el coeficient d'autorrelació parcial d'ordre m , i

$$S(m) = \sum_{t=m+1}^T e_{t,m} e'_{t,m}$$

la matriu de suma de quadrats d'aquests residus. El contrast de raó de versemblança de què la matriu de correlació parcial d'ordre m és nul·la, condicionat a què la matriu d'ordre $m - 1$ no ho sigui, depèn de la matriu de variàncies dels residus estimats mitjançant

$$\Lambda = \frac{|S(m)|}{|S(m-1)|}$$

El contrast analitza si la reducció de residus al quadrat és significativa o no. Sota la hipòtesi que $P(m)$ és zero, la distribució d'aquest estadístic es pot aproximar a una χ^2 central mitjançant la transformació:

$$\chi^2 = - \left(T - \frac{1}{2} - mk \right) \log \Lambda,$$

que es distribueix asimptòticament com una χ^2 amb k^2 graus de llibertat, el nombre de termes de la matriu $\hat{P}(m)$, fet que ens permet fer el contrast de què la matriu sigui nul·la.

1.2 Models ARMA vectorials

La generalització multivariant del teorema de Wold [19] indica que tot procés vectorial estacionari sense components deterministes pot ser representat com:

$$\tilde{Y}_t = \Psi(B)u_t = \sum_{i=1}^{\infty} \Psi_i u_{t-i}$$

on $\tilde{Y}_t = Y_t - \mu$, B és l'operador de retard, que aplicat sobre vectors complex $B(Y_t) = Y_{t-1}$, Ψ_i són matrius quadrades $k \times k$, amb $\Psi_0 = I_k$ la matriu identitat, i u_t vectors de soroll blanc de mitjana zero i matriu de covariàncies Σ_u definida positiva.

Anàlogament al cas univariant, els models ARMA vectorials surdeixen com una representació finita de l'operador $\Psi(B)$. En efecte, suposarem que es pot aproximar pel producte de dos operadors finits:

$$\tilde{Y}_t = \Phi(B)^{-1} \Theta(B)u_t$$

on

$$\Phi(B) = \Phi_0 - \Phi_1 B - \dots - \Phi_p B^p$$

és un operador en B i els termes Φ_i són matrius quadrades d'ordre k que contenen els coeficients autor-regressius del procés. L'operador de mitjana mòbil és

$$\Theta(B) = \Theta_0 + \Theta_1 B + \dots + \Theta_q B^q$$

De la mateixa manera que en el cas univariant, per a què els processos VARMA siguin estacionaris i la seva part de mitjanes mòbils invertible cal que satisfacin unes determinades condicions referides a les arrels de les equacions $|\Phi(z)| = 0$ i $|\Theta(z)| = 0$. Per als models ARMA unidimensionals, les condicions es detallen en el teorem següent:

Teorema 1.5. *Sigui $\Phi(B)Y_t = \Theta(B)u_t$ un model ARMA(p, q). Aleshores, el procés és estacionari si i només si les arrels de l'equació $|\Phi(z)| = 0$ estan fora del cercle unitat, i és invertible si i només si les arrels de l'equació $|\Theta(z)| = 0$ estan també fora del cercle unitat.*

Demostració. Vegem el cas d'invertibilitat, ja que l'estacionaritat es fa anàlogament.

Siguin z_1, \dots, z_q les arrels del polinomi $\Theta(B)$. Aleshores, podem escriure $\Theta(B) = (1 - \frac{B}{z_1}) \cdots (1 - \frac{B}{z_q})$. Els operadors $(1 - \frac{B}{z_i})$ són invertibles si i només si $|z_i| > 1$, ja que són el resultat d'una progressió geomètrica de raó $\frac{B}{z_i}$. En conclusió, l'operador Θ és invertible si i només si les arrels z_i tenen mòdul major que 1. \square

Per a trobar les dels models VARMA, enunciem primer un resultat previ:

Teorema 1.6 (Teorema de Smith, [7]). *Sigui $M(s)$ una matriu polinomial de rang r de dimensions $m_1 \times m_2$. Aleshores, $M(s)$ és equivalent o bé a una matriu polinomial $D_f(s)$ (per $m_1 < m_2$) o a $D_c(s)$ (for $m_2 < m_1$), amb*

$$D_f(s) = [D(s) \quad 0] \quad D_c = \begin{bmatrix} D(s) \\ 0 \end{bmatrix}$$

on $D(s) = \text{diag}(e_1(s), \dots, e_r(s), 0, \dots, 0)$ i els elements $e_i(s)$ són polinomis mòncics tals que $e_i(s)$ divideix $e_{i+1}(s)$.

Si $m_1 = m_2$, aleshores $M(s)$ és equivalent a una matriu quadrada diagonal $D(s)$

Amb aquest resultat, estem en condicions d'enunciar les condicions per a què un procés ARMA vectorial sigui estacionari i invertible:

Teorema 1.7. *Sigui $\Phi(B)Y_t = \Theta(B)u_t$ un procés estocàstic amb u_t soroll blanc gaussià. Aleshores, el procés és estacionari si i només si les arrels de l'equació $|\Phi(z)| = 0$ estan fora del cercle unitat, i és invertible si i només si les arrels de l'equació $|\Theta(z)| = 0$ estan també fora del cercle unitat.*

Demostració. Com en el cas dels models ARMA, veurem només la condició d'invertibilitat.

Pel Teorema de Smith, tenim que existeixen dues matrius polinomials $E(B)$ i $F(B)$ amb determinant constant diferent de zero i una matriu diagonal $D(B)$ tals que $\Theta(B) = E(B)D(B)F(B)$. Per tant, $\Phi(B)Y_t = E(B)D(B)F(B)u_t$. Com que $E(B)$ té inversa polinomial, tenim que $E^{-1}(B)\Phi(B)Y_t = D(B)F(B)u_t$. Com que $D(B)$ és diagonal, és invertible si i només si ho són els seus elements, de manera que hem reduït

el problema al cas univariant. Pel teorema 1.5, tenim que $D(B)$ és invertible quan les arrels de tots els polinomis en B que conformen la diagonal tenen mòdul major que 1. Atès que el determinant de $D(B)$ és el producte dels polinomis de la diagonal i coincideix amb el determinant de $\Theta(B)$, obtenim la condició que volíem. \square

Corol·lari 1.8. *Sigui $\Phi(B)Y_t = \Theta(B)u_t$ un procés estocàstic amb u_t soroll blanc gaussià, amb $\Phi(B) = I - \Phi_1 B$ i $\Theta(B) = I + \Theta_1 B$. Sigui $\bar{\lambda}_1$ el valor propi de mòdul major de Φ_1 i $\bar{\lambda}_2$ el de Θ_1 . Aleshores, $\{Y_t\}$ és invertible i estacionari si i només si $|\bar{\lambda}_1| < 1$ i $|\bar{\lambda}_2| < 1$.*

Demostració. Com que les arrels de les equacions $|\Phi(z)| = 0$ i $|\Theta(z)| = 0$ són els inversos dels valors propis de les matrius Φ_1 i Θ_1 respectivament, atès que $|\bar{\lambda}_1| < 1$ i $|\bar{\lambda}_2| < 1$, totes les arrels estan fora del cercle unitat, i per tant, el procés és invertible i estacionari. \square

Proposició 1.9. *Sigui $\Phi(B)Y_t = \Theta(B)u_t$ un procés estocàstic amb u_t soroll blanc gaussià. Considerem $f, g: \mathbb{C} \rightarrow \mathbb{C}$ tals que $f(z) = \det(\Phi(z))$ i $g(z) = \det(\Theta(z))$. Sigui $c: [0, 2\pi) \rightarrow \mathbb{C}$ la corba definida com $c(t) = \cos(t) + i\sin(t)$. Aleshores el procés és estacionari si i només si la corba tancada $f(c(t))$ no dona cap volta entorn del punt $z = 0$, i és invertible si i només si la corba tancada $g(c(t))$ no dona cap volta entorn de $z = 0$.*

Demostració. Pel principi de l'argument, el nombre de zeros de les funcions holomorfes (per ser polinòmiques) f i g en el cercle unitat coincideix amb l'índex de les funcions en la corba $c(t)$. Per tant, f i g no presenten zeros en el cercle unitat si, i només si, les corbes tancades $f(c(t))$ i $g(c(t))$ no donen cap volta ni tallen el punt $z = 0$. Pel teorema 1.7, si la funció f no té zeros en el cercle unitat el procés és estacionari i si no en té la funció g , és invertible. \square

Les matrius d'un procés VARMA han de complir determinades condicions per a què el model sigui identificable, és a dir, existeixi una equivalència bijectiva entre la representació anterior i la funció de covariàncies del procés. Per exemple, els processos:

$$\tilde{Y}_t = u_t - \Theta_1 u_{t-1}$$

i

$$\tilde{Y}_t = M u_t - \Theta_1^* u_{t-1}$$

on $u_t = M u_t$, $\Theta_1^* = \Theta_1 M^{-1}$ són indistingibles per a qualsevol matriu M no singular. És necessari, doncs, introduir condicions per a definir l'estructura ARMA vectorial de manera única. Les condicions necessàries i suficients van ser obtingudes per Hannan a [8]. En concret, el procés és identificable si:

- i. $\Phi_0 = \Theta_0 = I_k$ i Σ_u és una matriu qualsevol.
- ii. Φ_0 qualsevol, $\Theta_0 = I_k$ i Σ_u diagonal.
- iii. Θ_0 qualsevol, $\Phi_0 = I_k$ i Σ_u diagonal.
- iv. $\Sigma_u = I_k$, Φ_0 i Θ_0 qualssevol.

Aquestes formulacions difereixen en on estableixen la correlació contemporània: en el primer cas, es posa en Σ_u , en el segon, en Φ_0 , en el tercer en Θ_0 i en el darrer en ambdues, exigint també la normalització de Σ_u , $\Sigma_u = I_k$. Pel que fa aquest treball, suposarem $\Phi_0 = \Theta_0 = I_k$ i la matriu Σ_u definida positiva.

1.2.1 Els processos VAR(p)

El primer tipus de model VARMA que presentem és la generalització multivariant dels models autorregressius.

Definició 1.10. Donat un procés multivariant temporal $\{Y_t\}$, diem que és *vectorial autorregressiu* d'ordre p , VAR(p), amb constant quan es pot expressar com:

$$Y_t = v + \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + u_t$$

Sovint, però, es més còmode treballar amb els processos VAR(p) centrats, que es defineixen com:

$$Y_t - \mu = \Phi_1 (Y_{t-1} - \mu) + \dots + \Phi_p (Y_{t-p} - \mu) + u_t$$

Direm que és *estacionari* si les arrels de l'equació $|\Phi(z)| = 0$ estan fora del cercle unitat. Prenent la trasposta de l'equació del procés i multiplicant per la dreta per $Y_{t-m} - \mu$, s'obté:

$$(Y_{t-m} - \mu)(Y_t - \mu)' = (Y_{t-m} - \mu)(Y_{t-1} - \mu)' \Phi_1' + \dots + (Y_{t-m} - \mu)(Y_{t-p} - \mu)' \Phi_p' + (Y_{t-m} - \mu)u_t'$$

i prenent esperances, es conclou que les matrius de covariàncies verifiquen la relació:

$$\Gamma(m) = \sum_{i=1}^p \Gamma(m-i) \Phi_i' \quad m = 1, 2, 3, \dots$$

on $\Gamma(m) = -\Gamma'(-m)$.

Per un procés VAR, en general les matrius $\Gamma(m)$ seran diferents de zero per tot retard, i pel que fa les matrius de correlació parcial, tindrà les primeres p matrius diferents de zero i la resta nul·les.

Exemple 1.2. Considerem el procés autoregressiu d'ordre 1 definit per:

$$\begin{bmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \end{bmatrix} = \begin{bmatrix} 0.8260 & 0.0403 & -0.1977 \\ 0.9848 & -0.1209 & -0.7157 \\ 0.9320 & -0.8409 & -0.0101 \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \\ y_{t-1}^3 \end{bmatrix} + \begin{bmatrix} u_t^1 \\ u_t^2 \\ u_t^3 \end{bmatrix}$$

amb matriu de covariàncies del soroll

$$\Sigma_u = \begin{bmatrix} 2 & 0.5 & 0.25 \\ 0.5 & 3 & 0.1 \\ 0.25 & 0.1 & 1 \end{bmatrix}$$

Les primeres tres matrius d'autocorrelació simple estimades a partir d'una mostra simulada de 1000 observacions són

$$\hat{\rho}(1) = \begin{bmatrix} 0.7838 & 0.5427 & 0.3456 \\ 0.4444 & 0.0382 & -0.4808 \\ 0.1017 & -0.3891 & -0.1884 \end{bmatrix} \quad \hat{\rho}(2) = \begin{bmatrix} 0.6053 & 0.3718 & 0.1905 \\ 0.4558 & 0.6397 & 0.3873 \\ 0.0787 & 0.2770 & 0.4980 \end{bmatrix}$$

$$\hat{\rho}(3) = \begin{bmatrix} 0.4615 & 0.3316 & 0.2005 \\ 0.3246 & 0.0252 & -0.2130 \\ -0.0298 & -0.2753 & -0.2055 \end{bmatrix}$$

Observem que totes són clarament no nul·les i tenen pràcticament tots els seus coeficients significativament no nuls. En efecte, el valor crític per al test de significació per als coeficients amb un nivell de confiança de 0.95 és 0.062, és a dir, els coeficients es consideren significativament no nuls si prenen valors, en mòdul, superiors al de referència.

Les matrius d'autocorrelació parcial estimades són

$$\hat{P}(1) = \begin{bmatrix} 0.8031 & 0.0528 & -0.1804 \\ 0.9891 & -0.1699 & -0.6891 \\ 0.9308 & -0.8401 & 0.0004 \end{bmatrix} \quad \hat{P}(2) = \begin{bmatrix} -0.0253 & 0.0612 & -0.0334 \\ 0.0631 & -0.0201 & 0.0232 \\ 0.0183 & 0.0176 & -0.0061 \end{bmatrix}$$

i a partir de la segona ja es consideren nul·les. En efecte, el p.valor del test de significació per a $P(2)$ (veure secció 1.1.2), que contrasta si la matriu es pot considerar nula o no, és 0.2179. Per tant, es pot acceptar amb contundència la hipòtesi que $P(2) = 0$.

1.2.2 Els processos VMA(q)

Definició 1.11. Un procés $\{Y_t\}$ segueix un model de mitjanes mòbils vectorial d'ordre q quan:

$$Y_t = (I_k + \Theta_1 B + \dots + \Theta_q B^q)u_t$$

La condició d'*invertibilitat* és que els zeros de l'equació $|I_k + \Theta_1 z + \dots + \Theta_q z^q| = 0$ estiguin fora del cercle unitat. Les matrius d'autocovariàncies es calculen com:

$$\begin{aligned} \Gamma(m) &= E[(Y_t - \mu)(Y_{t+m} - \mu)'] \\ &= E[(u_t + \Theta_1 u_{t-1} + \dots + \Theta_q u_{t-q})(u'_{t+m} + u'_{t+m-1}\Theta'_1 + \dots + u'_{t+m-q}\Theta'_q)] \end{aligned}$$

que condueix al resultat:

$$\begin{aligned} \Gamma(m) &= \Sigma_u \Theta'_m + \Theta_1 \Sigma_u \Theta'_{m+1} + \dots + \Theta_{q-m} \Sigma_u \Theta'_q \quad m \leq q \\ \Gamma(m) &= 0 \quad m > q \end{aligned}$$

Per tant, un procés VMA només té q matrius d'autocovariàncies diferents de zero. Com passa en el cas univariant, un procés MA multivariant invertible és equivalent a un VAR(∞) i per això les seves matrius d'autocorrelació parcial, $P(m)$, seran diferents de zero per a qualsevol retard.

Exemple 1.3. Considerem el procés de mitjanes mòbils d'ordre 1 definit per:

$$\begin{bmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \end{bmatrix} = \begin{bmatrix} 0.8260 & 0.0403 & -0.1977 \\ 0.9848 & -0.1209 & -0.7157 \\ 0.9320 & -0.8409 & -0.0101 \end{bmatrix} \begin{bmatrix} u_{t-1}^1 \\ u_{t-1}^2 \\ u_{t-1}^3 \end{bmatrix} + \begin{bmatrix} u_t^1 \\ u_t^2 \\ u_t^3 \end{bmatrix}$$

amb matriu de covariàncies del soroll

$$\Sigma_u = \begin{bmatrix} 2 & 0.5 & 0.25 \\ 0.5 & 3 & 0.1 \\ 0.25 & 0.1 & 1 \end{bmatrix}$$

Les primeres matrius d'autocorrelació simple estimades per a una mostra simulada de 1000 observacions

són

$$\hat{\rho}(1) = \begin{bmatrix} 0.5029 & 0.4148 & 0.3715 \\ 0.1975 & 0.0583 & -0.4453 \\ -0.0116 & -0.1223 & 0.0517 \end{bmatrix} \quad \hat{\rho}(2) = \begin{bmatrix} 0.0662 & 0.0595 & 0.0342 \\ 0.0914 & 0.0595 & 0.0196 \\ 0.0118 & 0.0247 & -0.0057 \end{bmatrix}$$

$$\hat{\rho}(3) = \begin{bmatrix} 0.0615 & 0.0171 & 0.0093 \\ 0.0482 & -0.0000 & 0.0259 \\ 0.0080 & -0.0239 & -0.0158 \end{bmatrix}$$

La primera matriu és clarament no nula. A partir de la segona, però, ja es poden considerar nul·les. En efecte, tots els termes de la segona matriu llevat de 0.0662 i 0.0914 es poden considerar nuls ja que el valor de referència és 0.062 (veure exemple 1.2), i aquests dos són molt petits. De fet, el determinant de la matriu és molt proper a zero, $4.312 \cdot 10^{-5}$, de manera que reforça l'argument de poder-la considerar nul·la. En el cas de la tercera, directament tots els seus termes són no significatius estadísticament.

Pel que fa les matrius d'autocorrelació parcial, són totes no nul·les, i les dels primers tres retards són:

$$\hat{P}(1) = \begin{bmatrix} 0.5770 & -0.0291 & -0.1678 \\ 0.6772 & -0.1464 & -0.2810 \\ 0.8378 & -0.7751 & 0.0765 \end{bmatrix} \quad \hat{P}(2) = \begin{bmatrix} -0.2552 & -0.0127 & 0.1067 \\ 0.1751 & -0.4102 & 0.1397 \\ 0.0564 & -0.0523 & -0.1052 \end{bmatrix}$$

$$\hat{P}(3) = \begin{bmatrix} 0.1917 & 0.0423 & -0.0923 \\ 0.2223 & -0.0516 & -0.1544 \\ 0.3289 & -0.2818 & 0.0041 \end{bmatrix}$$

Els p.valors per al test de significació de $P(1)$, $P(2)$ i $P(3)$ (veure secció 1.1.2) són menors a 10^{-8} , de manera que clarament s'han de considerar diferents de la matriu nul·la.

1.2.3 Els processos VARMA(p,q)

En les dues seccions anteriors s'ha presentat com les propietats dels models ARMA univariants es generalitzen de forma elegant i natural per a les representacions multivariants. Així, de la mateixa manera que els processos AR(p) s'identifiquen per tenir només els primers p retards de la funció d'autocorrelació parcial significativament diferents de zero, els processos VAR(p) s'identifiquen per tenir només les primeres p matrius d'autocorrelació parcial no nul·les. Anàlogament, els processos VMA(q) presenten les primeres q matrius d'autocorrelació simple diferents de zero i les altres nul·les així com els MA(q) només tenen els primers q retards de la funció d'autocorrelació significatius. L'objectiu d'aquesta secció és presentar els models mixtes, els anomenats models VARMA. Veurem que com en el cas dels ARMA, l'estructura de

les matrius d'autocorrelació pot ser molt variada i no presenta un tall (en el sentit de passar a ser nul·les) en l'ordre del procés.

Definició 1.12. Definim un procés VARMA(p,q) com:

$$Y_t = v + \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + u_t + \Theta_1 u_{t-1} + \dots + \Theta_q u_{t-q}$$

on u_t és soroll blanc, de mitjana zero, amb matrius de covariàncies no singular Σ_u .

El procés és *estable* o *estacionari* si la part VAR(p) és estable, és a dir, si les arrels de l'equació en z $|I_k - \Phi_1 z - \dots - \Phi_p z^p| = 0$ estan fora del cercle unitat. El procés és *invertible* quan ho és la seva part de mitjanes mòbils, és a dir, quan les arrels de l'equació $|I_k + \Theta_1 z + \dots + \Theta_q z^q| = 0$ tenen mòdul major que 1.

La representació pura VMA i pura VAR dels processos VARMA

Notant per $\Phi(B) = I_k - \Phi_1 B - \dots - \Phi_p B^p$ i per $\Theta(B) = I_k + \Theta_1 B + \dots + \Theta_q B^q$, podem reescriure el procés VARMA(p,q) com

$$\Phi(B)Y_t = v + \Theta(B)u_t$$

Aleshores, si el procés és estable el podem expressar com un VMA d'ordre infinit:

$$Y_t = \Phi(1)^{-1}v + \Phi(B)^{-1}\Theta(B)u_t = \mu + \sum_{i=0}^{\infty} \Psi_i u_{t-i}$$

on μ és la mitjana.

Multiplicant per l'esquerra per $\Phi(B)$ i restant v resulta

$$\begin{aligned} \Phi(B)Y_t - v &= \Theta(B)u_t \\ &= (I_k + \Theta_1 B + \dots + \Theta_q B^q)u_t \\ &= \Phi(B)\left(\sum_{i=0}^{\infty} \Psi_i u_{t-i}\right) \\ &= \left(I_k + \sum_{i=1}^{\infty} \left[\Psi_i - \sum_{j=1}^i \Phi_j \Psi_{i-j}\right] B^i\right)u_t \end{aligned}$$

d'on, comparant coeficients, obtenim

$$\Theta_i = \Psi_i - \sum_{j=1}^i \Phi_j \Psi_{i-j} \quad i = 1, 2, \dots$$

amb $\Psi_0 := I_k$, $\Phi_j := 0$ per $j > p$ i $\Theta_i := 0$ per $i > q$. Arribem finalment a

$$\Psi_i = \Theta_i + \sum_{j=1}^i \Phi_j \Psi_{i-j} \quad i = 1, 2, \dots$$

D'aquesta manera podem obtenir, recursivament, els coeficients de la presentació en pur VMA.

Quan el procés és invertible en podem obtenir una representació com a pur VAR:

$$Y_t - \sum_{i=1}^{\infty} \Pi_i Y_{t-i} = \Theta(B)^{-1} \Phi(B) Y_t = \Theta(1)^{-1} v + u_t$$

on els Π_i són les matrius de la representació VAR pura, que es troben comparant coeficient a coeficient la igualtat anterior respecte l'operador retard B :

$$I_k - \sum_{i=1}^{\infty} \Pi_i B^i = \Theta(B)^{-1} \Phi(B)$$

Alternativament, podem raonar de manera anàloga al cas anterior i multiplicar aquesta expressió per $\Theta(B)$, obtenint:

$$\begin{aligned} (I_k + \Theta_1 B + \dots + \Theta_q B^q) \left[I_k - \sum_{i=1}^{\infty} \Pi_i B^i \right] &= I_k + \sum_{i=1}^{\infty} \left[\Theta_i - \sum_{j=1}^i \Theta_{i-j} \Pi_j \right] B^i \\ &= I_k - \Phi_1 B - \dots - \Phi_p B^p \end{aligned}$$

on $\Theta_0 := I_k$, $\Theta_i := 0$ per $i > q$ i $\Phi_i := 0$ per $i > p$. Comparant matrius, arribem a

$$\Pi_i = \Phi_i + \Theta_i - \sum_{j=1}^{i-1} \Theta_{i-j} \Pi_j \quad i = 1, 2, \dots$$

amb la convenció que la suma val 0 quan el subíndex inferior és major que el superior.

Per exemple, per al procés VARMA(1,1) definit com

$$Y_t = \Phi_1 Y_{t-1} + u_t + \Theta_1 u_{t-1}$$

obtenim la representació VAR pura a partir de les matrius:

$$\begin{aligned} \Pi_1 &= \Phi_1 + \Theta_1 \\ \Pi_2 &= \Phi_2 + \Theta_2 - \Theta_1 \Pi_1 = -\Theta_1 \Phi_1 - \Theta_1^2 \\ &\vdots \\ \Pi_i &= (-1)^{i-1} (\Theta_1^i + \Theta_1^{i-1} \Phi_1) \end{aligned}$$

i la representació VMA pura:

$$\begin{aligned}
\Psi_0 &= 0 \\
\Psi_1 &= \Theta_1 + \Phi_1 \\
\Psi_2 &= \Theta_2 + \Phi_1\Psi_1 + \Phi_2\Psi_0 = \Phi_1(\Theta_1 + \Phi_1) \\
&\vdots \\
\Psi_i &= \Phi_1^{i-1}\Theta_1 + \Phi_1^i
\end{aligned}$$

Matrius d'autocorrelació

Per a calcular la funció d'autocorrelació per al procés estacionari VARMA general (centrat):

$$Y_t = \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + u_t + \Theta_1 u_{t-1} + \dots + \Theta_q u_{t-q}$$

trasposem la seva expressió, la premultipliquem per Y_{t-m} i prenem esperances. Obtenim així les matrius de covariàncies creuades:

$$\begin{aligned}
\Gamma(m) &= E[Y_{t-m} Y_t'] \\
&= E[(\Phi_1 Y_{t-m-1} + \dots + \Phi_p Y_{t-m-p} + u_{t-m} + \Theta_1 u_{t-m-1} + \dots + \Theta_q u_{t-m-q}) \times \\
&\quad (Y'_{t-1} \Phi'_1 + \dots + \Phi'_p Y_{t-p} + u'_t + u'_{t-1} \Theta'_1 + \dots + u'_{t-q} \Theta'_q)] \\
&= \sum_{i=1}^p E[\Phi_1 Y_{t-m-i} Y'_{t-i} \Phi'_i] - \sum_{j=0}^q E[Y_{t-m} u'_{t-j} \Theta'_j] \\
&= \sum_{i=1}^p \Gamma(m-i) \Phi'_i - \sum_{j=0}^q E[Y_{t-m} u'_{t-j} \Theta'_j]
\end{aligned}$$

Per a calcular el segon membre, observem que en ser un procés estacionari, podem escriure $Y_t = \Phi(B)^{-1} \Theta(B) u_t = \Psi(B) u_t$, que implica:

$$E(Y_{t-m} u'_{t-j}) = E(\Psi(B) u_{t-m} u'_{t-j}) = \begin{cases} 0 & j < m \\ \Sigma_u & j = m \\ \Psi_{j-m} \Sigma_u & j > m \end{cases}$$

D'on finalment obtenim:

$$\begin{aligned}
\Gamma(m) &= \sum_{i=1}^p \Gamma(m-i) \Phi'_i - \sum_{j=1}^q \Psi_{j-m} \Sigma_u \Theta'_j \quad 0 \leq m \leq q \\
\Gamma(m) &= \sum_{i=1}^p \Gamma(m-i) \Phi'_i \quad m > q
\end{aligned}$$

Les matrius d'autocorrelacions simples es calculen a partir de la matriu $V = \text{Diag}\{\gamma_{11}(0)^{1/2}, \dots, \gamma_{kk}(0)^{1/2}\}$, com $\rho(m) = V^{-1}\Gamma(m)V^{-1}$.

1.3 Construcció de models VARMA per a sèries estacionàries

Donat un vector de K sèries estacionàries Y_t , es vol construir un model VARMA que en descriu l'estructura de dependències dinàmiques entre elles. Suposarem que, a priori, no especifiquem si algunes de les variables són exògenes, de manera que permetem relacions entre totes elles. El mètode per a construir models consta de les etapes següents:

1. Identificar l'estructura del model, és a dir, l'ordre del model VARMA estacionari.
2. Estimar els paràmetres (veure capítol 2) i seleccionar el millor model entre els estimats.
3. Realitzar contrastos diagnòstics per comprovar si el model seleccionat no té deficiències detectables.

1.3.1 Identificació de l'estructura del model

En aquesta primera etapa, fonamental, determinarem un conjunt de models possibles que després estimarem. En primer lloc, cal trobar el nombre K de sèries *independents* que tenim, en el sentit de sèries que no s'han construït exclusivament com a combinació lineal de les altres. Per exemple, si tenim un sistema de tres sèries y_t^1, y_t^2 i y_t^3 tal que la primera són els beneficis, la segona els ingressos i la tercera els costos, aleshores és clar que $y_t^1 = y_t^2 - y_t^3$, de manera que no tenim tres sèries independents sinó dues. Suposem que disposem de Q sèries tals que existeixen H relacions exactes contemporànies representades pels coeficients del vector h_i entre elles:

$$(Y_t - \bar{Y})'h_i = 0$$

Aleshores tindrem únicament $K = Q - H$ components independents. En general, quan disposem de poques sèries aquesta situació és poc freqüent, però quan es treballa amb un grup gran de sèries sovint ens trobem combinacions lineals que són gairebé constants o amb molt poca variabilitat.

Per a contrastar aquest punt, començarem calculant la matriu $\hat{\Gamma}(0) = \frac{1}{T} \sum_{i=1}^T (Y_t - \bar{Y})(Y_t - \bar{Y})'$. Tindrem que

$$\hat{\Gamma}(0)h_i = \frac{1}{T} \sum_{i=1}^T (Y_t - \bar{Y})(Y_t - \bar{Y})'h_i = 0$$

fet que implica que la matriu $\hat{\Gamma}(0)$ ha de tenir H valors propis iguals a zero. Si la relació és $(Y_t - \bar{Y})'h_i = w_i$, per a H sèries w_i té mitjana zero i variància propera a zero, tindrem que la matriu $\hat{\Gamma}(0)$ ha de tenir H

valors propis molt propers a zero. En lloc de modelar el vector Y_t de Q sèries podem prendre els vectors i valors propis de la matriu $\hat{\Gamma}(0)$ i modelar les $K = Q - H$ combinacions lineals que tenen alta variància. Això equival a realitzar components principals de dades i quedar-nos amb les K components d'alta variància. És important assenyalar que quan les components de les sèries no van en els mateixos ordres de magnitud, convé estandarditzar-les dividint cada sèries per la seva desviació típica. Això equival a obtenir les components de la matriu de correlacions estimada $\rho(0)$, enlloc de la de covariàncies.

En segon lloc, un cop disposem de K sèries *independents* a modelar, cal determinar l'estructura dels operadors $\Phi(B)$ i $\Theta(B)$, a partir de les matrius d'autocorrelació simple i parcial. Si l'ordre p de les matrius d'autocorrelació simple és tal que per a $n > p$ les matrius $\hat{\rho}(n)$ es poden considerar nul·les (tals que tots els seus elements es puguin considerar zero) i n és elevat (major que 3 o 4), aleshores es pot considerar un procés VAR(p). Cas que p sigui baix (1, 2 o 3), la hipòtesi més simple és que el procés és un VMA(p). Les matrius de correlació parcial tenen la propietat oposada. Si n'hi ha poques diferents de zero, ens indica un procés VAR. Si pel contrari hi ha matrius no nul·les fins a retards elevats, això suggereix una estructura VMA. Finalment, una informació adicional important per a la detecció de l'estructura del model multivariant és l'estudi de les matrius d'autocorrelació simple dels residus estimats en cada etapa, que poden indicar l'existència d'un model mixte VARMA. En efecte, si per exemple suposem que el model vertader és un VARMA(1,1)

$$(I - \Phi_1 B)Y_t = (I + \Theta_1 B)u_t$$

i erròniament l'intentem modelar per un VAR(1)

$$(I - \hat{\Phi}_1 B)Y_t = \varepsilon_t$$

aleshores els residus ε_t tindran estructura VARMA, indicant la possible existència d'un model mixte. En efecte, tenim que

$$\varepsilon_t = (I + \hat{\Theta}_1 B)u_t$$

i per tant,

$$\begin{aligned} cov[\varepsilon_t, \varepsilon_{t-1}] &= E[u_t + \hat{\Theta}_1 u_{t-1}, u_{t-1} + \hat{\Theta}_1 u_{t-2}] \\ &= E[u_t u_{t-1}] + E[\hat{\Theta}_1 u_t u_{t-2}] + E[\hat{\Theta}_1 u_{t-1} u_{t-1}] + E[\hat{\Theta}_1^2 u_{t-1} u_{t-2}] \\ &= E[\hat{\Theta}_1 u_{t-1} u_{t-1}] = \hat{\Theta}_1 \Sigma_u \hat{\Theta}_1' \end{aligned}$$

on hem utilitzat que $cov[u_t, u_{t-j}] = 0$ per $j \neq 0$ ja que u_t és soroll blanc.

1.3.2 Estimació i selecció de models

Atès que el capítol 2 té per objecte l'estimació dels paràmetres del model, només tractarem aquí la selecció del millor model entre uns quants candidats. Alhora d'escollir un bon model hem de tenir present que perseguim explicar la realitat d'una manera prou precisa (versemblança) i senzilla (parsimònia). Per això, per triar entre diferents models, típicament s'utilitza la minimització del criteri d'informació d'Akaike (AIC) o el bayesià (BIC). Donat un model VARMA(p,q), la definició dels criteris suara esmentats és:

$$AIC = \log|\hat{\Sigma}_u| + \frac{2}{T}(p+q)K^2$$

$$BIC = \log|\hat{\Sigma}_u| + \frac{1}{T}(p+q)K^2 \log T$$

Observem que aquests criteris mesuren les dues qualitats d'un bon model. D'una banda, i com veurem amb profunditat en el capítol 2, $\log|\hat{\Sigma}_u|$ és proporcional al logarisme de la versemblança canviat de signe, de manera que ens interessa que sigui quan més petit millor. D'altra banda, el segon sumand és en tots dos criteris directament proporcional al nombre de paràmetres estimats ($(p+q)K^2$), ajustat segons la mida T de la mostra.

Cal destacar que aquests criteris permeten escollir entre diferents models, però no ens diuen que el model millor entre els avaluats sigui un *bon* model, en el sentit que reculli l'estructura essencial del procés i que tingui una bona capacitat predictiva. En aquest sentit, és molt important l'etapa següent de diagnosi i avaluar la capacitat predictiva del model: es construeix el model deixant unes quantes de les darreres observacions sense utilitzar, i es comprova com de bones haurien estat les seves previsions.

1.3.3 Diagnosi

Els tests diagnòstics del model estan dissenyats per a identificar possibles errors en la seva formulació, i consisteixen en l'anàlisi dels residus estimats \hat{u}_t , per a comprovar si segueixen un procés de soroll blanc. Això implica que:

1. Cada component u_t^i ha de ser soroll blanc gaussià. Per a veure-ho, comprovarem que les funcions d'autocorrelació simple i parcial no tenen retards significatius, i analitzarem per mitjà de contrastos com el test d'ajustament khi quadrat o el test de Kolmogorov-Smirnov si les components del soroll s'ajusten a una distribució normal.
2. Conjuntament, el vector u_t només ha de tenir correlació contemporània per tal de ser soroll blanc. Per tant, les matrius $\hat{\rho}_{u_t}(l)$ han de tenir tots els elements no significativament diferents de zero per a $l \geq 1$ i les matrius i $\hat{P}_{u_t}(l)$ s'han de poder considerar nul·les.

Cas que observéssim que el procés \hat{u}_t té estructura dinàmica

$$\hat{u}_t = \Psi(B)u_t$$

reformularíem el model mitjançant

$$\Phi(B)Y_t = \Theta(B)u_t = \Theta(B)\Psi(B)u_t$$

i tornariem a l'etapa d'estimació.

1.4 Previsió

Donada una realització d'un procés vectorial ARMA al llarg del temps, $\{Y_s, s \leq t\}$, considerem el problema de predir els valors futurs Y_{t+l} , $l = 1, 2, \dots$. En aquesta secció de desenvolupament teòric, assumirem que coneixem i disposem del model exacte del procés $\{Y_t\}$, incloent els paràmetres. En el capítol 2 veurem que per a mostres prou grans, els errors d'estimació dels paràmetres a partir de les dades disponibles no desvien de manera important les previsions.

1.4.1 Predictor amb Error Quadràtic Mitjà Mínim

Abans de detallar els mètodes de previsió de models VARMA, fem un breu recordatori dels principis bàsics de les previsions en un context general. Signin Y i X dos vectors aleatoris k -dimensional i n -dimensional respectivament, i suposem que volem predir valors desconeguts de Y a partir del vector X , és a dir, volem estimar $\hat{Y} = g(X)$. La matriu d'error quadràtic mitjà del predictor \hat{Y} és

$$EQM = E[(Y - \hat{Y})(Y - \hat{Y})'|X]$$

Definició 1.13. Anomenem *predictor de mínim error quadràtic mig* (EQMM) de Y a la funció $\hat{Y} = g^*(X)$ que minimitza

$$E[(b'Y - b'\hat{Y})^2|X] = b'E[(Y - \hat{Y})(Y - \hat{Y})'|X]b = b'EQMb$$

on b és un vector k -dimensional qualsevol de constants.

Lema 1.14. Donats Y i X dos vectors aleatoris k -dimensional i n -dimensional respectivament, el predictor de EQMM de Y en funció de X és

$$\hat{Y} = g^*(X) = E[Y|X]$$

Demostració. Sumem i restem $E[Y|X]$ a l'expressió de l'error quadràtic mig i la desenvolupem:

$$\begin{aligned}
EQM &= E[(Y - \hat{Y} + E[Y|X] - E[Y|X])(Y - \hat{Y} + E[Y|X] - E[Y|X])'|X] \\
&= E[YY' - Y\hat{Y}' + YE[Y|X]' - YE[Y|X]' - \hat{Y}Y' + \hat{Y}\hat{Y}' - \hat{Y}E[Y|X]' + \hat{Y}E[Y|X]' + \dots \\
&\dots + E[Y|X]Y' - E[Y|X]\hat{Y}' + E[Y|X]E[Y|X]' - E[Y|X]Y' + \dots \\
&\dots + E[Y|X]\hat{Y}' - E[Y|X]E[Y|X]' + E[Y|X]E[Y|X]'|X] \\
&= E[(Y - E[Y|X])(Y - E[Y|X])'|X] + E[(E[Y|X] - \hat{Y})(E[Y|X] - \hat{Y})'|X] \\
&= VAR(Y|X) + E[(E[Y|X] - \hat{Y})(E[Y|X] - \hat{Y})'|X]
\end{aligned}$$

on l'última expressió s'obté a partir de la linealitat de l'esperança i del fet que $E[Y|X]$ és constant i per tant surt fora de les esperances.

Observem que el primer terme no depèn del predictor i per tant, hem de minimitzar el segon. I resulta que el podem fer zero prenent $\hat{Y} = E[Y|X]$. \square

L'error de predicció del predictor amb EQMM és $e = Y - E[Y|X]$.

Quan restringim les funcions $g(X)$ a funcions lineals, parlem de *predictors lineals* i de *predictors lineals òptims*.

1.4.2 Previsió puntual i per intervals i regions de confiança per als models VARMA

Suposem que tenim un procés K -dimensional de mitjana zero VARMA(p,q)

$$Y_t = \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + u_t + \Theta_1 u_{t-1} + \dots + \Theta_q u_{t-q}$$

estable i invertible. Aleshores, com hem vist en la secció 1.2.3 el procés té una representació pura VAR

$$Y_t = \sum_{i=1}^{\infty} \Pi_i Y_{t-i}$$

i pura VMA

$$Y_t = \sum_{i=0}^{\infty} \Psi_i u_{t-i}$$

Assumint que u_t és soroll blanc independent i aplicant el predictor òptim, és a dir, esperances condicionades per la informació fins a temps t , E_t , obtenim que en el pas h la previsió puntual òptima és:

$$Y_t(h) = \begin{cases} \Phi_1 Y_t(h-1) + \dots + \Phi_p Y_t(h-p) + \Theta_h u_t + \dots + \Theta_q u_{t+h-q} & h \leq q \\ \Phi_1 Y_t(h-1) + \dots + \Phi_p Y_t(h-p) & h > q \end{cases}$$

on notem $Y_t(j) := Y_{t+j}$ per a $j \leq 0$. Anàlogament, a partir de la representació $\text{VAR}(\infty)$, obtenim

$$Y_t(h) = \sum_{i=1}^{\infty} \Pi_i Y_t(h-i)$$

i de la $\text{VMA}(\infty)$

$$Y_t(h) = \sum_{i=h}^{\infty} \Psi_i u_{t+h-i} = \sum_{i=0}^{\infty} \Psi_{h+i} u_{t-i}$$

Totes tres expressions són equivalents, són predictors lineals òptims (que és també el predictor òptim per al model, atès que és lineal respecte les components temporals).

Dins aquest marc teòric, podem trobar també la millor estimació de les matrius de l'EQM. En efecte, l'expressió de l'error de previsió a l'instant $t+h$ és

$$e_t(h) = Y_{t+h} - Y_t(h) = \sum_{i=0}^{h-1} \Psi_i u_{t+h-i}$$

i per tant la previsió de la matriu de l'EQM és

$$\Sigma_Y(h) = E[e_t(h)e_t(h)'] = \sum_{i=0}^{h-1} \Psi_i \Sigma_u \Psi_i'$$

Atès que l'esperança de l'error de previsió és zero, no hi ha esbiaixament, la matriu de l'EQM estimada és la matriu de covariàncies de l'error previst. En conseqüència, si el procés és gaussià, i.e., per a tot t i h , Y_t, \dots, Y_{t+h} tenen una distribució conjunta multinormal i els u_t estan distribuïts amb normalitat, aleshores

$$e_t(h) \sim N(0, \Sigma_Y(h))$$

fet que ens permet construir intervals i regions de confiança per a les previsions. En efecte, aquest resultat implica que els errors de previsió de les components individuals són normals, de manera que notant per $\sigma_i(h)$ l'arrel quadrada de l'element i -èssim de la diagonal de $\Sigma_Y(h)$, tenim

$$\frac{Y_t^i(h) - Y_{t+h}^i}{\sigma_i(h)} \sim N(0, 1)$$

Finalment, notant per $Z \sim N(0, 1)$, si z_α és el valor tal que $P(Z > z_\alpha) = \alpha$, aleshores l'interval de confiança amb significació α per a la variable i -èssima és

$$[Y_t^i(h) - z_{\alpha/2}\sigma_i(h), Y_t^i(h) + z_{\alpha/2}\sigma_i(h)]$$

Podem estendre aquest raonament per a construir regions de confiança per a un conjunt de variables del model. Per exemple, si es vol una regió de previsió per a les primeres n variables del model VARMA, definim la matriu $n \times K$ -dimensional $F := [I_n \ 0]$ i recordem que

$$[Y_{t+h} - Y_t(h)]' F (F \Sigma_Y(h) F')^{-1} [Y_{t+h} - Y_t(h)] \sim \chi^2(n)$$

Aleshores, la distribució $\chi^2(n)$ es pot utilitzar per a determinar un $(1 - \alpha)100\%$ elipsoide de previsions per a les primeres n components.

A la pràctica, atès que construir-la pot ser molt difícil, s'opta per definir una regió de confiança a partir de diferents intervals de confiança. Tenint present que donats uns esdeveniments E_1, \dots, E_m , es té que $P[\bigcap_{i=1}^m E_i] \geq 1 - \sum_{i=1}^m P(\bar{E}_i)$, on \bar{E}_i és el complementari de E_i , aleshores es prenen intervals H_i de confiança de significació α/n , i s'obté

$$P(Y_{t+h} \in H_1 \times \dots \times H_n) \geq 1 - \sum_{i=1}^n P(Y_{t+h}^i \notin H_i) = 1 - \alpha$$

Per a processos VARMA que tenen mitjana μ qualsevol, només cal sumar la mitjana a les previsions del procés centrat, i la previsió de la matriu de l'EQM no queda alterada. És a dir, donat el procés de mitjana zero Y_t i $X_t = Y_t + \mu$, aleshores $X_t(h) = Y_t(h) + \mu$ i $\Sigma_X(h) = \Sigma_Y(h)$.

1.5 Models VARMA estacionals multiplicatius

Sovint, per exemple en sèries relacionades amb l'agricultura, la climatologia, l'economia o el turisme, ens trobem que contenen un patró estacional que es va repetint al llarg del temps. El període temporal s més petit per al qual es va repetint el patró s'anomena període estacional. Aquests patrons poden aparèixer també quan es modelitza un conjunt de sèries alhora. Per això, treballs com [20], [17]. [11] han extès el concepte de models estacionals multiplicatius ARMA per a processos vectorials.

Definició 1.15. Donat $\{Y_t\}$ un procés vectorial de dimensió K , diem que segueix un model estacional multiplicatiu de període s , $VARMA(p)(P)_s(q)(Q)_s$, si s'expressa com

$$\phi(B)\Phi(B^s)Y_t = \Theta(B)\Theta(B^s)u_t$$

on

$$\begin{aligned}\phi(B) &= I_K - \phi_1 B - \dots - \phi_p B^p \\ \Phi(B^s) &= I_K - \Phi_1 B^s - \dots - \Phi_P B^{Ps} \\ \theta(B) &= I_K + \theta_1 B + \dots + \theta_q B^q \\ \Theta(B^s) &= I_K + \Theta_1 B^s + \dots + \Theta_Q B^{Qs}\end{aligned}$$

i u_t és soroll blanc vectorial de mitjana zero.

Notem que a diferència dels models ARMA vectorials, aquesta expressió no és única, atès que el producte de matrius no és commutatiu i per tant l'ordre dels operadors determina el model resultant. És a dir, el procés VARMA(p)(P) $_s$ (q)(Q) $_s$ definit com $\phi(B)\Phi(B^s)Y_t = \Theta(B)\Theta(B^s)u_t$ no és, en general, el mateix que el procés VARMA(P) $_s$ (p)(Q) $_s$ (q) definit com $\Phi(B^s)\phi(B)Y_t = \Theta(B^s)\theta(B)u_t$. Per tant, alhora de detallar un model VARMA estacional multiplicatiu cal especificar l'ordre dels operadors.

Proposició 1.16. *Sigui $\{Y_t\}$ un model VARMA estacional multiplicatiu de període s i ordre p, P, q i Q . El procés és estacionari si, i només si, les arrels de les equacions $|\phi(B)| = 0$ i $|\Phi(B^s)| = 0$ tenen mòdul major que 1. El procés és invertible si, i només si, les arrels de les equacions $|\theta(B)| = 0$ i $|\Theta(B^s)| = 0$ tenen mòdul major que 1.*

Demostració. El resultat és immediat en virtut del teorema 1.7 i atès que el determinant és multiplicatiu. \square

Notem que l'estacionarietat i la invertibilitat d'aquests models no depèn de l'ordre dels operadors.

Destaquem també que el problema de l'ordre dels operadors no és solament un problema de representació, sinó de model, ja que les matrius que defineixen el procés seran molt diferents si es canvia l'ordre de les representacions.

Exemple 1.4. Per a veure la diferència que suposa l'ordre dels operadors, considerem un model VAR amb $p = 1$, $P = 1$ i $s = 12$, i amb matrius

$$\phi_1 = \begin{bmatrix} 0.5 & -0.5 \\ 0.7 & -0.2 \end{bmatrix} \quad \Phi_1 = \begin{bmatrix} 0.3 & 0.3 \\ -0.25 & 0.5 \end{bmatrix}$$

i el vector u_t soroll blanc de mitjana zero amb una certa matriu Σ_u de covariàncies.

Si la representació que prenem per al model és VAR(1)(1) $_{12}$, aleshores obtenim:

$$\begin{bmatrix} Y_t^1 \\ Y_t^2 \end{bmatrix} = \begin{bmatrix} 0.5 & -0.5 \\ 0.7 & -0.2 \end{bmatrix} \begin{bmatrix} Y_{t-1}^1 \\ Y_{t-1}^2 \end{bmatrix} + \begin{bmatrix} 0.3 & 0.3 \\ -0.25 & 0.5 \end{bmatrix} \begin{bmatrix} Y_{t-12}^1 \\ Y_{t-12}^2 \end{bmatrix} + \begin{bmatrix} 0.275 & -0.1 \\ 0.26 & 0.11 \end{bmatrix} \begin{bmatrix} Y_{t-13}^1 \\ Y_{t-13}^2 \end{bmatrix} + \begin{bmatrix} u_t^1 \\ u_t^2 \end{bmatrix}$$

mentre que si el representem com un VAR(1) $_{12}$ (1):

$$\begin{bmatrix} Y_t^1 \\ Y_t^2 \end{bmatrix} = \begin{bmatrix} 0.5 & -0.5 \\ 0.7 & -0.2 \end{bmatrix} \begin{bmatrix} Y_{t-1}^1 \\ Y_{t-1}^2 \end{bmatrix} + \begin{bmatrix} 0.3 & 0.3 \\ -0.25 & 0.5 \end{bmatrix} \begin{bmatrix} Y_{t-12}^1 \\ Y_{t-12}^2 \end{bmatrix} + \begin{bmatrix} 0.36 & -0.21 \\ 0.225 & 0.025 \end{bmatrix} \begin{bmatrix} Y_{t-13}^1 \\ Y_{t-13}^2 \end{bmatrix} + \begin{bmatrix} u_t^1 \\ u_t^2 \end{bmatrix}$$

1.5.1 Selecció del millor model estacional multiplicatiu

Pel que acabem de veure, l'ordre dels operadors pot canviar notablement els paràmetres del model en els factors multiplicatius. En la pràctica, l'estimació serà també diferent en les matrius que defineixen $\phi(B)$,

$\Phi(B^s)$, $\theta(B)$ i $\Theta(B^s)$. Com que a priori no coneixem el model subjacent que defineix el vector a modelar, ni, evidentment, els seus paràmetres, brolla de manera natural la pregunta de com seleccionar la millor representació per a un VARMA estacional multiplicatiu.

Criteri d'Informació d'Akaike

El criteri més àmpliament adoptat per a la tria entre diferents models és l'AIC normalitzat. Donada una mostra de mida T i model VARMA d'ordre p, P, q i Q , i període s , es defineix l'AIC normalitzat com:

$$AIC = \log(|\Sigma_u|) + \frac{2K^2(p + P + q + Q)}{T}$$

Atès que la selecció de l'ordre dels operadors no influeix en els paràmetres p, P, q i Q , la comparació dels valors de l'AIC dels diferents models és purament una comparació de la seva versemblança, representada pel terme $\log(|\Sigma_u|)$.

Criteri d'ajustament del soroll

Encara que el criteri de l'AIC és molt emprat quan cal comparar models, depenen de l'estructura dels processos i del nombre d'observacions per no ser molt adequat [20] Per això, en aquesta secció presentem un altre criteri que pugui ajudar a millorar la selecció de models.

Els sorolls u_t dels processos VARMA estan absolutament condicionats pels operadors. En el cas dels VARMA estacionals, el canvi en l'ordre dels operadors porta a matrius que poden ser molt diferents, com hem vist en l'exemple 2.3. Si observéssim els sorolls veritables del procés, aleshores podríem comparar-los amb els que estima la modelització estacional i ens quedaríem amb el model estacional que fes l'estimació més correcta. La pregunta que ens queda per respondre és, doncs, com podem obtenir aquests sorolls vertaders.

Com que treballem amb la hipòtesi que els processos són estacionaris i invertibles, admeten una representació de l'estil

$$Y_t = u_t + \sum_{i=1}^{\infty} \Pi_i Y_{t-i}$$

La convergència de la sèrie ens garanteix que les matrius Π_i es van fent properes a zero a mesura que i creix. Per tant, prenent n prou gran podem aproximar el procés com

$$Y_t = u_t + \sum_{i=1}^n \Pi_i(n) Y_{t-i}$$

D'aquesta manera obtenim l'estimació \hat{u}_t del soroll

$$\hat{u}_t = Y_t - \sum_{i=1}^{\infty} \hat{\Pi}_i(n) Y_{t-i}$$

Notem per $u^{(1)} = \{u_t^{(1)}\}$, $u^{(2)} = \{u_t^{(2)}\}$, $u^{(3)} = \{u_t^{(3)}\}$ i $u^{(4)} = \{u_t^{(4)}\}$ els sorolls estimats per als model estacionals (1), (2), (3) i (4) definits en la Proposició ???. Rerremarquem que $u^{(i)}$ són matrius de dimensió $K \times T$, on la columna j -èssima correspon al vector soroll j -èssim estimat pel model VARMA estacional multiplicatiu (i) .

Per a cada model, definim

$$d(i) = \| [(\hat{u}_1 - u_1^{(i)})', \dots, (\hat{u}_T - u_T^{(i)})'] \|_2$$

és a dir, la norma euclídea de la diferència entre el soroll vertader estimat a partir de l'aproximació pura VAR, i el soroll estimat per la representació estacional.

Finalment, per a escollir el millor model estacional multiplicatiu, triem el model (i) tal que $d(i) < d(j)$ per $j \neq i, j = 1, \dots, 4$.

A continuació mostrem un exemple de la fiabilitat d'aquest criteri.

Exemple 1.5. Considerem el procés VARMA estacional de període $s = 4$ definit per les matrius:

$$\begin{aligned} \phi_1 &= \begin{bmatrix} 0.2 & 0.5 \\ 0.7 & -0.5 \end{bmatrix} & \Phi_1 &= \begin{bmatrix} 0.5 & 0.5 \\ -0.25 & 0.25 \end{bmatrix} \\ \theta_1 &= \begin{bmatrix} 0.5 & -0.2 \\ 0.4 & 0.9 \end{bmatrix} \\ \Sigma_u &= \begin{bmatrix} 0.2 & 0.1 \\ 0.1 & 0.3 \end{bmatrix} \end{aligned}$$

Observem que en aquest cas l'ordre dels operadors és especialment important, atès que determina el signe de la influència de les sèries en el cinquè retard. En efecte, la matriu que correspon a l'operador B^5 per al procés VARMA(1)(1)₄(1)(0)₄ és

$$\phi_1 \Phi_1 = \begin{bmatrix} 0.025 & -0.225 \\ -0.475 & -0.225 \end{bmatrix}$$

mentre que per al procés VARMA(1)₄(1)(1)(0)₄ és:

$$\Phi_1 \phi_1 = \begin{bmatrix} -0.45 & 0 \\ -0.125 & -0.25 \end{bmatrix}$$

La taula següent resumeix els resultats del criteri d'ajustament del soroll, és a dir, de la distribució de diferències de normes euclídies de la resta de sorolls entre el model correcte (c) i l'alternatiu (a), de 300 simulacions (150 per cada model) per al procés anterior:

	$T = 50$	$T = 100$	$T = 500$	$T = 1000$
<i>% de seleccio correcta</i>	100.0	100.0	100.0	100.0
$ \hat{u} - u^{(c)} - \hat{u} - u^{(a)} $				
<i>Minim</i>	-527.8	-763.1	-2490.9	-4775.3
<i>Maxim</i>	-15.8	-87.4	-906.8	-2216.4
<i>Mitjana</i>	-121.5	-267.42	-1539.6	-3085.5

Observem en primer lloc que el criteri aconsegueix identificar correctament el model en el 100% dels casos, amb independència del nombre d'observacions. En segon lloc, veiem que el valor $d(c) - d(a)$ es va fent cada cop més negatiu en augmentar el nombre d'observacions. Això es deu a què per a mostres més grans els sorolls estimats amb l'aproximació autoregressiva tendeixen més als sorolls reals, de manera que $d(c)$ s'aproxima a zero, i $d(a)$ creix en estimar-se millor l'error d'equivocar la selecció, i a més en estimar els errors sobre més sorolls (hi ha tants sorolls com mida té la mostra).

Capítol 2

Estimació i previsió

En aquest capítol es presentaran alguns mètodes i algorismes per estimar els paràmetres dels models AR-MA vectorials i fer-ne previsions. S'estudiaran les propietats d'aquests algorismes, la seva convergència i l'error que cometem, en un intent de posar sobre avís el lector de la precaució que ha de tenir l'estadístic alhora de valorar les bondats de la modelització.

2.1 Estimació dels paràmetres d'un procés VAR(p)

En aquesta secció es presenta com estimar els paràmetres d'un model VAR(p) mitjançant mínims quadrats multivariants. Assumim que es disposa d'una sèrie temporal Y_1, \dots, Y_T , amb Y_i un vector amb K variables i p premostres per a cada variable, és a dir, tenim Y_{-p+1}, \dots, Y_0 (ja sigui suposant-les iguals a la mitjana, o bé utilitzant les primeres observacions com a premostres). Definim:

$$\begin{aligned} Y &:= (Y_1, \dots, Y_T) && (K \times T) \\ B &:= (v, \Phi_1, \dots, \Phi_p) && (K \times (Kp + 1)) \\ Z_t &:= \begin{bmatrix} 1 \\ Y_t \\ \vdots \\ Y_{t-p+1} \end{bmatrix} && ((Kp + 1) \times 1) \\ Z &:= (Z_0, \dots, Z_{T-1}) && ((Kp + 1) \times T) \\ U &:= (u_1, \dots, u_T) && (K \times T) \end{aligned}$$

El model VAR(p) es pot escriure com

$$Y = BZ + U$$

L'estimador de B per mínims quadrats, \hat{B} , és el resultat de resoldre per mínims quadrats el sistema $Y \approx BZ$, o equivalentment $Z'B' = Y'$. Multiplicant per Z i assumint que ZZ' és no singular, hipòtesi que es compleix amb probabilitat 1 si y_t té una distribució contínua, obtenim

$$\hat{B} = YZ'(ZZ')^{-1}$$

En el cas dels models VAR(p) centrats, ajustat per la mitjana, el procés no conté cap constant a estimar. Definint:

$$Y^0 := (Y_1 - \mu, \dots, Y_T - \mu) \quad (K \times T)$$

$$A := (\Phi_1, \dots, \Phi_p) \quad (K \times Kp)$$

$$X_t := \begin{bmatrix} Y_t - \mu \\ \vdots \\ Y_{t-p+1} - \mu \end{bmatrix} \quad (Kp \times 1)$$

$$X := (X_0, \dots, X_{T-1}) \quad (Kp \times T)$$

$$U := (u_1, \dots, u_T) \quad (K \times T)$$

podem expressar el model com $Y^0 = AX + U$ i l'estimador per mínims quadrats de A és $\hat{A} = Y^0 X' (X X')^{-1}$. Quan es desconeix la mitjana del procés, es pot estimar per mitjà de l'estimador mostrat $\hat{\mu} = \frac{1}{T} \sum_{i=1}^T Y_i$.

2.1.1 Estimació sotmesa a constriccions lineals

En l'estimació de paràmetres és important eliminar la influència de paràmetres no significatius que en l'estimació han surtit diferents de zero, atès que poden desvirtuar l'anàlisi i la interpretació de la modelització. Per això, és necessari estudiar la possibilitat d'estimar els paràmetres d'un procés sabent que alguns d'ells seran nuls o estaran subjectes a algún tipus de constricció lineal. Notant per $\beta := \text{vec}(B)$, on vec ¹ és un operador que posa les columnes de les matrius en una única columna, les constriccions

¹Donada una matriu $A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$, on a_i són les columnes de A , aleshores $\text{vec}(A) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$

lineals per a un procés autoregressiu amb constant es poden expressar com

$$\beta = R\gamma + r$$

on R és una matriu coneguda de dimensions $K(Kp+1) \times m$ i rang m , γ és un vector de dimensions $m \times 1$ de paràmetres lliures del model a estimar i r és un vector $K(Kp+1)$ dimensional conegut de constants. Així, per exemple, si volem estimar un procés VAR(3) amb $\Phi_1 = 0$, tindriem $m = 2K^2 + K$ i

$$R = \begin{bmatrix} 0 \\ I_m \end{bmatrix}, \quad \gamma = \text{vec}(v, \Phi_2, \Phi_3), \quad r = 0$$

Representant així les constriccions, obtenim de manera automàtica l'estimador de γ . En efecte,

$$\begin{aligned} y &:= \text{vec}(Y) = (Z' \otimes I_K) \text{vec}(B) + \text{vec}(U) \\ &= (Z' \otimes I_K)(R\gamma + r) + u \end{aligned}$$

i notant $z := y - (Z' \otimes I_K)r$, la solució per mínims quadrats del sistema és

$$\hat{\gamma} = (R'(Z' \otimes I_K)'(Z' \otimes I_K)R)^{-1}R'(Z' \otimes I_K)'z$$

on \otimes denota el producte de Kronecker de matrius ².

Exemple 2.1. Per a veure la bondat de l'estimació per mínims quadrats, es presenta un exemple construït a partir de la simulació del model

$$\begin{bmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \end{bmatrix} = \begin{bmatrix} -0.320 & 0.146 & 0.961 \\ 0.044 & -0.153 & 0.289 \\ -0.002 & 0.225 & -0.264 \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \\ y_{t-1}^3 \end{bmatrix} + \begin{bmatrix} -0.161 & 0.115 & 0.934 \\ 0.050 & 0.019 & -0.010 \\ 0.034 & 0.355 & -0.022 \end{bmatrix} \begin{bmatrix} y_{t-2}^1 \\ y_{t-2}^2 \\ y_{t-2}^3 \end{bmatrix} + \begin{bmatrix} u_t^1 \\ u_t^2 \\ u_t^3 \end{bmatrix}$$

amb $\Sigma_u = \begin{bmatrix} 0.0009 & 0.0010 & 0 \\ 0.0010 & 0.0088 & 0 \\ 0 & 0 & 0.006 \end{bmatrix}$.

En primer lloc, verifiquem l'estabilitat del procés. Per la proposició 1.9 és equivalent a veure que la imatge de la corba $|z| = 1$ per la funció $f(z) = \det(\Phi(z)) = \det(I_3 - \Phi_1 z - \Phi_2 z^2)$ no dona cap volta entorn de $z = 0$. Condició que la figura 2.1 ens permet comprovar. Com que el zero està fora de la corba,

²Donades dues matrius $A \in M_{n \times m}$, $B \in M_{s \times r}$, es defineix el producte de Kronecker de les dues, $A \otimes B \in M_{ns \times mr}$, com:

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \dots & \vdots \\ a_{n1}B & a_{n2} & \dots & a_{nn}B \end{pmatrix}$$

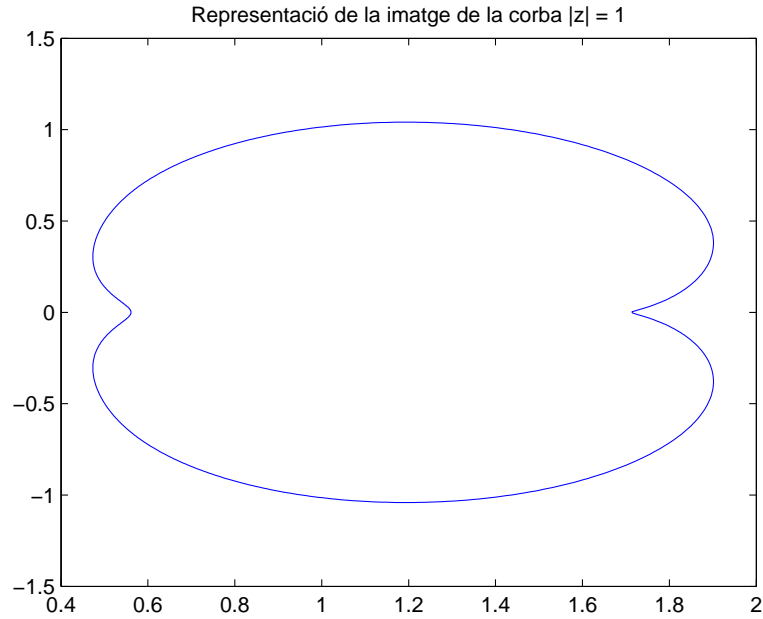


Figura 2.1: Representació de la imatge de la corba $|z| = 1$ per la funció $f(z) = \det(I_3 - \Phi_1 z - \Phi_2 z^2)$.

el procés és estacionari. Ara estem en condicions d'estimar-ne els paràmetres. En concret, per a una simulació de 1000 observacions l'algorisme obté:

$$\hat{\Phi}_1 = \begin{bmatrix} -0.3125 & 0.1424 & 0.9750 \\ 0.0635 & -0.1479 & 0.3014 \\ 0.0411 & 0.2041 & -0.2391 \end{bmatrix} \quad \hat{\Phi}_2 = \begin{bmatrix} -0.1707 & 0.1114 & 0.9298 \\ 0.0716 & 0.0260 & -0.0405 \\ 0.0144 & 0.3489 & -0.0827 \end{bmatrix}$$

De manera que, pel que fa l'error, tenim que els errors relatius -en norma infinit- per a les matrius estimades són: $e_r^{\hat{\Phi}_1} = \frac{\|\Phi_1 - \hat{\Phi}_1\|_\infty}{\|\Phi_1\|_\infty} = \frac{0.0889}{1.427} = 0.0623$ i $e_r^{\hat{\Phi}_2} = \frac{\|\Phi_2 - \hat{\Phi}_2\|_\infty}{\|\Phi_2\|_\infty} = \frac{0.0864}{1.21} = 0.0714$.

2.2 Estimació dels paràmetres d'un procés VMA(q)

Suposem que tenim un procés Y_t de mitjana zero (sinó prenem $Y_t - \mu$) generat per un VMA(q) gausià i invertible

$$Y_t = u_t + \Theta_1 u_{t-1} + \dots + \Theta_q u_{t-q}$$

amb u_t soroll blanc gausià amb matriu de covariàncies Σ_u , i del qual es disposa d'una mostra Y_1, \dots, Y_T . Aleshores podem reescriure'l com

$$Y = \bar{\mathfrak{M}}_q \begin{bmatrix} u_{-q+1} \\ \vdots \\ u_0 \\ u_1 \\ \vdots \\ u_T \end{bmatrix}$$

on

$$\bar{\mathfrak{M}}_q := \begin{bmatrix} \Theta_q & \Theta_{q-1} & \dots & \Theta_1 & 0 & 0 & \dots & 0 \\ 0 & \Theta_q & \dots & \Theta_2 & \Theta_1 & I_k & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \Theta_q & \dots & \Theta_2 & \Theta_1 & I_k \end{bmatrix}$$

és una matriu ($KT \times K(T+q)$).

Com que u_t és soroll blanc gausià, tenim que

$$\begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_T \end{bmatrix} \sim N(0, I_{T+1} \otimes \Sigma_u)$$

on \otimes denota el producte de Kronecker. Llavors, es té que

$$Y \sim N(0, \bar{\mathfrak{M}}_q(I_{T+1} \otimes \Sigma_u)\bar{\mathfrak{M}}_q')$$

i la funció de versemblança és

$$l(\Theta_1, \dots, \Theta_q, \Sigma_u | Y) \propto |\bar{\mathfrak{M}}_q(I_{T+1} \otimes \Sigma_u)\bar{\mathfrak{M}}_q'|^{-1/2} \exp\left\{\frac{1}{2}Y'[\bar{\mathfrak{M}}_q(I_{T+1} \otimes \Sigma_u)\bar{\mathfrak{M}}_q']^{-1}Y\right\}$$

Aquesta expressió pren una forma molt més senzilla si suposem que $u_{-q+1} = \dots = u_0 = 0$. En aquest cas,

$$l_0(\Theta_1, \dots, \Theta_q, \Sigma_u | Y) \propto |\Sigma_u|^{-T/2} \exp\left\{\frac{1}{2}Y'[\mathfrak{M}_q(I_{T+1} \otimes \Sigma_u)\mathfrak{M}_q']^{-1}Y\right\}$$

on

$$\mathfrak{M}_q := \begin{bmatrix} I_k & 0 & 0 & \dots & \dots & \dots & 0 & 0 \\ \Theta_1 & I_k & 0 & \dots & \dots & \dots & 0 & 0 \\ \Theta_2 & \Theta_1 & I_k & 0 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & & & & \\ \Theta_q & \Theta_{q-1} & & \ddots & \ddots & & & \\ 0 & \Theta_q & \ddots & & \ddots & \ddots & & \\ \vdots & & \ddots & \ddots & & \ddots & \ddots & \\ 0 & 0 & \dots & \Theta_q & \dots & \dots & \Theta_1 & I_k \end{bmatrix}$$

i, per tant,

$$\mathfrak{M}_q^{-1} = \begin{bmatrix} I_k & 0 & \dots & 0 \\ -\Pi_1 & I_k & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\Pi_{T-1} & -\Pi_{T-2} & \dots & I_k \end{bmatrix}$$

on els Π_i són les matrius de la representació del procés com pur VAR (veure secció següent). Tenint en compte que $u_t = y_t - \sum_{i=1}^{t-1} \Pi_i y_{t-1}$, s'obté una expressió més senzilla per la versemblança:

$$l_0(\Theta_1, \dots, \Theta_q, \Sigma_u | Y) \propto |\Sigma_u|^{-T/2} \exp \left(-\frac{1}{2} \sum_{t=1}^T u_t' \Sigma_u^{-1} u_t \right)$$

Les matrius $\hat{\Sigma}_u$ i $\hat{\Theta}_1, \dots, \hat{\Theta}_q$ que maximitzen l'expressió anterior són els estimadors maximo-versemblants dels paràmetres del model VMA(q). En la secció 2.3 es detalla l'algorisme que permet trobar-los, per al cas general d'un model VARMA(p,q), i es mostra un exemple detallat per als models VMA i VARMA.

2.3 Estimació dels paràmetres d'un model VARMA(p,q)

L'estimació dels paràmetres per als models VARMA es presentarà per mitjà d'una generalització de l'estimació maximo-versemblant per als VMA. Suposarem que treballem amb un procés VARMA centrat ($\tilde{Y}_t = Y_t - \mu$), de mitjana zero. Donat un procés VARMA(p,q), definim

$$\beta = \text{vec}[\Phi_1, \dots, \Phi_p, \Theta_1, \dots, \Theta_q]$$

on vec és un operador que posa les columnes de les matrius en una única columna.

La funció de versemblança del procés esdevé una funció de β , μ i Σ_u . De la mateixa manera que en

l'expressió de la versemblança per als processos VMA, podem obtenir una expressió aproximada senzilla suposant que

$$\tilde{Y}_{-p+1} = \dots = \tilde{Y}_0 = u_{-q+1} = \dots = u_0 = 0$$

En efecte, llevat de constant,

$$l_0(\mu, \beta, \Sigma_u) = |\Sigma_u|^{-T/2} \exp\left(-\frac{1}{2} \sum_{t=1}^T u(\mu, \beta)'_t \Sigma_u^{-1} u(\mu, \beta)_t\right)$$

on $u_t(\mu, \beta) = \tilde{Y}_t - \sum_{i=1}^{t-1} \Pi_i(\beta) \tilde{Y}_{t-i}$ amb $\Pi_i(\beta)$ els coeficients de la representació VAR pura del procés.

Els valors $\hat{\beta}$, $\hat{\mu}$ i $\hat{\Sigma}_u$ que maximitzen $l_0(\mu, \beta, \Sigma_u)$ són els estimadors maxiversemblants dels paràmetres del procés.

2.3.1 Equacions normals de la maximització de la versemblança

Per a trobar els paràmetres maximitzarem la funció de log-versemblança. Per tal de trobar les equacions normals de la maximització, és a dir, les condicions de primer ordre d'extrem, derivem l'expressió en funció de μ , β i Σ_u :

$$\begin{aligned} 0 &= \frac{\partial Lnl_0}{\partial \mu'} = - \sum_{t=1}^T u'_t \Sigma_u^{-1}(\mu, \beta) \frac{\partial u(\mu, \beta)_t}{\partial \mu'} = \sum_{t=1}^T u'_t \Sigma_u^{-1}(\mu, \beta) \left[I_k - \sum_{i=1}^{t-1} \Pi_i(\beta) \right] \\ 0 &= \frac{\partial Lnl_0}{\partial \beta'} = - \sum_{t=1}^T u(\mu, \beta)'_t \Sigma_u^{-1}(\mu, \beta) \frac{\partial u(\mu, \beta)_t}{\partial \beta'} \\ 0 &= \frac{\partial Lnl_0}{\partial \Sigma_u} = -\frac{T}{2} \Sigma_u^{-1}(\mu, \beta) + \frac{1}{2} \Sigma_u^{-1}(\mu, \beta) \left[\sum_{t=1}^T u'_t(\mu, \beta) u_t(\mu, \beta) \right] \Sigma_u^{-1}(\mu, \beta) \end{aligned}$$

De la tercera equació en resulta

$$\hat{\Sigma}_u(\mu, \beta) = \frac{1}{T} \sum_{t=1}^T u(\mu, \beta)'_t u(\mu, \beta)_t$$

Substituint $\hat{\Sigma}_u(\mu, \beta)$ en la log-versemblança obtenim:

$$\begin{aligned} lnl_0(\mu, \beta) &= -\frac{T}{2} \ln |\hat{\Sigma}_u(\mu, \beta)| - \frac{1}{2} tr \left[\hat{\Sigma}_u^{-1}(\mu, \beta) \sum_{t=1}^T u_t(\mu, \beta) u_t(\mu, \beta)' \right] \\ &= -\frac{T}{2} \ln |\hat{\Sigma}_u(\mu, \beta)| - \frac{TK}{2} \end{aligned}$$

Per tant, enlloc de resoldre el sistema d'equacions normals o maximitzar $ln l_0$, podem minimitzar

$$\ln |\hat{\Sigma}_u(\mu, \beta)| \quad o \quad |\hat{\Sigma}_u(\mu, \beta)|$$

2.3.2 Algorisme de màxima versemblança o dels “scores”

Per a maximitzar la versemblança es construirà un algorisme iteratiu inspirat en els mètodes del gradient, que s’anomena mètode dels “scores”. Assumim que el nostre problema de minimitzar la funció $-\ln l_0$ o equivalentment $\ln|\hat{\Sigma}_u(\mu, \gamma)|$, és el de minimitzar una funció de classe \mathcal{C}^2 respecte els paràmetres β , que notem per $h(\beta)$.

Donat un vector β_i , buscarem una direcció d de descens i una passa s de manera que $h(\beta_i + sd) < h(\beta_i)$. Donada D_i una matriu definida positiva, aleshores, notant per $\nabla h(\beta)$ el gradient de la funció h , $d = -D_i \nabla h(\beta_i)^T$ és de descens. En efecte, desenvolupant per Taylor

$$\begin{aligned} h(\beta_i + sd) &= h(\beta_i - sD_i \nabla h(\beta_i)^T) \\ &= h(\beta_i) - s \nabla h(\beta_i) D_i \nabla h(\beta_i)^T + \frac{s^2}{2} \nabla h(\beta_i) D_i^T \nabla^2 h(\beta_i) D_i \nabla h(\beta_i)^T + \dots \end{aligned}$$

d’on, per a valors de s suficientment petits, donat que s^2 i les seves potències superiors seran menyspreables respecte de s , obtenim:

$$h(\beta_i - sD_i \nabla h(\beta_i)^T) - h(\beta_i) \approx -s \nabla h(\beta_i) D_i \nabla h(\beta_i)^T < 0$$

atesa la definició positiva de D_i i sempre i quant el gradient sigui diferent de zero (en cas contrari, s’ha torbat ja l’extrem).

Per a la matriu D_i , idealment es voldria escollir la inversa de la hessiana de la funció h , és a dir, $D_i = \nabla^2 h(\beta_i)$ (o una aproximació definida positiva), atès que a prop del mínim serà definida positiva i que es pot veure que garanteix una convergència local de segon ordre. Malhauradament, calcular la hessiana de la funció $-\ln l_0$ és molt complicat, i per això treballarem amb una estimació $\hat{I}(\beta_i)$ de la matriu d’informació de Fisher, que és l’esperança de la hessiana:

$$I(\beta_i) = E \left[\frac{\partial^2 (-\ln l_0)}{\partial \beta \partial \beta'} \right]_{\beta_i}$$

Així doncs, els paràmetres β s’actualitzen en la iteració i -èssim com

$$\beta_{i+1} = \beta_i - s_i \hat{I}(\beta_i)^{-1} \left[\frac{\partial (-\ln l_0)}{\partial \beta} \right]_{\beta_i}$$

Per a poder posar en funcionament l’algorisme ens queden algunes preguntes fonamentals a respondre. En primer lloc, ens cal saber com trobar un vector de paràmetres inicial β_0 prou proper al mínim. En segon lloc, detallar el càlcul del gradient de la log-versemblança i de l’estimació de la matriu d’informació. Finalment, l’algorisme dóna una estimació de β condicionada a una matriu Σ_u i un vector μ donats, atès que tant el gradient com la hessiana els involucren. Per tant, cal detallar també com estimar-los.

Estimació de valors inicials

Els coeficients d'un procés VARMA(p,q) en forma estàndard

$$Y_t = \Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + u_t + \Theta_1 u_{t-1} + \dots + \Theta_q u_{t-q}$$

poden ser estimats mitjançant mínims quadrats multivariants si els retards de u_t fossin donats. Assumirem en aquesta secció que s'ha restat del procés la mitjana mostral \bar{Y} . La idea bàsica de l'estimació inicial és aproximar el model VARMA(p,q) amb un model VAR(n) amb n prou gran (notablement major que p i q),

$$Y_t = \sum_{i=1}^n \Pi_i(n) Y_{t-i} + u_t(n)$$

i estimar a partir d'aquí els residus

$$\hat{u}_t(n) := Y_t - \sum_{i=1}^n \hat{\Pi}_i(n) Y_{t-i}$$

on $\hat{\Pi}_i(n)$ s'estimen per mínims quadrats.

Un cop feta aquesta estimació, aleshores ja podem estimar tots els coeficients del VARMA(p,q) per mínims quadrats. En efecte, construïm un model multivariant de regressió

$$Y = [\Phi \quad \Theta] X_n + U$$

on definim

$$Y := [Y_1, \dots, Y_T]$$

$$\Phi := [\Phi_1, \dots, \Phi_p]$$

$$\Theta := [\Theta_1, \dots, \Theta_q]$$

$$Y_{t,n} := \begin{bmatrix} Y_t \\ \vdots \\ Y_{t-p+1} \\ \hat{u}_t(n) \\ \vdots \\ \hat{u}_{t-q+1}(n) \end{bmatrix}$$

$$X_n := [Y_{0,n}, \dots, Y_{T-1,n}]$$

i U és la matriu de residus. El que volem és doncs aproximar $Y \approx [\Phi \quad \Theta] X_n$ per mínims quadrats. Per tant, obtenim l'estimació

$$[\hat{\Phi} \quad \hat{\Theta}] := Y X_n' (X_n' X_n)^{-1}$$

Podem obtenir també una estimació inicial de Σ_u fent

$$\hat{\Sigma}_u(n) := \hat{U}\hat{U}'/T$$

on $\hat{U} = Y - [\hat{\Phi} \quad \hat{\Theta}]X_n$.

Càlcul del gradient i de la matriu d'informació

L'expressió del gradient de la log-versemblança entesa com a funció de β és:

$$\frac{\partial \ln l_0}{\partial \beta'} = - \sum_{t=1}^T u_t' \Sigma_u^{-1} \frac{\partial u_t}{\partial \beta'}$$

En la iteració i -èssima de l'algorisme, tindrem una estimació dels residus i de la matriu de covariàncies, $u(\hat{\mu}, \beta_i)_t$ i $\hat{\Sigma}_u$ respectivament. Per tant, per calcular el gradient només ens cal veure com calcular $\frac{\partial u_t}{\partial \beta'}$. El següent lema detalla com fer el càlcul.

Lema 2.1. *Sigui*

$$u_t = Y_t - \Phi_1 Y_{t-1} - \dots - \Phi_p Y_{t-p} - \Theta_1 u_{t-1} - \dots - \Theta_q u_{t-q}$$

$$\beta := \text{vec}[\Phi_1, \dots, \Phi_p, \Theta_1, \dots, \Theta_q]$$

Definim $\frac{\partial u_0}{\partial \beta'} = \frac{\partial u_{-1}}{\partial \beta'} = \dots = \frac{\partial u_{-q+1}}{\partial \beta'} = 0$ i $Y_0 = \dots = Y_{-p+1} = u_0 = \dots = u_{-q+1} = 0$. *Aleshores*

$$\frac{\partial u_t}{\partial \beta'} = -[(Y'_{t-1}, \dots, Y'_{t-p}, u'_{t-1}, \dots, u'_{t-q}) \otimes I_K] - \left[\Theta_1 \frac{\partial u_{t-1}}{\partial \beta'} + \dots + \Theta_q \frac{\partial u_{t-q}}{\partial \beta'} \right]$$

per a $t = 1, \dots, T$.

Demostració. Notem $\Phi_i = (\phi_{1,i}, \dots, \phi_{K,i})$ i $\Theta_j = (\theta_{1,j}, \dots, \theta_{K,j})$, on $\phi_{r,i}$ i $\theta_{s,j}$ són les columnes r -èssima i s -èssima de les matrius Φ_i i Θ_j respectivament.

Reexpressem u_t en funció de β . Això és:

$$u_t = Y_t - \Phi_1 Y_{t-1} - \dots - \Phi_p Y_{t-p} - \Theta_1 u_{t-1} - \dots - \Theta_q u_{t-q}$$

$$= Y_t - (\phi_{1,1}, \dots, \phi_{K,1})Y_{t-1} - \dots - (\phi_{1,p}, \dots, \phi_{K,p})Y_{t-p} - (\theta_{1,1}, \dots, \theta_{K,1})u_{t-1} - \dots - (\theta_{1,q}, \dots, \theta_{K,q})u_{t-q}$$

$$= Y_t - \sum_{i=1}^p \sum_{v=1}^K Y_{t-i}^v \phi_{v,i} - \sum_{j=1}^q \sum_{w=1}^K u_{t-j}^w \theta_{w,j}$$

$$= Y_t - \sum_{i=1}^p \sum_{v=1}^K Y_{t-i}^v I_K \phi_{v,i} - \sum_{j=1}^q \sum_{w=1}^K u_{t-j}^w I_K \theta_{w,j}$$

$$= Y_t - [(Y'_{t-1}, \dots, Y'_{t-p}, u'_{t-1}, \dots, u'_{t-q}) \otimes I_K] \beta$$

on Y_{t-i}^v i u_{t-j}^w són la sèrie v -èssima i el soroll w -èssim en els instants $t-i$ i $t-j$ respectivament.

Ara la fórmula del lema resulta immediatament derivant i aplicant la regla de la cadena. \square

Pel que fa la matriu d'informació, tenim que

$$\begin{aligned} \frac{\partial^2(-\ln l_0)}{\partial\beta\partial\beta'} &= \frac{\partial}{\partial\beta'} \left[\sum_{t=1}^T \frac{\partial u_t'}{\partial\beta} \Sigma_u^{-1} u_t \right] \\ &= \sum_{t=1}^T \frac{\partial u_t'}{\partial\beta} \Sigma_u^{-1} \frac{\partial u_t}{\partial\beta'} + (u_t' \Sigma_u^{-1} \otimes I_k) \frac{\partial \text{vec}[\partial u_t' / \partial\beta]}{\partial\beta'} \end{aligned}$$

Prenent esperances, l'últim terme desapareix, ja que $E(u_t) = 0$ i $u_t' \Sigma_u^{-1} \otimes I_k$ és independent de $\frac{\partial \text{vec}[\partial u_t' / \partial\beta]}{\partial\beta'}$, donat que el darrer no conté termes en Y_t ni u_t , com hem vist al lema anterior. En conclusió,

$$E \left[\frac{\partial^2(-\ln l_0)}{\partial\beta\partial\beta'} \right] = \sum_{t=1}^T E \left[\frac{\partial u_t'}{\partial\beta} \Sigma_u^{-1} \frac{\partial u_t}{\partial\beta'} \right]$$

Estimant l'esperança dels termes de la dreta a partir de les mitjanes mostrals, s'arriba a l'estimador de la matriu d'informació:

$$\hat{I}(\beta) = \sum_{t=1}^T \frac{\partial u(\hat{\mu}, \beta)'_t}{\partial\beta} \Sigma_u^{-1} \frac{\partial u(\hat{\mu}, \beta)_t}{\partial\beta'}$$

Estimació de μ i de Σ_u

Per a l'estimació de la mitjana, utilitzarem la mitjana mostral de la mostra

$$\hat{\mu} = \bar{Y} = \frac{1}{T} \sum_{t=1}^T Y_t$$

atès que és un estimador amb les mateixes propietats asimptòtiques que l'estimació maximovertsemblant de la mitjana.

Per a Σ_u , hem vist com donar-ne una estimació inicial a partir del càlcul de β_0 per mitjà d'un model de mínims quadrats. De les equacions normals es dedueix que una bona possibilitat és utilitzar, en la iteració i -èssima,

$$\hat{\Sigma}_u := \frac{1}{T} \sum_{i=1}^T \hat{u}_i(\hat{\mu}, \beta_i) \hat{u}_i(\hat{\mu}, \beta_i)'$$

2.3.3 Estimació sotmesa a constriccions lineals

Com hem vist en la secció 2.1.1, sovint cal imposar constriccions lineals sobre els paràmetres del model, especialment per fixar com a nuls aquells paràmetres que no són significatius. Representem novament les

constriccions per

$$\beta = R\gamma + r$$

on R és una matriu coneguda de dimensions $(K^2(p+q) + K) \times m$ i rang m , γ és un vector de dimensions $m \times 1$ de paràmetres lliures del model a estimar i r és un vector $K^2(p+q) + K$ dimensional conegut de constants. L'algorisme d'estimació de γ és exactament igual al de β que hem descrit amb dues salvetats. En primer lloc, en l'estimació de l'aproximació inicial per mínims quadrats caldrà imposar les constriccions lineals sobre els paràmetres de la manera descrita en la secció 2.1.1. En segon lloc, l'expressió de les derivades $\partial u_t / \partial \gamma'$ esdevé

$$\begin{aligned} \frac{\partial u_t}{\partial \gamma'} &= [\Phi_1 Y_{t-1} + \dots + \Phi_p Y_{t-p} + \Theta_1 u_{t-1} + \dots + \Theta_q u_{t-q}]' \otimes I_K \\ &\times [I_{K^2} \ 0 \ \dots \ 0] R - [(Y'_{t-1}, \dots, Y'_{t-p}, u'_{t-1}, \dots, u'_{t-q}) \otimes I_K] [O \ I_{K^2(p+q)}] R \\ &- \left[\Theta_1 \frac{\partial u_{t-1}}{\partial \gamma'} + \dots + \Theta_q \frac{\partial u_{t-q}}{\partial \gamma'} \right] \end{aligned}$$

Exemple 2.2. Per a veure la qualitat de l'estimació per maximoversemblança en els models VMA(q) es presenta un exemple construït a partir de la simulació del model

$$\begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.3 \\ 0.3 & 0.2 \end{bmatrix} \begin{bmatrix} u_{t-1}^1 \\ u_{t-1}^2 \end{bmatrix} + \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix} \begin{bmatrix} u_{t-2}^1 \\ u_{t-2}^2 \end{bmatrix} + \begin{bmatrix} u_t^1 \\ u_t^2 \end{bmatrix}$$

amb $\Sigma_u = \begin{bmatrix} 0.0077 & 0.0013 \\ 0.0013 & 0.0077 \end{bmatrix}$.

Per a veure la seva invertibilitat, calculem les arrels de l'equació $|\Theta(z)| = |I_3 + \Theta_1 z + \Theta_2 z^2| = 0$. Dites arrels són:

$$z_1 = 2.041742 \quad z_2 = -1.397324 \quad z_3 = 1.427791 - 2.593106i \quad z_4 = 1.427791 + 2.593106i$$

Com que totes tenen mòdul major que 1, el procés es pot invertir.

Pel que fa l'estimació dels paràmetres, el resultat de l'algorisme per a una mostra simulada de 1000 observacions és

$$\hat{\Theta}_1 = \begin{bmatrix} 0.531 & 0.305 \\ 0.316 & 0.167 \end{bmatrix} \quad \hat{\Theta}_2 = \begin{bmatrix} 0.195 & -0.037 \\ 0.054 & 0.239 \end{bmatrix}$$

L'error relatiu comès en l'estimació és: $e_r^{\hat{\Theta}_1} = \frac{\|\Theta_1 - \hat{\Theta}_1\|_\infty}{\|\Theta_1\|_\infty} = \frac{0.049}{0.8} = 0.0613$ i $e_r^{\hat{\Theta}_2} = \frac{\|\Theta_2 - \hat{\Theta}_2\|_\infty}{\|\Theta_2\|_\infty} = \frac{0.093}{0.2} = 0.465$. Observem que l'error $e_r^{\hat{\Theta}_2}$ és força elevat, però això és degut a haver pres l'estimació sense aplicar primer un test de significació als paràmetres. Cas de fer-ho, els dos paràmetres que han estat estimats

propers a zero es poden considerar zero i l'error relatiu és aleshores molt menor. En efecte, aplicant l'algorisme d'optimització subjecte a les constriccions obtenim:

$$\hat{\Theta}_1 = \begin{bmatrix} 0.530 & 0.305 \\ 0.316 & 0.169 \end{bmatrix} \quad \hat{\Theta}_2 = \begin{bmatrix} 0.193 & 0 \\ 0 & 0.224 \end{bmatrix}$$

i el nou error relatiu per a la segona matriu és $e_r^{\hat{\Theta}_2} = 0.12$.

Exemple 2.3. Vegem finalment un altre exemple per a un model mixte, un VARMA(1,1):

$$\begin{bmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \end{bmatrix} = \begin{bmatrix} 2.5558 & -1.6339 & -1.0733 \\ 3.2402 & -1.8082 & -1.8938 \\ 1.7328 & -1.2544 & -0.4976 \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \\ y_{t-1}^3 \end{bmatrix} + \begin{bmatrix} 1.4926 & 0.5950 & -1.3249 \\ 2.3122 & 0.9469 & -1.7828 \\ 1.7395 & 0.9314 & -1.7196 \end{bmatrix} \begin{bmatrix} u_{t-1}^1 \\ u_{t-1}^2 \\ u_{t-1}^3 \end{bmatrix} + \begin{bmatrix} u_t^1 \\ u_t^2 \\ u_t^3 \end{bmatrix}$$

En primer lloc, comprovem les condicions d'invertibilitat i estacionarietat. Per fer-ho, calculem els valors propis de les matrius de la part de mitjanes mòbils i autoregressiva respectivament. Els de la part VMA són

$$\lambda_1^{\Theta_1} = 0.80025 \quad \lambda_2^{\Theta_1} = 0.31973 \quad \lambda_3^{\Theta_1} = -0.40008$$

i els de la part VAR

$$\lambda_1^{\Phi_1} = -0.44999 \quad \lambda_2^{\Phi_1} = 0.19996 \quad \lambda_3^{\Phi_1} = 0.50003$$

Pel Corol·lari 1.8, com que tenen mòdul menor que 1 el procés és invertible i estable.

Pel que fa l'estimació dels paràmetres, per a una mostra de 1000 observacions l'algorisme obté:

$$\hat{\Phi}_1 = \begin{bmatrix} 2.5795 & -1.6020 & -1.1378 \\ 3.3478 & -1.8750 & -1.8654 \\ 1.7180 & -1.1992 & -0.5313 \end{bmatrix} \quad \hat{\Theta}_1 = \begin{bmatrix} 1.4254 & 0.4306 & -1.2988 \\ 2.0881 & 0.8438 & -1.9302 \\ 1.6392 & 0.7339 & -1.6981 \end{bmatrix}$$

Els errors relatius en norma infinit per a les matrius són $e_r^{\hat{\Phi}_1} = 0.0292$ i $e_r^{\hat{\Theta}_1} = 0.0941$

2.4 Anàlisi de la convergència i l'error de l'algorisme

En aquesta secció estudiarem el funcionament de l'algorisme en funció de la mida de la mostra, del nombre de sèries i de la magnitud de la matriu de covariàncies del soroll. Si bé és cert que l'algorisme ja està dotat d'un estudi de l'error, en el sentit que estima l'error estàndard dels paràmetres i per tant situa l'estadístic en la magnitud del possible error comès, l'anàlisi que presentem en aquesta secció ens permet tenir una idea del funcionament global de l'algorisme a priori, dels motius pels quals es produeixen els errors, i ens dota de possibles eines per a minimitzar-los.

2.4.1 Metodologia

Per a l'estudi, es realitzaran simulacions de processos estocàstics. Dites simulacions requereixen de variables multinomials (soroll blanc) i d'un generador de nombres aleatoris que permeti reproduir distribucions uniformes (veure Annex A). Fixa't un procés, es faran 50 simulacions del mateix i per a cadascuna s'estimaràn els seus paràmetres i se'n calcularà l'error relatiu en norma infinit. Donada una matriu A i una seva estimació \hat{A} , definim l'error relatiu com $e_r^{\hat{A}} = \frac{\|\hat{A}-A\|_{\infty}}{\|A\|_{\infty}}$. De les 50 simulacions per al procés s'estimarà la mitjana mostral de l'error relatiu, de manera que se'n tindrà una estimació precisa i robusta. Aquest procés es repetirà 50 vegades, és a dir, es simularan 50 processos diferents, i es mostraran boxplots dels resultats.

2.4.2 Algorisme de mínims quadrats per a processos VAR

Comencem l'estudi de la convergència dels algorismes d'estimació analitzant el cas dels processos VAR. En concret, analitzarem els errors produïts en l'estimació de processos VAR(1) estacionaris.

En primer lloc, es presenta com evoluciona l'error de l'algorisme en funció del nombre d'observacions per a processos amb $K = 3$ sèries. S'han pres matrius de covariàncies per al soroll blanc diagonals amb coeficients de l'ordre d'entre 10^{-3} i 10^{-2} (com es veura més endavant, l'ordre de magnitud no té importància en l'estimació dels paràmetres) i matrius de paràmetres Φ_1 del procés amb coeficients molt variats generades com

$$\Phi_1 = SDS^{-1}$$

on D és una matriu diagonal amb els valors propis de Φ_1 , i S es genera amb coeficients aleatoris en $[0, 1]$.

La figura 2.2 mostra clarament com el nombre d'observacions és decisiu per a la bona estimació dels paràmetres. Concretament, les mitjanes dels errors relatius són

<i>Observacions</i>	30	70	100	300	1000
<i>Mitjana d'error</i>	21.34%	12.66%	10.28%	5.98%	3.27%

En segon lloc, estudiem l'evolució de l'error segons el nombre K de sèries, prenent 500 observacions i amb matrius de sorolls qualssevol, simètriques i definides positives, amb coeficients de l'ordre d'entre 10^{-3} i 10^{-2} . En la figura 2.3 veiem que el nombre de sèries no està massa relacionat amb l'error comés en l'estimació de processos VAR amb múnims quadrats. De fet, la mitjana de l'error per $K = 9$ és del 6.4% mentre que per a $K = 3$ és del 8.75%.

Finalment, estudiem el comportament de l'algorisme segons la magnitud dels termes de la matriu Σ_u . En aquesta anàlisi, la matriu Σ_u es genera com $\Sigma_u = S \cdot S'$, de manera que té coeficients d'ordres variats i

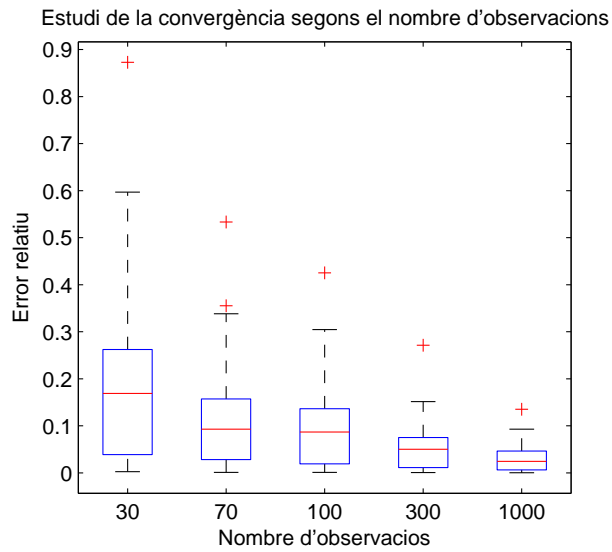


Figura 2.2: Estudi de l'error en l'estimació dels paràmetres segons el nombre d'observacions per a processos VAR(1)

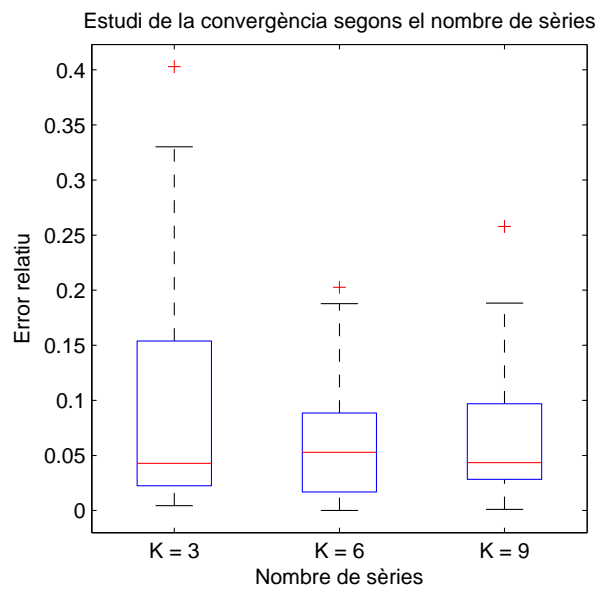


Figura 2.3: Estudi de l'error en l'estimació dels paràmetres segons el nombre K de sèries per a processos VAR(1)

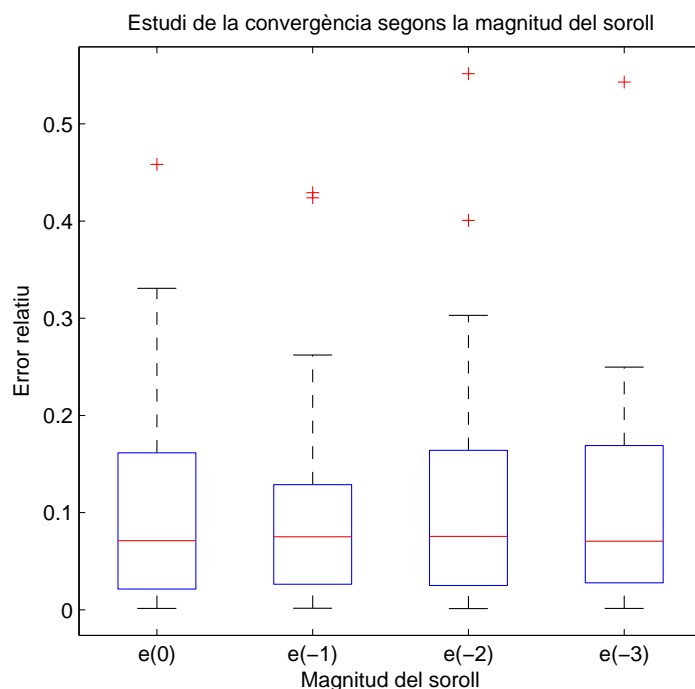


Figura 2.4: Estudi de l'error en l'estimació dels paràmetres segons l'ordre de magnitud dels coeficients de la matriu de covariàncies del soroll per a processos VAR(1). Els termes $e(i)$ representen la potència de 10 per la qual ha estat multiplicada la matriu Σ_u . Així, per exemple, $e(-2)$ vol dir que s'ha utilitzat la matriu $10^{-2}\Sigma_u$.

similars als de la matriu Φ , i després es divideix entre una potència i de 10, que notem per $e(i)$. La figura 2.4, juntament amb les mitjanes dels errors, mostren com la magnitud del soroll no altera la robustesa de l'algorisme.

<i>Exponent</i>	0	-1	-2	-3
<i>Mitjana d'error</i>	10.76%	10.52%	11.27%	10.43%

2.4.3 Algorisme de màxima versemblança per a processos VMA

En aquesta secció repetim l'anàlisi anterior per a processos VMA(1). Els resultats de la convergència segons el nombre d'observacions es resumeixen en la figura 2.5 i la taula següents:

<i>Observacions</i>	30	70	100	300	1000
<i>Mitjana d'error</i>	31.28%	15.24%	12.55%	6.92%	3.67%

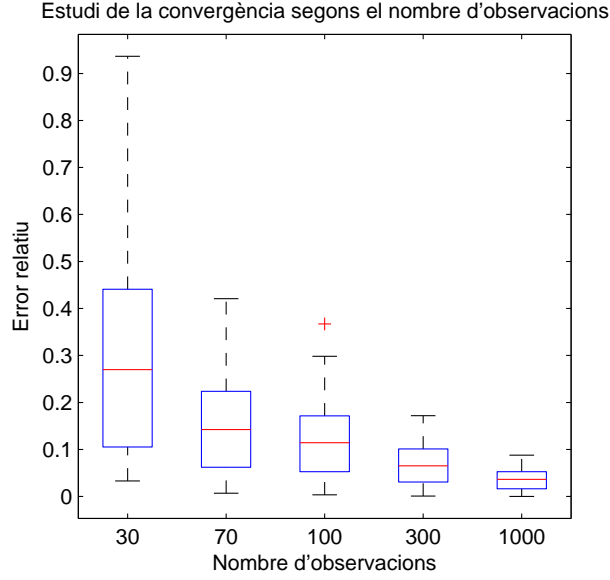


Figura 2.5: Estudi de l'error en l'estimació dels paràmetres segons el nombre d'observacions per a processos VMA(1)

Observem que els errors són força més alts que per als processos VAR quan la mostra té poques observacions, mentre que són molt semblants per a mostres grans. Aquest fet es deu a la necessitat d'estimar els sorolls a partir de sorolls passats i coeficients estimats. En efecte, per a un procés general VMA(q), tenim que

$$u_t = y_t - \sum_{i=1}^q \Theta_i u_{t-i}$$

Notant per δ_{t-i} l'error comès en l'estimació \hat{u}_{t-i} de l'error u_{t-i} , és a dir, $\hat{u}_{t-i} = u_{t-i} + \delta_{t-i}$, i per Δ_i l'error comès en l'estimació de Θ_i , aleshores obtenim que l'estimació del soroll u_t és:

$$\begin{aligned} \hat{u}_t &= y_t - \sum_{i=1}^q \hat{\Theta}_i \hat{u}_{t-i} \\ &= y_t - \sum_{i=1}^q (\Theta_i + \Delta_i)(u_{t-i} + \delta_{t-i}) \\ &= y_t - \sum_{i=1}^q \Theta_i u_{t-i} + \sum_{i=1}^q \Theta_i \delta_{t-i} + \sum_{i=1}^q \Delta_i u_{t-i} + \sum_{i=1}^q \Delta_i \delta_{t-i} \\ &= u_t + \left(\sum_{i=1}^q \Theta_i \delta_{t-i} + \sum_{i=1}^q \Delta_i u_{t-i} + \sum_{i=1}^q \Delta_i \delta_{t-i} \right) \end{aligned}$$

En conclusio, l'error δ_t és

$$\delta_t = \sum_{i=1}^q \Theta_i \delta_{t-i} + \sum_{i=1}^q \Delta_i u_{t-i} + \sum_{i=1}^q \Delta_i \delta_{t-i}$$

Veiem, doncs, que l'error es va multiplicant per les matrius i els errors d'estimació de les matrius del procés. Això provoca, d'una banda, que per a mostres amb poques observacions, com Δ_i de l'aproximació inicial pot ser gran i l'error inicial en l'estimació dels primers u_t es propaga multiplicant-lo per les matrius Θ_i , l'error δ_t creix molt ràpid i desestabilitza l'estimació de Σ_u , fet que impedeix convergir l'algorisme i s'ha de prendre l'aproximació inicial com a valor òptim. D'altra banda, i especialment quan els coeficients de les matrius Θ_i són grans (i per tant ho és la seva norma infinit), l'algorisme també s'atura pel mateix fet, per la desestabilització en l'estimació de Σ_u .

Notis que aquest error és inevitable per a qualsevol algorisme que necessiti estimar els sorolls a partir de sorolls anteriors, i més si es parteix de la hipòtesi que $u_{t-q+1} = \dots = u_0 = 0$.

Exemple 2.4. Vegem un exemple del que acabem d'explicar, mitjantçant el procés definit a partir de

$$\Theta_1 = \begin{bmatrix} 15.2980 & 29.1543 & -32.0139 \\ 3.4895 & 6.6284 & -7.1894 \\ 9.9844 & 18.9950 & -20.7773 \end{bmatrix} \quad \Sigma_u = \begin{bmatrix} 0.8 & 0 & 0 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0.8 \end{bmatrix}$$

Per a una mostra de 1000 observacions, en l'aproximació inicial de l'algorisme s'obté l'estimació

$$\hat{\Theta}_1 = \begin{bmatrix} 15.3084 & 29.1916 & -31.9653 \\ 3.4566 & 6.6568 & -7.2150 \\ 9.9477 & 19.0013 & -20.7478 \end{bmatrix} \quad \hat{\Sigma}_u = \begin{bmatrix} 0.9135 & -0.0138 & 0.0889 \\ -0.0138 & 0.7960 & 0.0714 \\ 0.0889 & 0.0714 & 0.7983 \end{bmatrix}$$

amb un error relatiu $e_r^{\Theta_1} = 1.3\%$. Ara bé, si per a fer l'estimació de Σ_u estimem els sorolls recursivament a partir de la matriu $\hat{\Theta}_1$ i amb $u_0 = 0$, la matriu que s'obté és:

$$\hat{\Sigma}_u = \begin{bmatrix} 14.9207 & 3.3758 & 9.5243 \\ 3.3758 & 1.6065 & 2.2188 \\ 9.5243 & 2.2188 & 7.2702 \end{bmatrix}$$

i el que és més important, si fem el mateix procediment prenent Θ_1 i no una seva aproximació, s'obté:

$$\hat{\Sigma}_u = \begin{bmatrix} 2.5501 & 0.5254 & 1.3093 \\ 0.5254 & 0.9607 & 0.3375 \\ 1.3093 & 0.3375 & 1.8285 \end{bmatrix}$$

Com que $\|\Theta_1\|_\infty = 76.4662$, l'error en l'estimació dels residus elimina qualsevol possibilitat d'estimar correctament la Σ_u si s'utilitza el model VMA(1) per a estimar-los. Per contra, l'aproximació inicial és

molt millor pel fet que no estima els sorolls u_t a partir d'altres sorolls, sinó a partir de les observacions (que no tenen error) i dels coeficients del procés en la seva representació pura VAR. En aquest cas, notant de nou per δ_{t-i} l'error en u_{t-i} i per Δ_i l'error en l'estimació de Π_i , tenim:

$$\begin{aligned}\hat{u}_t &= y_t - \sum_{i=1}^n \Pi(n)_i y_{t-i} \\ &= y_t - \sum_{i=1}^n (\Pi_i + \Delta_i) y_{t-i} \\ &= y_t - \sum_{i=1}^n \Pi_i y_{t-i} + \sum_{i=n+1}^{\infty} \Pi_i y_{t-i} - \sum_{i=n+1}^{\infty} \Pi_i y_{t-i} + \sum_{i=1}^n \Delta_i y_{t-i} \\ &= u_t - \left(\sum_{i=n+1}^{\infty} \Pi_i y_{t-i} - \sum_{i=1}^n \Delta_i y_{t-i} \right)\end{aligned}$$

on, per tant, l'error δ_t és

$$\delta_t = \sum_{i=n+1}^{\infty} \Pi_i y_{t-i} + \sum_{i=1}^n \Delta_i y_{t-i}$$

on el primer terme de la dreta fa referència a l'error que es comet per aproximar la representació infinita en ordre finit n , i el segon a l'error d'estimació dels coeficients de la representació pura VAR. Observem, en primer lloc, que l'error no es propaga en el càlcul dels successius residus. En segon lloc, com que els processos que tractem són invertibles, la sèrie de la representació pura VAR és convergent, de manera que prenent un n prou gran podem fer el primer terme de l'error tant petit com vulguem. Finalment, l'error relacionat amb l'estimació dels paràmetres, per a variables centrades i escalades per controlar la variabilitat, també tendeix a ser petit.

L'estudi de la convergència segons el nombre de sèries es resumeix a partir de la gràfica 2.6. En aquest cas, i a diferència de la robustesa de l'algorisme de mínims quadrats, l'error augmenta d'acord amb l'augment del nombre de sèries. Les mitjanes dels errors són

<i>Series</i>	3	6	9
<i>Mitjana d'error</i>	7.47%	12.18%	14.03%

Cal destacar que l'increment d'error no es produeix tant pel major nombre de paràmetres a estimar, sinó per la tendència a simular matrius Σ_u més mal condicionades. El mal condicionament de la matriu porta a que petites perturbacions afectin de manera decisiva la seva inversa, i per tant, la direcció de descens de l'algorisme d'optimització. A més a més, el mal condicionament està directament relacionat amb el fet que els sorolls u_t^1, \dots, u_t^K tinguin variabilitats molt diverses, i per tant, l'algorisme pugui cometre

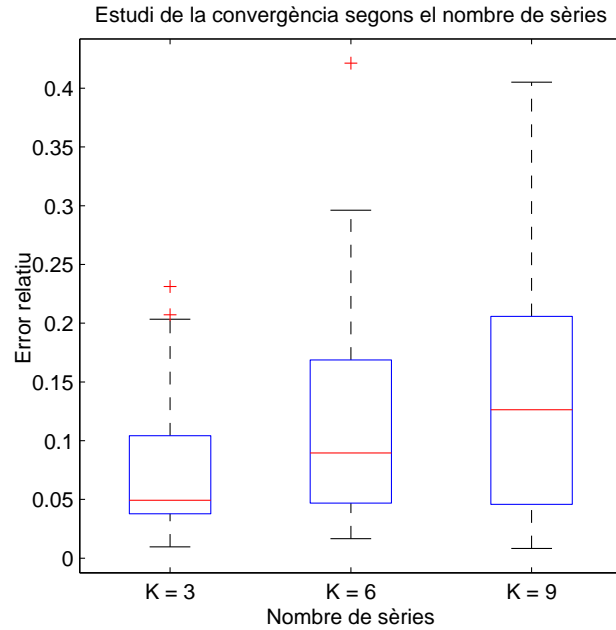


Figura 2.6: Estudi de l'error en l'estimació dels paràmetres segons el nombre K de sèries per a processos VMA(1)

errors importants en els paràmetres que estan vinculats als sorolls més estables, pel fet que són menys importants.

Exemple 2.5. Considerem el procés VMA(1) definit per:

$$\Theta_1 = \begin{bmatrix} 0.5 & 0.2 \\ 0.3 & 0.5 \end{bmatrix} \quad \Sigma_u = \begin{bmatrix} 0.01 & 0.2 \\ 0.2 & 5 \end{bmatrix}$$

La matriu Σ_u està força mal condicionada: el seu nombre de condició és aproximadament 2508. L'estimació dels paràmetres que fa l'algorisme amb una mostra de 1000 observacions és:

$$\hat{\Theta}_1 = \begin{bmatrix} 0.4803 & 0.2019 \\ 0.0454 & 0.5236 \end{bmatrix} \quad \hat{\Sigma}_u = \begin{bmatrix} 0.0097 & 0.1894 \\ 0.1894 & 4.6623 \end{bmatrix}$$

essent l'error relatiu comès en l'estimació de Θ_1 del 34.75%. Veiem que el terme $\Theta_1(2, 1) = 0.3$ està molt mal estimat, $\hat{\Theta}_1(2, 1) = 0.0454$. Com que la variància de u_{t-1}^1 és molt petita en relació a la de u_{t-1}^2 , la sèrie $y_t^2 = u_t^2 + 0.2u_{t-1}^1 + 0.5u_{t-1}^2$ queda governada principalment pel soroll u_t^2 , i per això en la maximització de la versemblança el terme que fa referència a la primera component del soroll és poc important.

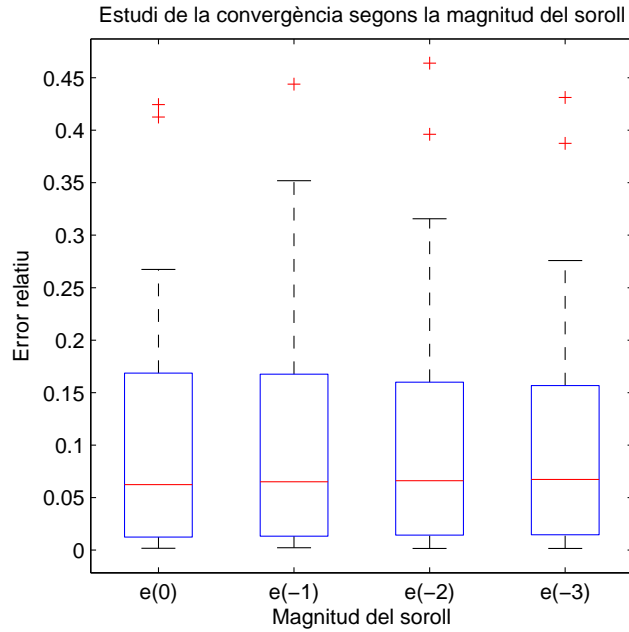


Figura 2.7: Estudi de l'error en l'estimació dels paràmetres segons l'ordre de magnitud dels coeficients de la matriu de covariàncies del soroll per a processos VMA(1).

Quan la matriu Σ_u està ben condicionada, com per exemple

$$\Sigma_u = \begin{bmatrix} 1 & 0.2 \\ 0.2 & 0.5 \end{bmatrix}$$

l'estimació que s'obté per als paràmetres és:

$$\hat{\Theta}_1 = \begin{bmatrix} 0.5083 & 0.1894 \\ 0.2730 & 0.4937 \end{bmatrix} \quad \hat{\Sigma}_u = \begin{bmatrix} 0.9831 & 0.2163 \\ 0.2163 & 0.5056 \end{bmatrix}$$

essent l'error relatiu $e_r^{\Theta_1} = 3.52\%$.

Finalment, l'estudi de l'error segons la magnitud del soroll (figura 2.7) ens mostra les mateixes conclusions que en el cas VAR. Com hem vist, la importància no està en si les variàncies i covariàncies són grans, sinó en el condicionament de la matriu Σ_u .

Exponent	0	-1	-2	-3
Mitjana d'error	10.43%	11.18%	11.65%	11.47%

2.4.4 Algorisme de màxima versemblança per a processos VARMA

En la generalització de l'estudi al cas de models mixtes VARMA(1,1) ens centrarem de manera exhaustiva en l'anàlisi de la convergència segons el nombre d'observacions, ja que sobre la implicació del nombre de sèries i de la magnitud del soroll ha quedat palès que només incideixen en la mesura que empitjora el condicionament de la matriu Σ_u . En el nostre estudi apareixeran, com tot seguit veurem, les problemàtiques que hem vist per als casos VAR i VMA, i d'altres de noves pròpies de l'estructura mixta. En concret, el resultat més important d'aquesta secció és la no convergència de l'algorisme sota determinades condicions.

L'estudi de l'evolució de l'error segons el nombre d'observacions es resumeix en la figura 2.8 i la taula següent:

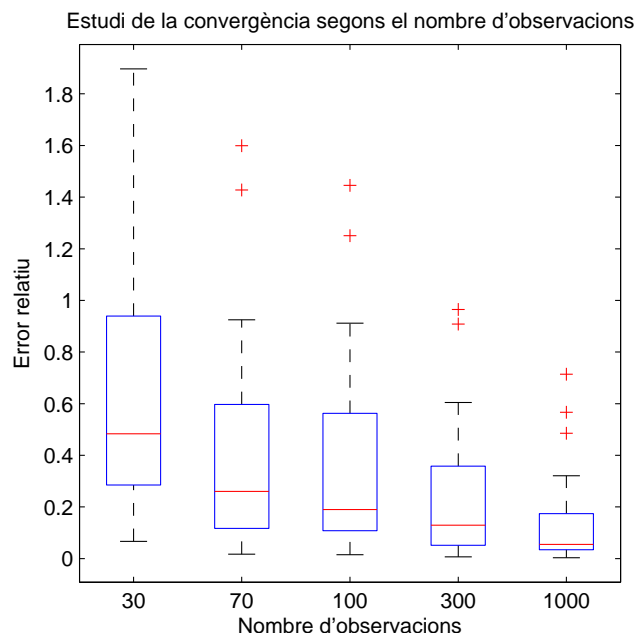


Figura 2.8: Estudi de l'error en l'estimació dels paràmetres segons el nombre d'observacions per a processos VARMA(1,1)

Observacions	30	70	100	300	1000
Mitjana d'error	61.71%	40.13%	36.49%	22.70%	13.75%

Observem que l'error s'ha disparat respecte els casos VAR i VMA. Aquest augment es deu a tres causes. En primer lloc, a l'increment de la propagació de l'error en l'estimació dels sorolls, que ja havíem vist per al cas VMA. Per als processos VARMA(p,q), notant per δ_{t-i} , Δ_i i ξ_i els errors comesos en l'estimació de

u_{t-i} , Θ_i i Φ_i respectivament, l'estimació del soroll u_t és:

$$\begin{aligned}\hat{u}_t &= y_t - \sum_{i=1}^p \hat{\Phi}_i y_{t-i} - \sum_{i=1}^q \hat{\Theta}_i \hat{u}_{t-i} \\ \hat{u}_t &= y_t - \sum_{i=1}^p (\Phi_i + \xi_i) y_{t-i} - \sum_{i=1}^q (\Theta_i + \Delta_i) (u_{t-i} + \delta_i) \\ \hat{u}_t &= u_t - \sum_{i=1}^p \xi_i y_{t-i} - \sum_{i=1}^q \Theta_i \delta_i - \sum_{i=1}^q \Delta_i u_{t-i} - \sum_{i=1}^q \Delta_i \delta_i\end{aligned}$$

d'on

$$\delta_t = -\left(\sum_{i=1}^p \xi_i y_{t-i} + \sum_{i=1}^q \Theta_i \delta_i + \sum_{i=1}^q \Delta_i u_{t-i} + \sum_{i=1}^q \Delta_i \delta_i\right)$$

En segon lloc, gran part de l'increment de l'error es deu a fer simulacions de processos amb matrius amb termes molt dispersos, per exemple amb els de la part autoregressiva molt petits i els de mitjanes mòbils grans. Per a matrius amb termes semblants, com havíem vist en l'exemple 2.3, els errors es comporten de manera molt semblant al cas VMA o VAR, però quan es dona la situació contrària, l'error es descontrola.

Exemple 2.6. Considerem el procés

$$\begin{bmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \end{bmatrix} = \begin{bmatrix} -0.0448 & 0.0549 & -0.0159 \\ -0.0709 & 0.0831 & -0.0040 \\ 0.3140 & 0.1128 & -0.2734 \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \\ y_{t-1}^3 \end{bmatrix} + \begin{bmatrix} -5.4017 & -0.8625 & 3.9456 \\ -6.9728 & -0.6627 & 4.7693 \\ -8.0259 & -1.1820 & 5.8510 \end{bmatrix} \begin{bmatrix} u_{t-1}^1 \\ u_{t-1}^2 \\ u_{t-1}^3 \end{bmatrix} + \begin{bmatrix} u_t^1 \\ u_t^2 \\ u_t^3 \end{bmatrix}$$

amb matriu de covariàncies del soroll

$$\Sigma_u = \begin{bmatrix} 0.07 & 0 & 0 \\ 0 & 0.07 & 0 \\ 0 & 0 & 0.07 \end{bmatrix},$$

per a una mostra de 1000 observacions l'algorisme estima els paràmetres

$$\begin{aligned}\hat{\Phi}_1 &= \begin{bmatrix} -3.0349 & 0.7597 & 1.4896 \\ -4.2798 & 1.3217 & 1.8950 \\ -4.5932 & 1.5601 & 1.9437 \end{bmatrix} & \hat{\Theta}_1 &= \begin{bmatrix} -2.3905 & -1.5420 & 2.5207 \\ -2.6196 & -1.8529 & 2.8756 \\ -3.1216 & -2.5422 & 3.6828 \end{bmatrix} \\ \hat{\Sigma}_u &= \begin{bmatrix} 0.1320 & 0.0735 & 0.0928 \\ 0.0735 & 0.1620 & 0.1133 \\ 0.0928 & 0.1133 & 0.2030 \end{bmatrix}\end{aligned}$$

Aquest exemple és especialment interessant atès que presenta dos fenòmens importants. D'una banda, la dificultat d'estimació per la propagació de l'error, no només perquè els errors Δ_i i ξ_i siguin grans, sinó

que com $\|\Theta_1\|_\infty = 15.06$, fins i tot utilitzant els valors reals de Φ_1 i Θ_1 , l'estimació maximovertsemblant de Σ_u conté un error molt gran:

$$\hat{\Sigma}_u = \begin{bmatrix} 0.1590 & 0.1082 & 0.1329 \\ 0.1082 & 0.2073 & 0.1658 \\ 0.1329 & 0.1658 & 0.2641 \end{bmatrix}$$

No obstant, en l'aproximació pura VAR que realitza l'algorisme dels "scores" per a tenir uns valors inicials per a Φ_1 i Θ_1 , es produeix un error molt petit que, tal i com havíem vist en la secció 2.4.3, permet estimar la matriu de covariàncies de forma molt precisa:

$$\hat{\Sigma}_u = \begin{bmatrix} 0.0676 & 0.0064 & 0.0004 \\ 0.0064 & 0.0743 & 0.0043 \\ 0.0004 & 0.0043 & 0.0718 \end{bmatrix}$$

D'altra banda, una nova font important d'errors en l'estimació és la possibilitat de reduir la dimensionalitat de les sèries. En la secció 1.3.1 havíem introduït la necessitat d'estudiar el nombre de sèries que poden considerar-se independents com a pas previ a la modelització. En aquest exemple, la matriu estimada de covariàncies de les sèries és:

$$\hat{\Sigma}_Y = \begin{bmatrix} 3.3414 & 4.1394 & 4.6529 \\ 4.1394 & 5.3273 & 5.8999 \\ 4.6529 & 5.8999 & 6.7065 \end{bmatrix}$$

que té per valors propis

$$\lambda_1 = 0.0759 \quad \lambda_2 = 0.0777 \quad \lambda_3 = 15.2216$$

Per tant, observem que encara que hi ha 3 sèries, la variabilitat del procés en concentra en una sola dimensió, en la combinació de les sèries que correspon al vector propi de valor propi λ_3 .

Finalment, la tercera font més important d'error és la relació que es dona entre les matrius Φ_1 i Θ_1 quant a vectors propis. Com més semblants són els seus vectors propis, més error es pot produir en l'algorisme, arribant a ser incapaç de convergir a la solució real quan són els mateixos. Aquest punt és molt important perquè en certa manera invalida les condicions d'identificabilitat de Hannan ([8]), en el sentit que si bé des del punt de vista matemàtic els processos certament són únics, des del punt de vista estadístic no. Com veurem en l'exemple següent, processos amb matrius molt diferents poden ser no distingibles per mitjà de la seva versemblança.

Exemple 2.7. Considerem el procés VARMA(1,1)

$$\begin{bmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \end{bmatrix} = \begin{bmatrix} 0.2129 & -0.0189 & -0.0106 \\ 0.1255 & 0.1069 & -0.1062 \\ 0.2884 & -0.1700 & -0.1007 \end{bmatrix} \begin{bmatrix} y_{t-1}^1 \\ y_{t-1}^2 \\ y_{t-1}^3 \end{bmatrix} + \begin{bmatrix} -0.8168 & 0.6976 & -0.1895 \\ -1.2387 & 1.0941 & -0.0493 \\ -0.8874 & 0.5246 & 0.6499 \end{bmatrix} \begin{bmatrix} u_{t-1}^1 \\ u_{t-1}^2 \\ u_{t-1}^3 \end{bmatrix} + \begin{bmatrix} u_t^1 \\ u_t^2 \\ u_t^3 \end{bmatrix}$$

amb

$$\Sigma_u = \begin{bmatrix} 0.088 & 0 & 0 \\ 0 & 0.088 & 0 \\ 0 & 0 & 0.088 \end{bmatrix}$$

La base de vectors propis de les matrius Φ_1 i Θ_1 és, en tots dos casos:

$$S = \begin{bmatrix} 0.0457 & 0.6925 & -0.5092 \\ 0.3598 & 0.6542 & -0.8607 \\ 0.9319 & 0.3041 & -0.0019 \end{bmatrix}$$

L'estimació que realitza l'algorisme per a una mostra de 1000 observacions és:

$$\hat{\Phi}_1 = \begin{bmatrix} -0.7665 & 0.5709 & -0.1530 \\ -1.1927 & 0.9379 & -0.3012 \\ -0.4716 & 0.2930 & -0.1603 \end{bmatrix} \quad \hat{\Theta}_1 = \begin{bmatrix} 0.1165 & 0.1067 & -0.0698 \\ 0.0679 & 0.2466 & 0.1645 \\ -0.1533 & 0.0488 & 0.6428 \end{bmatrix}$$

que, observem, no tenen res a veure amb les matrius originals. Però el que és més rellevant és que a partir d'aquestes matrius es realitza una estimació de la matriu de covariàncies del soroll per màxima versemblança molt precisa:

$$\hat{\Sigma}_u = \begin{bmatrix} 0.0821 & 0.0010 & -0.0010 \\ 0.0010 & 0.0903 & 0.0022 \\ -0.0010 & 0.0022 & 0.0910 \end{bmatrix}$$

i amb $|\hat{\Sigma}_u| = 6.723 \cdot 10^{-4} < 6.815 \cdot 10^{-4} = |\Sigma_u|$, és a dir, més versemblant que la matriu real del procés. Per tant, l'algorisme convergeix i no pot distingir des del punt de vista de la versemblança els valors reals de Φ_1 i Θ_1 .

No obstant, l'aproximació en pur VAR del procés (prenent $p = 16$) és molt robusta i comet molt poc error tant en l'estimació dels coeficients Π_i com de la matriu Σ_u . En efecte, els primers quatre coeficients

de la representació VAR estimats són

$$\hat{\Pi}_1 = \begin{bmatrix} -0.6281 & 0.7098 & -0.1916 \\ -1.1419 & 1.1836 & -0.1896 \\ -0.5943 & 0.3819 & 0.5369 \end{bmatrix} \quad \hat{\Pi}_2 = \begin{bmatrix} 0.2106 & -0.2473 & 0.0627 \\ 0.3994 & -0.4164 & -0.0432 \\ 0.4286 & -0.2658 & -0.4114 \end{bmatrix}$$

$$\hat{\Pi}_3 = \begin{bmatrix} -0.0947 & 0.0649 & -0.0326 \\ -0.2466 & 0.1924 & 0.0444 \\ -0.3689 & 0.2429 & 0.3145 \end{bmatrix} \quad \hat{\Pi}_4 = \begin{bmatrix} -0.0069 & 0.0364 & 0.0259 \\ 0.1390 & -0.0957 & -0.0715 \\ 0.2474 & -0.2043 & -0.2781 \end{bmatrix}$$

mentre que els reals són

$$\Pi_1 = \begin{bmatrix} -0.6039 & 0.6787 & -0.2001 \\ -1.1132 & 1.2010 & -0.1555 \\ -0.5990 & 0.3546 & 0.5492 \end{bmatrix} \quad \Pi_2 = \begin{bmatrix} 0.1698 & -0.2163 & 0.0491 \\ 0.4404 & -0.4558 & -0.0507 \\ 0.4374 & -0.2582 & -0.4529 \end{bmatrix}$$

$$\Pi_3 = \begin{bmatrix} -0.0856 & 0.0924 & -0.0104 \\ -0.2499 & 0.2181 & 0.0939 \\ -0.3646 & 0.2150 & 0.3645 \end{bmatrix} \quad \Pi_4 = \begin{bmatrix} 0.0353 & -0.0359 & -0.0049 \\ 0.1494 & -0.1136 & -0.0976 \\ 0.2921 & -0.1722 & -0.2954 \end{bmatrix}$$

L'error relatiu d'aquests quatre primers termes se situa entre el 3.25% del primer i el 8.70% del tercer.

Pel que fa l'estimació de Σ_u , obtenim

$$\hat{\Sigma}_u = \begin{bmatrix} 0.0839 & -0.0003 & -0.0000 \\ -0.0003 & 0.0885 & 0.0012 \\ -0.0000 & 0.0012 & 0.0831 \end{bmatrix}$$

En no dependre de l'aleatorietat del soroll, l'aproximació pura VAR pot estimar molt millor l'estructura del procés.

Tot i els errors que es produeixen en els paràmetres del procés, el fet que l'estimació hagi convergit des del punt de vista de la màxima versemblança permet obtenir previsions molt semblants a les que s'obtidrien si es disposés dels paràmetres reals. En efecte, en les taules següents es pot observar que les previsions fetes a partir de les matrius estimades i de les matrius reals per a les sèries són molt semblants (en la secció 2.5 es detalla com es realitzen aquestes previsions).

Previsions amb les matrius estimades

	$t = T + 1$	$t = T + 2$	$t = T + 3$	$t = T + 4$
y_t^1	-0.2902	-0.0865	-0.0255	-0.0029
y_t^2	-0.5302	-0.1781	-0.0392	-0.0078
y_t^3	0.1328	-0.0494	0.0064	0.0105

Previsions amb les matrius reals

	$t = T + 1$	$t = T + 2$	$t = T + 3$	$t = T + 4$
y_t^1	-0.3494	-0.0669	-0.0129	-0.0034
y_t^2	-0.5340	-0.1507	-0.0251	-0.0102
y_t^3	0.1212	-0.0217	0.0094	0.0086

A més a més, si construïm els intervals de confiança al 95% per a les previsions, observem que en tots els casos els valors de les sèries hi són continguts:

Valors reals de les sèries

	$t = T + 1$	$t = T + 2$	$t = T + 3$	$t = T + 4$
y_t^1	-0.3215	0.3035	0.0550	-0.3167
y_t^2	-0.5559	0.1198	-0.1183	0.0154
y_t^3	0.4009	0.1689	0.1094	-0.1134

Limit inferior

	$t = T + 1$	$t = T + 2$	$t = T + 3$	$t = T + 4$
y_t^1	-0.8902	-0.8948	-0.8427	-0.8193
y_t^2	-1.1365	-1.3696	-1.2406	-1.2082
y_t^3	-0.4588	-0.8649	-0.8110	-0.8061

Limit superior

	$t = T + 1$	$t = T + 2$	$t = T + 3$	$t = T + 4$
y_t^1	0.3098	0.7217	0.7916	0.8136
y_t^2	0.0762	1.0133	1.1622	1.1926
y_t^3	0.7244	0.7661	0.8238	0.8271

2.5 Estimació de previsions

Suposem que hem estimat els paràmetres d'un procés VARMA(p,q) centrat del qual es disposava d'una mostra de mida T , és a dir, que tenim la representació

$$Y_t - \hat{\mu} = \hat{\Phi}_1(Y_{t-1} - \hat{\mu}) + \dots + \hat{\Phi}_p(Y_{t-p} - \hat{\mu}) + u_t + \hat{\Theta}_1 u_{t-1} + \dots + \hat{\Theta}_q u_{t-q}$$

Aleshores, pels resultats vistos a la secció 1.4, el predictor lineal òptim per a la passa h és

$$\hat{Y}_t(h) = \begin{cases} \hat{\mu} + \hat{\Phi}_1(\hat{Y}_t(h-1) - \hat{\mu}) + \dots + \hat{\Phi}_p(\hat{Y}_t(h-p) - \hat{\mu}) + \hat{\Theta}_h \hat{u}_t + \dots + \hat{\Theta}_q \hat{u}_{t+h-q} & h \leq q \\ \hat{\mu} + \hat{\Phi}_1(\hat{Y}_t(h-1) - \hat{\mu}) + \dots + \hat{\Phi}_p(\hat{Y}_t(h-p) - \hat{\mu}) & h > q \end{cases}$$

on $\hat{Y}_t(h-j) = Y_{h-j}$ per $h-j < T$, i els sorolls $\hat{u}_t = 0$ per $t > T$ (és el valor de l'esperança condicionada als valors passats).

Per a poder construir intervals i regions de confiança ens cal poder estimar la matriu Σ_Y . Utilitzant la representació VAR pura del procés, tenim

$$\hat{Y}_t(h) = \hat{\mu} + \sum_{i=1}^{h-1} \hat{\Pi}_i(\hat{Y}_t(h-i) - \hat{\mu}) + \sum_{i=h}^{\infty} \hat{\Pi}_i(\hat{Y}_{t+h-i} - \hat{\mu})$$

És clar que no disposem de tota la informació que requereix la suma infinita, sinó només fins a $i = T$. Tot i això, de moment considerarem la suma infinita. Així, el predictor de l'error de previsió és

$$Y_{t+h} - \hat{Y}_t(h) = [Y_{t+h} - Y_t(h)] + [Y_t(h) - \hat{Y}_t(h)]$$

on $Y_t(h)$ és el predictor lineal òptim basat en els coeficients coneguts, enlloc d'estimats. Els dos sumands estan no correlacionats atès que el primer es pot escriure en termes de u_s per a $s > t$ i el segon només conté Y_s amb $s \leq t$. Aleshores la matriu de l'EQM predita és

$$\Sigma_{\hat{Y}}(h) = EQM[Y_t(h)] + EQM[Y_t(h) - \hat{Y}_t(h)] = \Sigma_Y(h) + E[Y_t(h) - \hat{Y}_t(h)][Y_t(h) - \hat{Y}_t(h)]'$$

on $\Sigma_Y(h) = \sum_{i=0}^{h-1} \Psi_i \Sigma_u \Psi_i'$ on Ψ_i són els coeficients del model pur VMA. Es pot demostrar que una bona aproximació de $EQM[Y_t(h) - \hat{Y}_t(h)]$ és $\Omega(h)/T$, on

$$\Omega(h) = E \left[\frac{\partial Y_t(h)}{\partial \beta'} \Sigma_{\eta} \frac{\partial Y_t(h)'}{\partial \beta} \right]$$

amb $\eta = [\mu \quad \beta]'$, i en cas que η s'estimi per maximaversemblança, aleshores

$$\Sigma_{\hat{\eta}} = \begin{bmatrix} \Sigma_{\hat{\mu}} & 0 \\ 0 & \Sigma_{\hat{\beta}} \end{bmatrix}$$

Aleshores

$$\frac{\partial Y_t(h)}{\partial \eta'} \Sigma_{\hat{\eta}} \frac{\partial Y_t(h)'}{\partial \eta} = \frac{\partial Y_t(h)}{\partial \mu'} \Sigma_{\hat{\mu}} \frac{\partial Y_t(h)'}{\partial \mu} + \frac{\partial Y_t(h)}{\partial \beta'} \Sigma_{\hat{\beta}} \frac{\partial Y_t(h)'}{\partial \beta}$$

Ens cal trobar una expressió per a les derivades parcials respecte μ i β .

Lema 2.2. *Sigui Y_t un procés VARMA(p,q) invertible i estable, centrat, amb representació VAR pura*

$$Y_t = \mu + \sum_{i=1}^{\infty} \Pi_i(Y_{t-i} - \mu) + u_t$$

Aleshores tenim:

$$\begin{aligned}\frac{\partial Y_t(h)}{\partial \mu'} &= (I_K - \sum_{i=1}^{\infty} \Pi_i) && \text{per } h = 1 \\ &= (I_K - \sum_{i=1}^{\infty} \Pi_i) + \sum_{i=1}^{h-1} \Pi_i \frac{\partial Y_t(h-1)}{\partial \mu'} && \text{per } h = 2, 3, \dots \\ \frac{\partial Y_t(h)}{\partial \beta'} &= \sum_{i=1}^{h-1} ((Y_t(h-1) - \mu)' \otimes I_K) \frac{\partial \text{vec}(\Pi_i)}{\partial \beta'} + \sum_{i=1}^{h-1} \Pi_i \frac{\partial Y_t(h-1)}{\partial \beta'} \\ &+ \sum_{i=h}^{\infty} ((Y_{t+h-1} - \mu)' \otimes I_K) \frac{\partial \text{vec}(\Pi_i)}{\partial \beta'}, && \text{per } h = 1, 2, \dots\end{aligned}$$

amb

$$\frac{\partial \text{vec}(\Pi_i)}{\partial \beta'} = - \sum_{j=0}^{i-1} (H'(M')^{i-1-j} \otimes JM^j) \begin{bmatrix} 0 & I_{Kq} \otimes J' \\ I_{Kq} \otimes J' & 0 \end{bmatrix}$$

i les matrius H , M i J definides de manera que $-\Pi_i = JM^i H$:

$$\begin{aligned}M &:= \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \\ H &:= \begin{bmatrix} I_K & 0 & \dots & 0 & I_K & 0 & \dots & 0 \end{bmatrix}' \quad ((Kq + Kp) \times K) \\ J &:= [I_K \ 0 \ \dots \ 0] \quad (K \times K(p+q))\end{aligned}$$

i les matrius $M_{i,j}$:

$$\begin{aligned}M_{11} &:= \begin{bmatrix} -\Theta_1 & \dots & -\Theta_{q-1} & -\Theta_q \\ I_K & & 0 & 0 \\ & \ddots & \vdots & \vdots \\ 0 & \dots & I_K & 0 \end{bmatrix} \\ M_{12} &= \begin{bmatrix} -\Phi_1 & \dots & -\Phi_{p-1} & -\Phi_p \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix} \\ M_{21} &:= 0 \quad (Kp \times Kq) \\ M_{22} &:= \begin{bmatrix} 0 & \dots & 0 & 0 \\ I_K & & 0 & 0 \\ & \ddots & \vdots & \vdots \\ 0 & \dots & I_K & 0 \end{bmatrix}\end{aligned}$$

Les fórmules que proporciona aquest lema permeten calcular recursivament les derivades parcials de $Y_t(h)$ respecte els coeficients del model per $h = 1, 2, \dots$

Un estimador de $\Omega(h)$ s'obté canviant els termes desconeguts per les seves estimacions, i truncant la suma infinita fins a $i = T$. Notant les derivades parcials estimades per

$$\frac{\hat{\partial}Y_t(h)}{\partial\mu'} \quad \frac{\hat{\partial}Y_t(h)}{\partial\beta'}$$

l'estimador per $\Omega(h)$ és:

$$\hat{\Omega}(h) = \frac{1}{T} \sum_{t=1}^T \left[\frac{\hat{\partial}Y_t(h)}{\partial\mu'} \hat{\Sigma}_{\hat{\mu}} \frac{\hat{\partial}Y_t(h)'}{\partial\mu} + \frac{\hat{\partial}Y_t(h)}{\partial\beta'} \hat{\Sigma}_{\hat{\beta}} \frac{\hat{\partial}Y_t(h)'}{\partial\beta} \right]$$

amb els estimadors

$$\begin{aligned} \hat{\Sigma}_{\hat{\mu}} &:= \hat{\Phi}(1)^{-1} \hat{\Theta}(1) \hat{\Sigma}_u \hat{\Theta}(1)' \hat{\Phi}(1)'^{-1} \\ \hat{\Sigma}_{\hat{\beta}} &:= \left(\frac{1}{T} \hat{I}(\hat{\beta}) \right)^{-1} \end{aligned}$$

Finalment, obtenim l'estimador de la matriu amb l'EQM com

$$\hat{\Sigma}_{\hat{Y}}(h) := \hat{\Sigma}_Y(h) + \frac{1}{T} \hat{\Omega}(h)$$

essent l'estimador de $\Sigma_Y(h)$

$$\hat{\Sigma}_Y(h) = \sum_{i=0}^{h-1} \hat{\Psi}_i \hat{\Sigma}_u \hat{\Psi}_i'$$

Amb aquests resultats i el marc teòric desenvolupat a la secció 1.4, sota hipòtesi de normalitat, podem estimar intervals i regions de confiança per a les previsions.

Exemple 2.8. Considerem el procés estacionari i invertible VARMA(2,1), de mitjana zero, definit a partir de les matrius:

$$\begin{aligned} \Phi_1 &= \begin{bmatrix} 1 & 1.3 \\ 0.45 & -0.25 \end{bmatrix} & \Phi_2 &= \begin{bmatrix} -0.125 & -0.525 \\ -0.475 & -0.435 \end{bmatrix} \\ \Theta_1 &= \begin{bmatrix} 0.5 & -0.25 \\ 0.4 & 0.9 \end{bmatrix} \\ \Sigma_u &= \begin{bmatrix} 0.02 & 0.01 \\ 0.01 & 0.03 \end{bmatrix} \end{aligned}$$

Els estimadors per màximaversemblança per a una mostra de 500 observacions del procés són

$$\hat{\Phi}_1 = \begin{bmatrix} 1.0201 & 1.3252 \\ 0.4866 & -0.2747 \end{bmatrix} \quad \hat{\Phi}_2 = \begin{bmatrix} -0.1311 & -0.5747 \\ -0.5192 & -0.4983 \end{bmatrix}$$

$$\hat{\Theta}_1 = \begin{bmatrix} 0.5321 & -0.3044 \\ 0.3607 & 0.8507 \end{bmatrix}$$

$$\hat{\Sigma}_u = \begin{bmatrix} 0.0201 & 0.0099 \\ 0.0099 & 0.0286 \end{bmatrix}$$

Les previsions a llarg termini per a les sèries en les properes 3 observacions, amb els respectius intervals de confiança amb nivell de significació $\alpha = 0.05$ es mostren a la taula següent:

	<i>Previsió</i>	<i>Limit inferior</i>	<i>Limit superior</i>	<i>Valor real</i>	<i>Error relatiu</i>
$y^1(501)$	0.9215	0.6389	1.2041	1.1501	19.87%
$y^1(502)$	0.4392	-0.2796	1.1579	0.5532	20.60%
$y^1(503)$	0.3750	-0.8380	1.5880	0.4434	15.4%
$y^2(501)$	-0.2045	-0.5416	0.1326	-0.1972	3.70%
$y^2(502)$	-0.0461	-0.5436	0.4515	-0.1339	65.34%
$y^2(503)$	-0.1400	-0.6399	0.3599	-0.3303	57.58%

Observem que l'error relatiu se situa, en mitjana, entorn del 30.41%. Encara que pugui semblar un error elevat, cal tenir present que es tracta d'un fenomen estocàstic, governat de manera important per la part de mitjanes mòbils. El fet de modelitzar el procés ens permet, no només entendre la seva estructura i les interrelacions entre les sèries, sinó de cara a les previsions treballar amb uns errors relatius molt inferiors als que obtindríem si treballéssim amb altres mètodes, com la previsió per mitjà de la mitjana mostral, que per ser aproximadament zero, ens portaria a uns errors relatius propers al 100%.

Capítol 3

Cas pràctic

En aquest darrer capítol es mostrarà una aplicació real dels models VARMA, dins el camp de l'economia. En concret, s'analitzaran detalladament les sèries de Renta Disponible, Consum Personal i Renta del Capital, dels Estats Units d'Amèrica.

3.1 Anàlisi exploratòria

Les sèries que modelarem i estudiarem han estat extretes de la Federal Reserve Bank at St Louis (FED), i recullen les dades mensuals desestacionalitzades, en milers de milions de dòlars, de la Renta Disponible, el Consum Personal i la Renta del Capital per càpita dels Estats Units, des de Gener de 1959 fins a Febrer de 2012. La primera de les sèries, la Renta Personal Disponible (DSPI), correspon a la renda total rebuda pel ciutadà després de pagar impostos. El Consum Personal (PCE) és la riquesa gastada en l'adquisició de béns i serveis pels ciutadans, i la Renta del Capital (PII) recull el conjunt de rentes obtingudes a partir dels interessos en inversions. La figura 3.1 mostra les dades que proporciona la FED

Atès que les sèries presenten tendència, és a dir, no tenen la mitjana constant, no poden considerar-se estacionàries. Per això, treballarem amb els seus percentatges de creixement (en tant per u), calculats a partir de l'expressió:

$$rDPSI := \log(DSPI_t) - \log(DSPI_{t-1})$$

$$rPCE := \log(PCE_t) - \log(PCE_{t-1})$$

$$rPII := \log(PII_t) - \log(PII_{t-1})$$

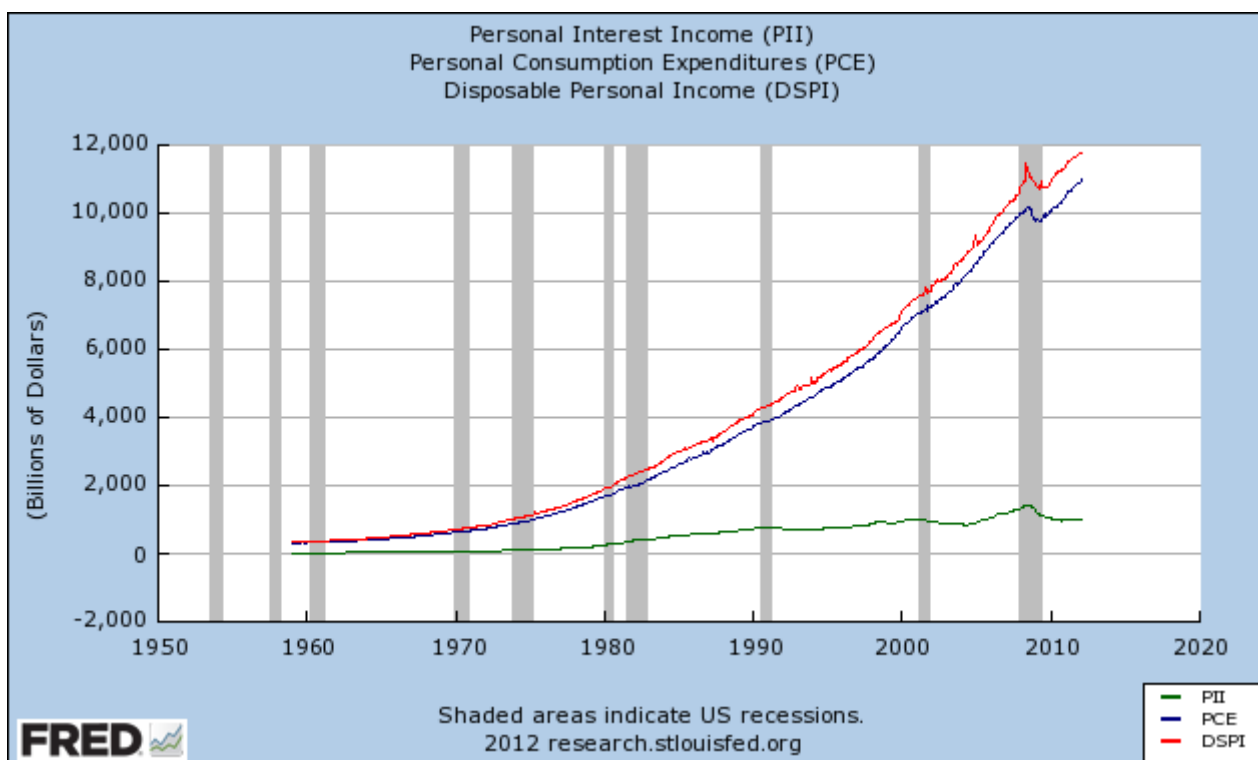


Figura 3.1: Representació de les sèries que recullen l'evolució de la Renta Disponible, el Consum Personal i la Renta del Capital per persona.

Observem que aquesta és una aproximació del percentatge (en tant per u) de creixement. En efecte, fent el desenvolupament de Taylor de l'expressió $\log(X_t) - \log(X_{t-1})$, tenim:

$$\log(X_t) - \log(X_{t-1}) = \log\left(\frac{X_t}{X_{t-1}}\right) = \log\left(1 + \frac{X_t - X_{t-1}}{X_{t-1}}\right) \approx \frac{X_t - X_{t-1}}{X_{t-1}}$$

És important assenyalar que les transformacions fetes són adequades, no només per l'eliminació de la tendència que suposen, sinó també per la reducció i estabilització de la variància de les sèries. La taula següent mostra la variància de les diferents sèries i les seves transformacions:

<i>Transformacions</i>	<i>PCE</i>	<i>DSPI</i>	<i>PII</i>
–	10797304	12449070	162572.2
<i>log</i>	1.2912	1.2587	1.7356
$(1 - B)\log$	$3.1993 \cdot 10^{-5}$	$5.2120 \cdot 10^{-5}$	$9.2392 \cdot 10^{-5}$

En la figura 3.2, podem observar que la transformació de les sèries presenta ja l'aparència d'un procés estacionari.

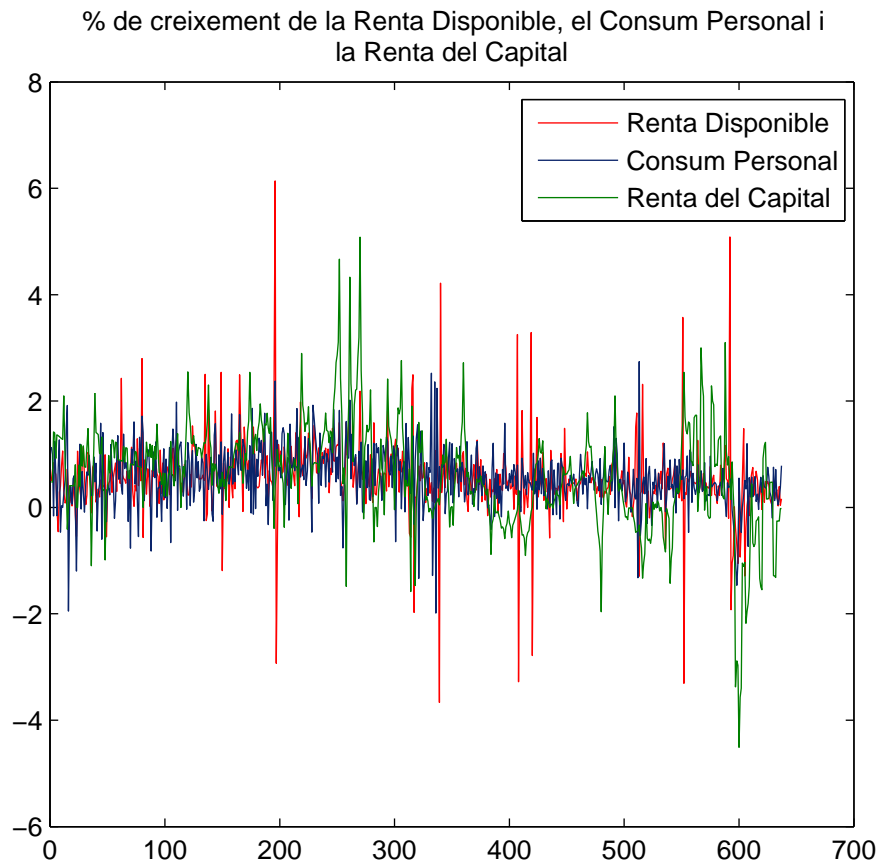


Figura 3.2: Sèries dels % de creixement de la Renta Disponible, el Consum Personal i la Renta del Capital per persona, aproximats a partir de la diferenciació del logaritme de les sèries originals.

3.2 Identificació del model

Per a identificar el model VARMA que s'ajusta a aquestes sèries ens cal estudiar l'estructura de les matrius de correlació simple i parcial de les mateixes. Com hem vist en la Secció 1.3.1, si només les primeres matrius de correlació simple són significatives i les de correlació parcial són totes diferents de zero, provarem la modelització com a procés VMA. Si per contra tenim que les matrius de correlació són totes diferents de zero però només són significatives les primeres matrius d'autocorrelació parcial, ajustarem les dades a un procés VAR. Per a altres estructures, provarem un model mixte.

Les matrius les construirem a partir de prendre les sèries com el vector

$$Y_t = \begin{bmatrix} rPCE_t \\ rDSPI_t \\ rPII_t \end{bmatrix}$$

Pel que fa les matrius de correlació parcial, a partir de la novena matriu es poden considerar no significatives. tal i com mostra la taula següent, que recull els resultats del test de significació

$$H_0 : P(m) = 0$$

$$H_1 : P(m) \neq 0$$

per a les matrius (veure secció 1.1.2):

<i>Retard</i>	<i>Valor del test</i>	<i>p.valor</i>
1	674.7660	$< 10^{-15}$
2	37.9231	$1.7994 \cdot 10^{-5}$
3	97.8374	$< 10^{-15}$
4	82.1741	$5.9730 \cdot 10^{-14}$
5	13.1973	0.1539
6	27.8543	0.001
7	47.5882	$3.0488 \cdot 10^{-7}$
8	19.012	0.0252
9	6.7421	0.6640
10	2.4988	0.9890

Pel que fa les matrius de correlació simple, podem considerar que són totes no nul·les. La figura 3.3 mostra els coeficients de les matrius de correlació per als primers 16 retards, juntament amb els valors

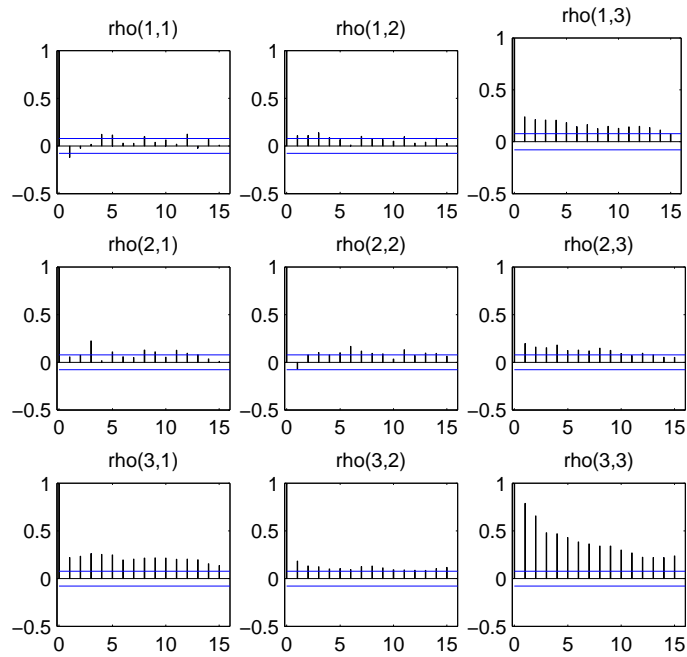


Figura 3.3: Coeficients de les matrius de correlació simple per a les sèries, en funció del retard. Cada gràfica $\rho(i,j)$ representa els valors del coeficient de la fila i i la columna j en les matrius de correlació dels primers 16 retards, juntament amb els llinars de significació amb un nivell de confiança del 95%.

de significació amb un nivell de significació del 0.05 (atès que disposem de 637 observacions, el valor de referència per a considerar els coeficients de les matrius no nuls amb un nivell de significació del 0.05 és $1.96/\sqrt{637} = 0.0777$).

En conclusió, si treballem amb un nivell de confiança per als test del 95%, podem assumir que l'estructura del procés és autoregressiva d'ordre 8. Tot i això, observem que la matriu vuitena podria considerar-se nula treballant amb un nivell de confiança del 99%, és a dir, és significativa per a un test amb nivell de confiança del 95% però per poc. Per això, assajarem primer amb un VAR(7) per tenir menys paràmetres (criteri de parsimònia). Com veurem, aquest model s'ajustarà a les dades.

3.3 Estimació del model

Per a l'estimació del model reservarem les observacions del darrer quadrimestre de manera que puguem verificar la seva capacitat predictiva en la fase de diagnosi, i treballarem amb el vector centrat, després d'estimar la seva mitjana mostral $\hat{\mu}$:

$$\hat{\mu} = \begin{bmatrix} 0.00563 \\ 0.00559 \\ 0.00609 \end{bmatrix}$$

El model estimat per als percentatges de creixement, un cop eliminats el coeficients que es poden considerar no significatius, és:

$$\begin{aligned} Y_t - \hat{\mu} &= \begin{bmatrix} -0.1889 & 0.1218 & 0 \\ 0 & -0.2246 & 0.1809 \\ 0 & 0 & 0.7930 \end{bmatrix} (Y_{t-1} - \hat{\mu}) + \begin{bmatrix} 0 & 0.1228 & 0 \\ 0.1363 & -0.1623 & 0 \\ 0.088 & 0 & 0.1342 \end{bmatrix} (Y_{t-2} - \hat{\mu}) + \dots \\ &+ \begin{bmatrix} 0 & 0.1280 & 0 \\ 0.3031 & -0.1294 & 0 \\ 0.1390 & 0 & -0.5229 \end{bmatrix} (Y_{t-3} - \hat{\mu}) + \begin{bmatrix} 0 & 0.0910 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.4751 \end{bmatrix} (Y_{t-4} - \hat{\mu}) + \dots \\ &+ \begin{bmatrix} 0.1542 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -0.2722 \end{bmatrix} (Y_{t-6} - \hat{\mu}) + \begin{bmatrix} 0.1196 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.2068 \end{bmatrix} (Y_{t-7} - \hat{\mu}) + u_t \end{aligned}$$

i la matriu de covariàncies estimada per al soroll és

$$\hat{\Sigma}_u = \begin{bmatrix} 0.2746 & 0.0966 & 0.0167 \\ 0.0966 & 0.4297 & 0.0330 \\ 0.0167 & 0.0330 & 0.2683 \end{bmatrix} \cdot 10^{-4}$$

Per a veure l'estacionarietat del procés estimat, considerem la funció $f(z) = \det(I_3 - \Phi_1 z - \Phi_2 z^2 - \Phi_3 z^3 - \Phi_4 z^4 - \Phi_6 z^6 - \Phi_7 z^7)$. Recordem que per la Proposició 1.9, el procés és estacionari si la imatge de la corba del pla complex $|z| = 1$ per la funció f no dóna cap volta entorn del zero. I, en efecte, la figura 3.4 mostra com es verifica aquesta condició.

Així doncs, el model ajustat és un procés VAR(7) estacionari.

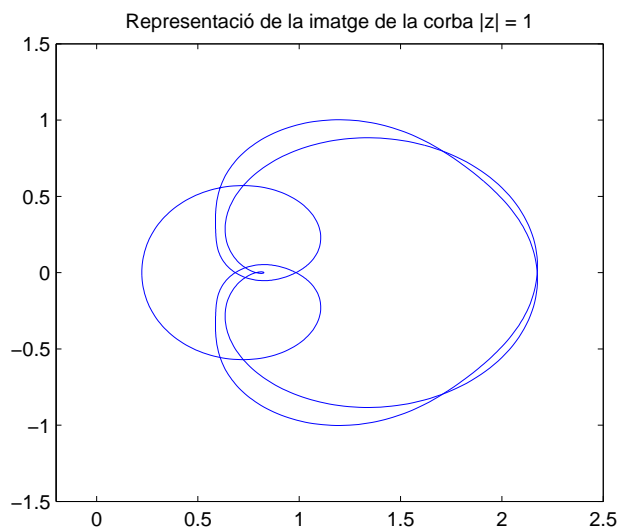


Figura 3.4: Representació de la imatge de la corba del pla complex $|z| = 1$, per la funció $f(z) = \det(I_3 - \Phi_1 z - \Phi_2 z^2 - \Phi_3 z^3 - \Phi_4 z^4 - \Phi_6 z^6 - \Phi_7 z^7)$.

3.4 Diagnosi

Per a concloure la modelització de les sèries, ens cal comprovar la correctesa del model. En aquest sentit, hem de verificar que el soroll estimat és en efecte soroll blanc. Analitzarem si, conjuntament, el vector u_t pot ser considerat un procés vectorial de soroll blanc, i si les seves components són normals.

Pel que fa l'estructura del procés, observem que les matrius de correlació parcial es poden considerar nul·les des del primer retard. En efecte, el p.valor del test de significació de la primera matriu és 0.9456, de manera que podem acceptar amb molta contundència que la primera matriu és no significativa. A partir de la primera matriu, els coeficients estimats per a les altres són dels mateixos ordres de magnitud, i per això podem considerar-les també nul·les. Els p.valors dels test de significació per a les matrius dels primers 10 retards són tots superiors al 0.5, de manera que no hi ha cap indicati per a rebutjar la hipòtesi

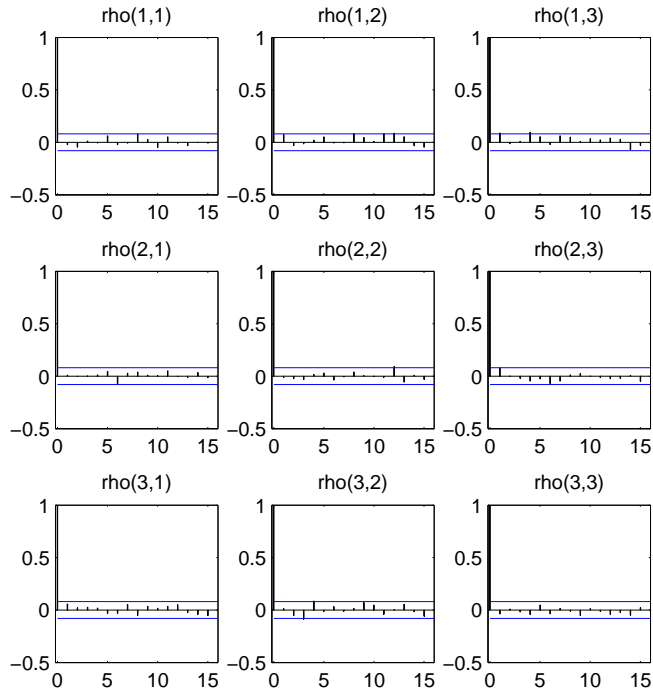


Figura 3.5: Coeficients de les matrius de correlació simple del soroll, en funció del retard. Cada gràfica rho(i,j) representa els valors del coeficient de la fila i i la columna j en les matrius de correlació dels primers 16 retards, juntament amb els líndars de significació amb un nivell de confiança del 95%

nula, és a dir, la hipòtesi que són no significatives. Mostrem aquí les primeres matrius estimades:

$$\hat{P}(1) = \begin{bmatrix} -0.0158 & 0.0074 & -0.0141 \\ -0.0089 & -0.0028 & -0.0086 \\ -0.0067 & -0.0011 & -0.0020 \end{bmatrix} \quad \hat{P}(2) = \begin{bmatrix} -0.0162 & -0.0008 & 0.0048 \\ -0.01543 & -0.0034 & 0.0088 \\ -0.0112 & 0.0019 & 0.0033 \end{bmatrix}$$

$$\hat{P}(3) = \begin{bmatrix} -0.0104 & -0.0075 & 0.0001 \\ -0.0082 & -0.0037 & 0.0040 \\ 0.0033 & 0.0028 & 0.0006 \end{bmatrix} \quad \hat{P}(4) = \begin{bmatrix} 0.0001 & -0.0106 & -0.0092 \\ 0.0040 & -0.0080 & -0.0004 \\ -0.0148 & 0.0063 & -0.0053 \end{bmatrix}$$

Quant a les seves matrius de correlació poden ser considerades també totes nul·les des del primer retard. En efecte, tal i com mostra la figura 3.5 que conté els valors dels coeficients de les matrius, tots els seus coeficients poden ser considerats nuls. Apareixen alguns coeficients no nuls en algunes de les matrius (5 coeficients de 135 estimats, un 3.7%), fet que es pot entendre com a conseqüència de treballar amb un

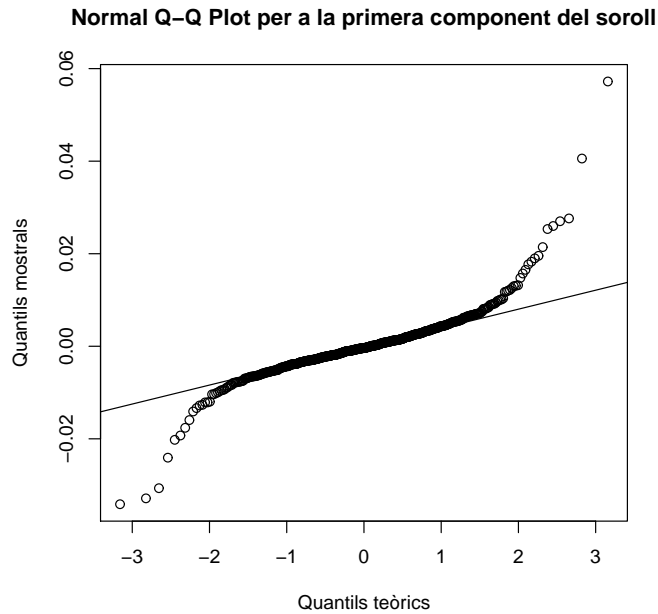


Figura 3.6: QQ-plot de la primera component del soroll.

una confiança del 95%, en el sentit que en un 5% dels casos, estimacions de paràmetres nuls podrien ser considerades no nul·les.

En conclusió, no podem rebutjar que el procés vectorial u_t sigui soroll blanc.

No obstant, no podem considerar-lo gaussià, ja que com mostren els QQ-plots de les seves components u_t^1 , u_t^2 i u_t^3 (figures 3.6, 3.7 i 3.8) presenten cues pesades i no s'ajusten a una distribució normal.

En les sèries econòmiques i financeres és habitual la presència de cues pesades, que en general està relacionada amb el fet que la variància és condicionalment heterocedàstica, és a dir, no és constant, i hi ha moments de volatilitat (com per exemple les crisis). Les figures 3.9 i 3.10, que representen l'autocorrelació simple i parcial dels quadrats de les components dels residus, verifiquen l'existència d'una interestructura dinàmica entre ells, és a dir, la presència d'heterocedasticitat condicionada. Aquest fet no ens porta a descartar la modelització de la sèrie, però obre la porta a considerar un model d'heterocedasticitat condicionada i a analitzar amb més deteniment la capacitat predictiva del model, quant als intervals de confiança, atès que l'estimació de la variància i dels valors superiors i inferior dels intervals està determinada per la hipòtesi de normalitat.

Normal Q-Q Plot per a la segona component del soroll

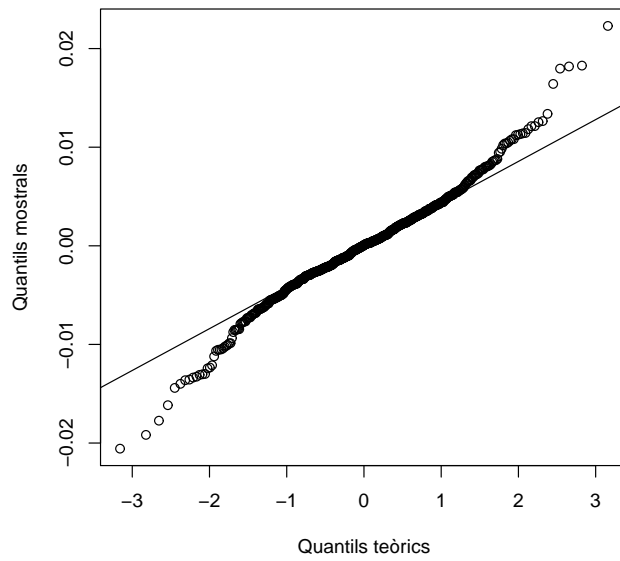


Figura 3.7: QQ-plot de la segona component del soroll.

Normal Q-Q Plot per a la tercera component del soroll

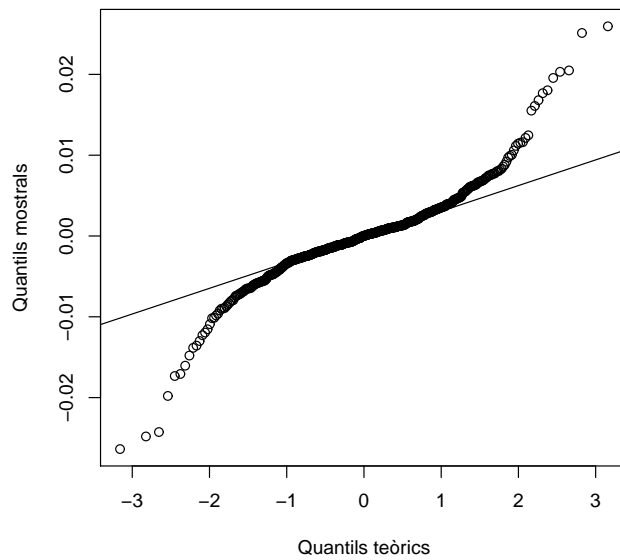


Figura 3.8: QQ-plot de la tercera component del soroll.

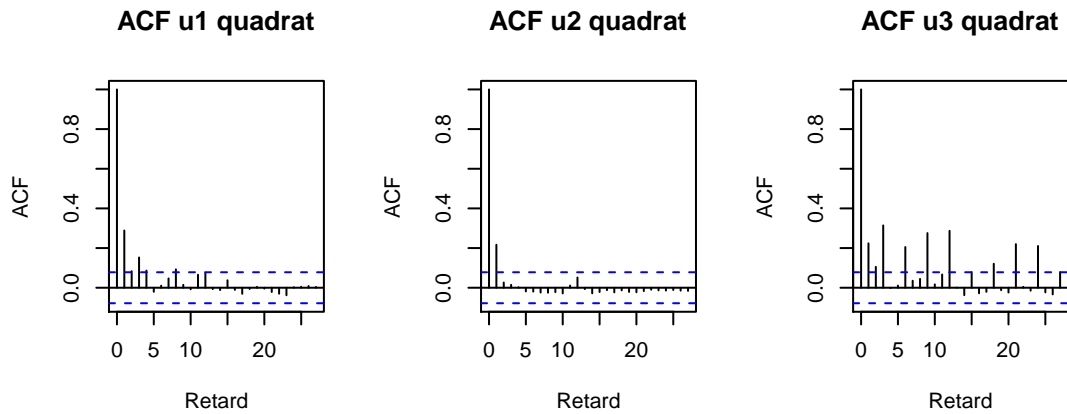


Figura 3.9: Funció d'autocorrelació simple del quadrat de les tres components del soroll.

Per aquest motiu, comprovem la capacitat predictiva de la sèrie per mitjà de les observacions que hem reservat. En les taules següents mostrem els valors predits i els intervals de confiança al 95% per a les sèries $rPII$, $rDSPI$ i $rPCE$, juntament amb els valors reals. Destaquem novament que aquests intervals es construeixen sota la hipòtesi de variància constant, i que precisament hem vist que aquesta no es satisfia.

	<i>Valor real</i>	<i>Previsió rPCE</i>	<i>Limit inferior</i>	<i>Limit superior</i>
<i>Novembre 2011</i>	0.005	0.0043	-0.0063	0.0148
<i>Desembre 2011</i>	0.0018	0.0027	-0.0081	0.0134
<i>Gener 2012</i>	0.0038	0.0040	-0.0069	0.0148
<i>Febrer 2012</i>	0.0078	0.0036	-0.0073	0.0145

	<i>Valor real</i>	<i>Previsió rDSPI</i>	<i>Limit inferior</i>	<i>Limit superior</i>
<i>Novembre 2011</i>	0.0003	0.0006	-0.0126	0.0138
<i>Desembre 2011</i>	0.0039	0.0018	-0.0121	0.0155
<i>Gener 2012</i>	0.0004	0.0008	-0.0131	0.0147
<i>Febrer 2012</i>	0.0016	0.0040	-0.0102	0.0181

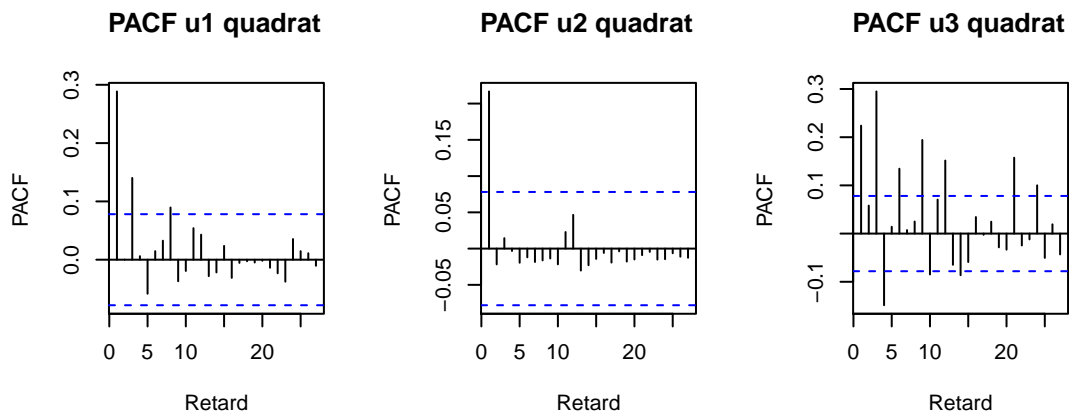


Figura 3.10: Funció d'autocorrelació parcial del quadrat de les tres components del soroll.

	<i>Valor real</i>	<i>Previsió rPII</i>	<i>Limit inferior</i>	<i>Limit superior</i>
<i>Novembre 2011</i>	-0.0026	-0.0038	-0.0143	0.0066
<i>Desembre 2011</i>	-0.0026	-0.0025	-0.0156	0.0105
<i>Gener 2012</i>	-0.0002	-0.0033	-0.0184	0.0118
<i>Febrer 2012</i>	-0.0002	-0.0005	-0.0158	0.0147

Observem que les prediccions capten en tot moment la tendència a l'alça o a la baixa dels creixements i n'aproximen de forma precisa els valors. A més a més, totes les previsions no només són properes al valor real, sinó que estan contingudes en els intervals de confiança.

3.5 Efectes multiplicadors

Acabem l'anàlisi d'aquestes sèries estudiant les relacions que hi ha entre elles. En aquest sentit, ens plantejem quin impacte (efectes multiplicadors) té una variació en una d'elles sobre l'evolució del conjunt de les sèries.

Considerem el vector

$$\nu = \begin{bmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{bmatrix}$$

que podem entendre com una variació en les sèries dels percentatges de creixement de la Renta Disponible, del Consum i de la Renta del Capital en l'instant t . Aleshores, en l'instant $t+1$ el valor del procés centrat

amb l'efecte de l'alteració ν , que notem per $\tilde{Y}_{t+1}(\nu)$, serà:

$$\begin{aligned}\tilde{Y}_{t+1}(\nu) &= \Phi_1(Y_t + \nu) + \Phi_2 Y_{t-2} + \Phi_3 Y_{t-3} + \Phi_4 Y_{t-4} + \Phi_6 Y_{t-6} + \Phi_7 Y_{t-7} + u_{t+1} \\ &= \Phi_1 \nu + (\Phi_1 Y_t + \Phi_2 Y_{t-2} + \Phi_3 Y_{t-3} + \Phi_4 Y_{t-4} + \Phi_6 Y_{t-6} + \Phi_7 Y_{t-7} + u_{t+1})\end{aligned}$$

on veiem que el primer terme de la dreta correspon a l'alteració produïda per ν , i el segon al valor que hagués prè la sèrie en l'instant $t + 1$ si no s'hagués produït ν . A més a més, aquest efecte coincideix amb el primer coeficient de la representació pura VMA del procés (veure secció 1.2.3). Per a l'instant $t + 2$, tenim:

$$\begin{aligned}\tilde{Y}_{t+2}(\nu) &= \Phi_1(\tilde{Y}_{t+1}(\nu)) + \Phi_2(Y_t + \nu) + \Phi_3 Y_{t-2} + \Phi_4 Y_{t-3} + \Phi_6 Y_{t-5} + \Phi_7 Y_{t-6} + u_{t+2} \\ &= \Phi_1^2 \nu + \Phi_2 \nu + Y_{t+2}\end{aligned}$$

d'on, l'efecte de ν en l'instant $t + 2$ és $\Phi_1^2 \nu + \Phi_2 \nu$, que és precisament el segon coeficient de la representació pura VMA. Raonant de la mateixa manera, veiem que en l'instant $t + k$ la sèrie esdevé

$$\tilde{Y}_{t+k}(\nu) = Y_{t+k} + \sum_{j=1}^k \Phi_j \Psi_{k-j} \nu$$

on Ψ_{k-j} és l'efecte en l'instant $k - j$ de ν , i coincideix amb el coeficient $k - j$ de la representació pura VMA, i $\Phi_j = 0$ per $j > 7$.

Estem en condicions d'estudiar els efectes que tenen variacions en cadascuna de les sèries, per separat. Considerarem l'impacte que té un increment d'un 1% en una de les sèries en l'evolució de les altres dues i d'ella mateixa. La figura 3.9 mostra els efectes multiplicadors que s'estableixen entre les sèries del creixement de la Renta Disponible (rDSPI), el Consum Personal (rPCE) i la Renta del Capital (rPII). Analtzem, en primer lloc, les repercussions que té un increment en el creixement del Consum Personal. Veiem que a curt termini aquest increment exogen implica una caiguda de la Renta Disponible, atès, precisament, que s'ha consumit renta. Per contra, té un efecte positiu en la Renta del Capital, que es pot entendre fruit de l'efecte estímulo de demanda, i per tant de creixement econòmic, que té l'augment de consum [12], [13]. A mig i llarg termini, l'efecte es dissipa i la sèrie retorna al seu equilibri de creixement a llarg termini. En segon lloc, un increment en la Renta Disponible provoca un augment superior al mitjà del Consum Personal, produït per l'anomenat efecte riquesa [12]. Els ciutadans, en veure incrementada la seva renta disponible incrementen de manera notable el consum per percebre's més rics. Aquest augment va acompanyat d'una caiguda del creixement de la Renta Disponible, que s'ha utilitzat per aquest consum (veure les gràfiques $rDSPI \rightarrow rPCE$ i $rDSPI \rightarrow rDSPI$ de la figura 3.11). Pel mateix motiu que abans, té també un efecte positiu en les Rentes del Capital, que es produeix amb aquest augment del Consum

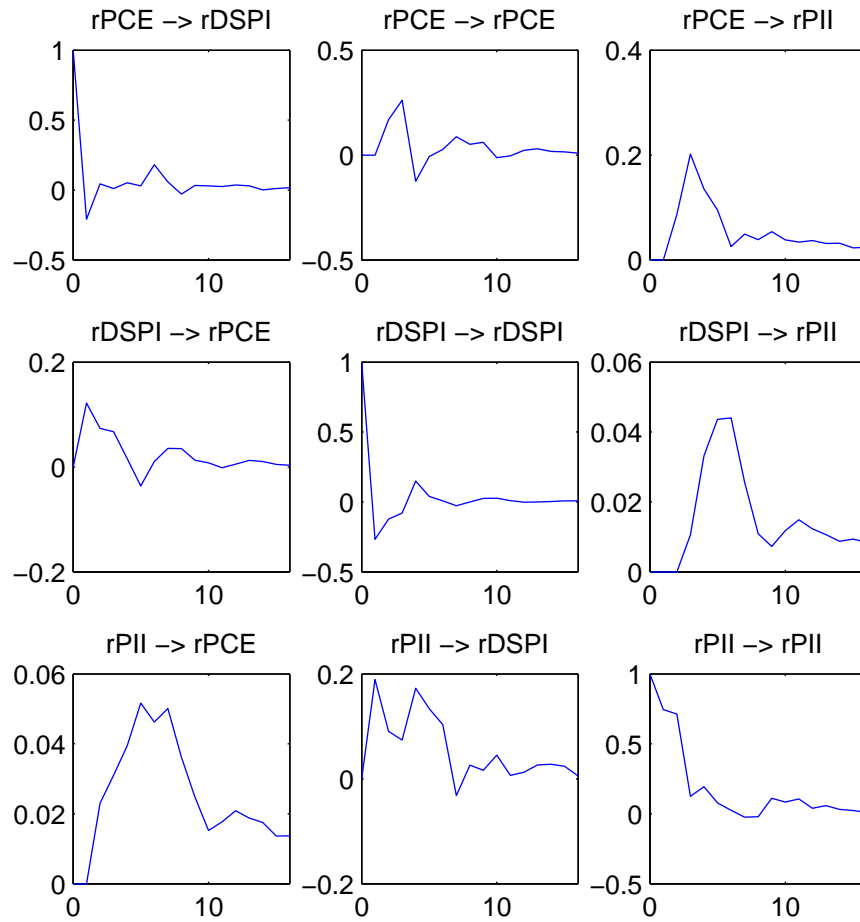


Figura 3.11: Representació dels efectes multiplicatius de les sèries. Les gràfiques mostren com impacta un increment d'un 1% en el creixement d'una sèrie en el % de creixement de les altres al llarg del temps. És a dir, els valors $\tilde{Y}_{t+k}(\nu) - Y_{t+k}$, amb $\nu = [0.01 \ 0 \ 0]^T$ quan estudiem els impulsos de la sèrie $rPCE$, $[0 \ 0.01 \ 0]^T$ per als de $rDSPI$ i $[0 \ 0 \ 0.01]^T$ per als de $rPII$. Així, per exemple, la gràfica $rPCE \rightarrow rDSPI$ representa l'alteració al llarg dels instants $t, t+1, \dots, t+16$ de la sèrie $rDSPI$ quan la $rPCE$ incrementa un 1% en l'instant t .

Personal, i per això té un retard de dos mesos. Finalment, l'increment de les Rentes del Capital té un efecte positiu en els creixements de les altres variables, confirmant els models de creixement econòmic [2], [18], que lliguen la salut i l'evolució d'una economia principalment amb la seva productivitat i nivell d'inversió. Per això, les Rentes del Capital, que són indicador del nivell d'inversió, alimenten el creixement de l'economia en general.

3.6 Conclusions

L'anàlisi de les sèries de Renta Disponible, Consum Personal i Renta del Capital per càpita ens ha permès tenir una visió global de l'anàlisi multivariant per mitjà de models VARMA de sèries temporals.

En la primera fase de l'estudi, l'anàlisi exploratòria, s'han realitzat transformacions sobre les sèries per tal que poguessin considerar-se la realització d'un procés estocàstic estacionari. En aquest sentit, cal destacar la importància de la transformació logarítmica per a l'estabilització de la variància i la diferenciació per a l'eliminació de la tendència. La combinació de totes dues és molt freqüent en el camp de l'economia i les finances donat que, com s'ha demostrat, aproxima els percentatges de creixement de les sèries.

En la segona fase s'ha modelitzat el procés. Per mitjà de les matrius de correlació simple i parcial hem identificat una estructura autoregressiva, justificada pel fet que les matrius de correlació simple es podien considerar totes no nul·les (figura 3.3) mentre que les de correlació parcial eren no significatives a partir del vuitè o novè retard (segons si es treballa amb un nivell de confiança del 99% o del 95%). L'estimació dels paràmetres del model ens ha portat a comprovar que, en efecte, es tractava d'un procés estacionari ja que les arrels de l'equació $\det(\hat{\Phi}(z)) = 0$ estaven fora del cercle unitat. L'estudi dels residus del procés ens ha permès comprovar que es tractava de soroll blanc, atès que totes les matrius de correlació simple i parcial eren nul·les (figura 3.5), però no era gaussià fruit de la presència de cues pesades ens els QQ-plots (figures 3.6, 3.7 i 3.8). Aquest fenomen és habitual en les sèries econòmiques i especialment en les financeres, fruit de què la variància no és constant sinó, tal i com han posat de manifest les gràfiques de les funcions d'autocorrelació simple i parcial del quadrat de les components dels residus (figures 3.9 i 3.10), condicionalment heterocedàstica, és a dir, són sèries amb volatilitat. No obstant, la bona capacitat predictiva del model, que s'ha pogut comprovar per mitjà de comparar les previsions fetes per al període del Novembre 2011 al Febrer 2012 amb els valors reals que es van observar, ens ha permès validar la modelització.

Finalment s'ha interpretat el significat del model obtingut per mitjà de l'estudi dels efectes multiplicadors que s'estableixen entre les sèries, és a dir, l'impacte que té el canvi d'alguna d'elles sobre l'evolució futura de si mateixa i de les altres. En aquest sentit, hem pogut veure que la sèrie que té un major impacte

en el creixement global de les tres és la de la Renta del Capital (figura 3.11), fet que avala les teories de creixement econòmic que situen en el cor d'una economia el seu nivell d'inversió en capital.

Apèndix A

Simulació de variables

En aquest apèndix es detalla com s'han simulat les distribucions aleatòries per a generar processos aleatòris.

El generador de nombres aleatoris que s'ha implementat és el $G9$ descrit per L'Ecuyer, Couture i Blouin a [9]. Es tracta d'un generador múltiple recursiu d'ordre 5 de la forma:

$$z_j = 107374182z_{j-1} + 104480z_{j-5} \pmod{2^{31} - 1}$$

El seu període és de l'ordre $(2^{31} - 1)^5 - 1 \approx 2^{155}$, molt superior als generadors típics de softwares com MATLAB o R.

Per a verificar que genera nombres que poden considerar-se aleatoris, s'han realitzat 30 mostres de 1000 nombres i s'han aplicat els tests khi-quadrat i Kolmogorov-Smirnov. Els resultats es resumeixen en la taula següent:

	<i>p.valor mig</i>	<i>p.valor maxim</i>	<i>p.valor minim</i>
<i>Khi - test</i>	0.4958	0.9589	0.0898
<i>KS - test</i>	0.5091	0.9750	0.0717

A més a més, el generador passa satisfactòriament tots els test de Diehard.

A partir d'aquest generador es simulen distribucions uniformes $U[a, b]$ per mitjà de la transformació

$$u := a + \frac{z}{2^{31} - 1}(b - a)$$

Per a la distribució multinomial de matriu de covariàncies Σ i mitjana μ , s'ha implementat l'algorisme MN detallat a [6]. Vegem el seu funcionament.

Sigui $Z = (Z_1, \dots, Z_r)'$ un vector aleatori multinormal amb densitat de probabilitat

$$f(z) = (2\pi)^{-r/2} |\Sigma|^{-1/2} e^{-(z-\mu)'\Sigma^{-1}(z-\mu)} \quad -\infty < z_i < \infty \quad i = 1, \dots, r$$

on $z = (z_1, \dots, z_r)'$, $\mu = (\mu_1, \dots, \mu_r)$ és el vector de mitjanes, i Σ denota la matriu $r \times r$ de covariàncies (simètria i definida positiva). Notem $Z \sim MN(\mu, \Sigma)$. Ara si $X = (X_1, \dots, X_r)'$ és una $MN(0, Id)$, aleshores podem representar

$$Z = cX + \mu$$

on c és la matriu triangular inferior que factoritza Σ . L'esquema de l'algorisme MN és:

$i \leftarrow 1$

Fins $i > r$:

Generar $X_i \sim N(0, 1)$

$Z_i \leftarrow \mu_i; j \leftarrow 1.$

Fins $j > i$: $Z \leftarrow Z_i + c_{ij}X_j; j \leftarrow j + 1$

$i \leftarrow i + 1$

Retorna $Z = (Z_1, \dots, Z_r)$

on la generació de $N(0, 1)$ es realitza per mitjà del mètode de Box-Muller [1].

Bibliografía

- [1] Banks, J. (1998): *Handbook of Simulation: Principles, Methodology, Advances, Applications and Practice*, John Wiley.
- [2] Barro, R.J., Sala-i-Martin, X. (2003): *Economic Growth*, The MITT Press.
- [3] Box, G. E. P., Cox, D. R. (1964): *An analysis of transformations*, Journal of the Royal Statistical Society, Series B 26.
- [4] Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1976): *Time Series Analysis, Forecasting and Control*, Third Edition, Holden-Day, Series G.
- [5] Dufour, JM., Pelletier D. (2002): *Linear methods for estimating VARMA models with a macroeconomic application*, Joint Statistical Meetings- Business & Economic Statistics.
- [6] Fishman, G.S. (1996): *Monte Carlo: concepts, algorithms and applications*, Springer-Verlag.
- [7] Gohberg, I., Lancaster, P., Rodman, L: *Matrix Polynomials*, Classics in Applied Mathematics 58, SIAM 2009.
- [8] Hannan, E.J. (1969): *The identification of Vector Mixed Autoregressive Moving Average Systems*, Biometrika, 56, 223-225.
- [9] L'Ecuyer, P., Blouin, F., Couture, R. (1993): *A search for good multiple recursive random number generators*, Transactions on Modeling and Computer Simulation.
- [10] Lütkepohl, H.(1993): *Introduction to Multiple Time Series Analysis*, Springer-Verlag.
- [11] Lütkepohl, H.(2005): *New introduction to Multiple Time Series Analysis*, Springer-Verlag.
- [12] Mankiw, N. Gregory (2008): *Principios de Economía*, Paraninfo.

- [13] Mishkin, Frederic S. (2008): *Moneda, banca y mercados financieros*, Pearson Addison Wesley.
- [14] Peña, D.(2005): *Análisis de series temporales*, Alianza Editorial.
- [15] Peña, D., Tiao, G., Tsay, R.(2001): *A Course in Time Series Analysis*, Wiley Series in Probability and Statistics.
- [16] Reinsel, G.(1993): *Elements of Multivariate Time Series Analysis*, Springer-Verlag.
- [17] Reinsel, G.C.(1997): *Elements of Multivariate Time series analysis* (2nd. ed.), New York, Wiley.
- [18] Sala-i-Martin, X. (2000): *Apuntes de Crecimiento Económico*, Antoni Bosch Editor.
- [19] Wold, H. (1938): *A Study in the Analysis of Stationary Time Series*, Almqvist & Wiksell.
- [20] Yozgatligil, C., Wei W (2009): *Representation of Multiplicative Seasonal Vector Autoregressive Moving Average Models*, The American Statistician, Vol. 63, No. 4.